EVALUATION OF *STAPHYLOCOCCUS AURUES* RNPA PROTEIN AS AN ANTIBACTERIAL TARGET

by

Lisha Ha

A Thesis

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Master of Science
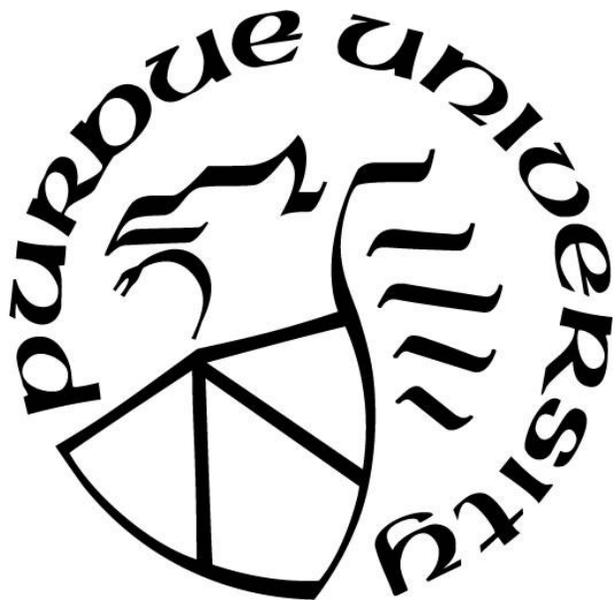
Department of Medicinal Chemistry and Molecular Pharmacology

West Lafayette, Indiana

August 2019

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Daniel P. Flaherty, Chair

> Department of Medicinal Chemistry and Molecular Pharmacology

Dr. V. Jo Davisson

> Department of Medicinal Chemistry and Molecular Pharmacology

Dr. Markus A. Lill

> Department of Medicinal Chemistry and Molecular Pharmacology

Dr. Mohamed Seleem

> Department of Comparative Pathobiology

Dr. Chittaranjan Das

> Department of Chemistry

Approved by:

> Dr. Andy Hudmon
>
> > Head of the Graduate Program

ACKNOWLEDGMENTS

I would like to begin by thanking my advisor, Dr. Daniel Flaherty for always being patient, understanding, and supportive. As my academic advisor, Dr. Flaherty has provided me with so many opportunities to learn new techniques and skills, and to brush up old knowledge. He would always be initiative in communicating with me, lead me through my research, provide me with constructive suggestions, and prompt me to think critically. He's frank with me about my strengths and weaknesses, and gives me time and help to improve. As a friend, he genuinely cares about my well-being and understand my difficulties as an international student.

I also want to thank my collaborators: Dr. Paul Dunman and Dr. Jennifer Colquhoun for providing the important preliminary data which laid the foundation of my research, and for their responsiveness whenever I needed help. Dr. Chittaranjan Das and his previous student John Hausman for the initial training in protein expression and purification and crystallography. Dr. Nicholas Noinaj, who introduced me into the field of crystallography as the instructor of my graduate crystallography course, him and his lab members, specifically Jeremy Baker helped me along the way of my crystallographic experiments. Dr. Mohamed Seleem and his graduate student Ahmed Hassan for the training and help on microbiological tests for my side projects.

I would also like to thank Dr. V. Jo Davisson, Dr. Markus Lill, Dr. Mohamed Seleem and Dr. Chittaranjan Das for serving as the members of my advisory committee.

My thanks also go to the present and past members of the Flaherty's lab, especially Aaron Krabill, Chad Hewitt and Jason Scott. They have created a caring, supportive and safe work environment where thoughts and experiences are freely exchanged and problems are solved together.

There are so many faculties and graduate students at Purdue that I have worked with, hung out with and made friends with. There are many friends outside West Lafayette in America or back home in China who had been there for me whenever I needed them. I can't name every of them because of the page limitation, but I'm truly grateful to have them in my life. I do want to thank Dr. Susan Holladay who's my TA instructor especially, as being my loving Purdue Mom.

Last but not least, my deepest gratitude goes to my beloved parents who support me through the thick and thin of my life, and kept me company as much as possible either in person or through the internet despite of the over 6,000-mile distance between us. I won't be who I am today without them, and there are no words that can express my appreciation for their devotion and sacrifices.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Author: Ha, Lisha. MS
Institution: Purdue University
Degree Received: August 2019
Title: Evaluation of *Staphylococcus aureus* RnpA Protein as an Antibacterial Target
Committee Chair: Dr. Daniel P. Flaherty

Staphylococcus aureus (*S. aureus*) is a Gram-positive pathogen that causes a wide range of infections in both hospitals and communities, of which the total mortality rate is higher than AIDS, tuberculosis, and viral hepatitis combined. The drug resistant *S. aureus* is a member of the "ESKAPE" pathogens that require immediate and sustained actions of novel method to combat. However, the current antimicrobial development against *S. aureus* is in stagnation, which underscores the urgent need for novel antimicrobial scaffolds and targets. *S. aureus* Ribonuclease P protein (RnpA) is an essential protein that plays important roles in both tRNA maturation and mRNA degradation pathways. The goal of this research was to evaluate RnpA as an antimicrobial target using biophysical methods. The crystal structures of wild-type RnpA in three different constructs were determined, among which the tag-free RnpA construct has a structural model of 2.0 Å resolution and R_{crys}/R_{free}= 0.214/0.234, and its crystals are reproducible. This crystal structure of tag-free *S. aureus* RnpA shows a globular representation with key structural motifs, including the "RNR" Ribonuclease P RNA binding region and a substrate binding central cleft, which shares high similarity to previously solved RnpA structures from other species despite of their low sequence identity. Meanwhile, in a screen of *S. aureus* RnpA mutants performed by our collaborator, RnpA^{P89A} was found lacking the mRNA degradation activity while retaining the tRNA maturation function, and causing defects in cell viability. We therefore studied this mutant using differential scanning fluorimetry, crystallography, and circular dichroism. It was shown that

RnpA^{P89A} is thermally less stable than wild-type RnpA by ~2.0 ˚C, but no secondary structural or 3D conformational differences were found between the two proteins. Although the mutant RnpA^{P89A} requires further characterization, the results of the studies in this thesis have begun to shed light on the relatively new role of *S. aureus* RnpA in mRNA degradation, and will serve as useful tools in future structure-based drug discovery for multi-drug resistant *S. aureus* treatment.

CHAPTER 1. INTRODUCTION

1.1 *Staphylococcus aureus* infection and resistance

Staphylococcus aureus (*S. aureus*) is a Gram-positive bacteria commonly found in the respiratory tract and on the skin of humans. Pathogenic *S. aureus* is able to cause a wide range of infections, including bacteremia, infective endocarditis, pleuropulmonary, osteoarticular, skin and soft tissue, and device-related infections[1]. The various types of *S. aureus* infections contribute to its high morbidity. *S. aureus* infection has been a leading cause of nosocomial infections for the past decade[1-3]. Recent data indicated that during 2011-2014, 12% of the hospital-associated infections (HAIs) reported in US were caused by *S. aureus*, which accounts for over 48,000 cases[3]. *S. aureus* infection also has a mortality rate of 10-67% depending on the type of infection, which is higher than AIDS, tuberculosis and viral hepatitis combined[4]. The most problematic manifestation of *S. aureus* infection is bacteremia, which has a case fatality rate of 15-50%[1]. Such danger of this microbial killer is further boosted by the fast-pace epidemic of drug-resistant *S. aureus* stains against all front line antibiotics worldwide[5-9]. The uncontrollable characteristic of drug-resistant *S. aureus* has made it a member of the notorious cohort of bacteria known as the "ESKAPE" pathogens as recognized by the Infectious Disease Society of America and the Center of Disease Control and Prevention (CDC)[10-12].

Two particular threatening resistant *S. aureus* strains are methicillin-resistant *Staphylococcus aureus* (MRSA) and vancomycin-resistant *Staphylococcus aureus* (VRSA)[13,14]. MRSA arose two years after the introduction of methicillin in 1960, and it has been shown to be resistant to all the β-lactams and cephalosporins[14,15]. The mechanism of methicillin resistance is related to the expression of a low affinity penicillin-binding protein (PBP2a) encoded by a MecA gene[15]. CDC

listed MRSA as a serious threat to public health that requires immediate and sustained actions[14]. Even if only cases of invasive infections were counted, it was estimated that MRSA had a morbidity of 80,000 and mortality of 11,000 in 2011[14]. MRSA has expanded its territory from hospitals to the community healthcare facilities. During 2012-2014, while 40% *S. aureus* HAIs in US were caused by MRSA[1,3], these cases accounted for only about 15% of the total 8,437 MRSA isolates[16]. The rest of the isolates came from the community, indicating the outbreak of community-acquired MRSA (CA-MRSA) in US, which was also observed worldwide[5,7,8,16,17]. The current regimen for MRSA treatment includes linezolid, daptomycin (IV), tigecycline (IV) and telavancin (IV) depending on the type of infection[18]. Yet adverse effects, resistant strains, and high drug costs are limiting their usage[18]. Vancomycin has been considered the last-resort treatment for drug-resistant *S. aureus*; however, with the emergence of VRSA, there is a real possibility that we will be left with few or no effective treatment against this pathogen. The currently stagnant antimicrobial development situation regarding drug resistant *S. aureus* underscores the need for novel antimicrobial scaffolds and targets. Below I introduce the Ribonuclease P protein (RnpA), which plays an essential role in RNA processing pathways, as a promising novel antimicrobial target for the treatment of *S. aureus*.

1.2 RnpA as a part of the Ribonuclease P holoenzyme

RnpA was first discovered as a part of the bacterial Ribonuclease P holoenzyme (RNase P) by Professor Sidney Altman in 1972[19]. RNase P is an endonuclease that exists in all three forms of life: Bacteria, Eukarya and Archaea. Most RNase P have been characterized as ribonucleases with both the RNA and the protein subunits[20,21]. Specifically, the architecture of bacterial RNase P is an RNA ribozyme (P RNA) that is about 400 nucleotides in length that is associated with one small protein (RnpA) of approximately 120 amino acids. It is well established that RNase P catalyzes

the maturation of 5' terminus of all transfer RNAs (tRNA)[22]. The tRNAs are initially synthesized as inactive precursor-tRNAs (ptRNA) with 5' leader and 3' trailer sequences. The ptRNAs go through extensive post-transcriptional modification to become matured tRNAs that participate in translation within cells. RNase P cleaves the 5' leader sequences at an early step in the biosynthetic "production chain" of tRNA[23,24]. (**Fig. 1.1**)
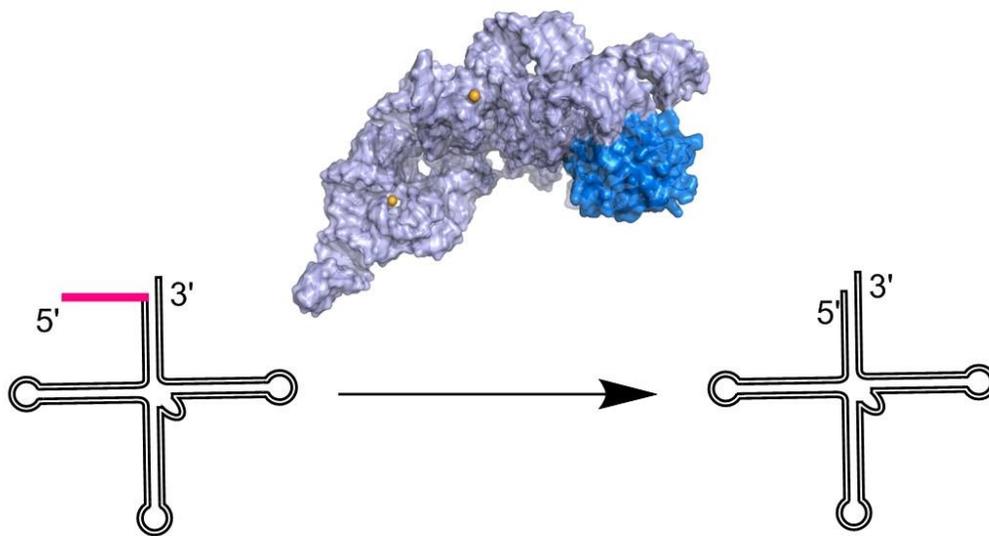
Figure 1.1 Ribonuclease P holoenzyme cleaves the 5' leader sequence of tRNA precursor. P RNA (light blue surface), RnpA (marine surface), 5'leader sequence (pink solid line).

Studies have shown that between the two subunits of bacterial RNase P, catalytic activity resides in the P RNA. Under *in vitro* high ionic conditions, P RNA alone is able to carry out cleavage of the 5' leader sequence from ptRNA[20]. However, both P RNA and RnpA are essential in physiologically relevant low salt conditions for bacterial cell viability[24,25]. The protein subunit, RnpA, plays an ancillary but crucial role in the functioning of RNase P. As mentioned above, RnpA is able to lower the dependence of P RNA on high condition ionic levels to exhibit catalytic activity *in vitro*. Previous biochemical studies have also shown that RnpA: 1) stabilizes the active conformation of P RNA[26-28], 2) improves binding specificity and affinity of RNase P to substrate

ptRNA[29,30], 3) helps with the release of product[27], and 4) improves the catalytic reaction efficiency[26,31,32]. These results were supported, and the understanding for bacterial RnpA was deepened, by the success of structural studies in the past two decades. So far, the RnpA structures of *Bacillus subtilis* (*B. subtillis*)[21], *S. aureus*[23] and *Thermotaga maritima* (*T. maritima*)[33,34] have been characterized by nuclear magnetic resonance (NMR) and/or X-ray crystallography. Reiter et al. successfully crystallized the *T. maritima* RNase P in complex with a post-cleavage tRNA[35] (**Fig. 1.2A**). This holoenzyme structure shows the orientation and contacts of P RNA, RnpA and tRNA. In particular, the highly conserved "RNR" motif region of RnpA forms electrostatic contacts with the phosphate backbone of the P RNA (**Fig. 1.2B**). These studies showed that there is conservation of bacterial RnpA secondary and tertiary structures across bacterial species despite the fact that RnpAs from different species share low sequence identities (23-51%)[23]. Bacterial RnpA is a globular protein that has an α-β sandwich fold. There is a β-sheet surface of four strands in the middle, and on one side of the β-sheet surface a helix to its N-terminal end lies perpendicular across the bottom of the strands. Together the helix and the β-sheet surface create a cleft that serves as the substrate binding site for 5' leader sequence of ptRNA (**Fig. 1.2C**). On the other side of the sheet are two other helices, and within one of them resides the highly conserved P RNA binding region, the "RNR" motif. This motif clusters basic residues that are positively charged in physiological conditions. Given that RNA is negatively charged at physiological pH, RnpA binds to P RNA with electrostatic forces and with nanomolar affinity[23,36]. The conservation of RnpA's structural and surface electrostatic elements across different bacterial species corroborates with the observations that bacterial RnpAs are interchangeable, despite of the low sequence identity[20,37]. For example, growth of *Escherichia coli* (*E. coli*) with *rnpA* gene (encodes *E. coli* RnpA) deleted

can be rescued by the complementation of a heterologous *rnpA* gene from *S. aureus, Pseudomonas aeruginosa,* or *B. subtillis*[38].

Previous studies have presented the essentiality of RnpA as a key component of bacterial RNase P. Targeting bacterial RnpA within the tRNA processing pathway is therefore a promising antimicrobial production strategy.
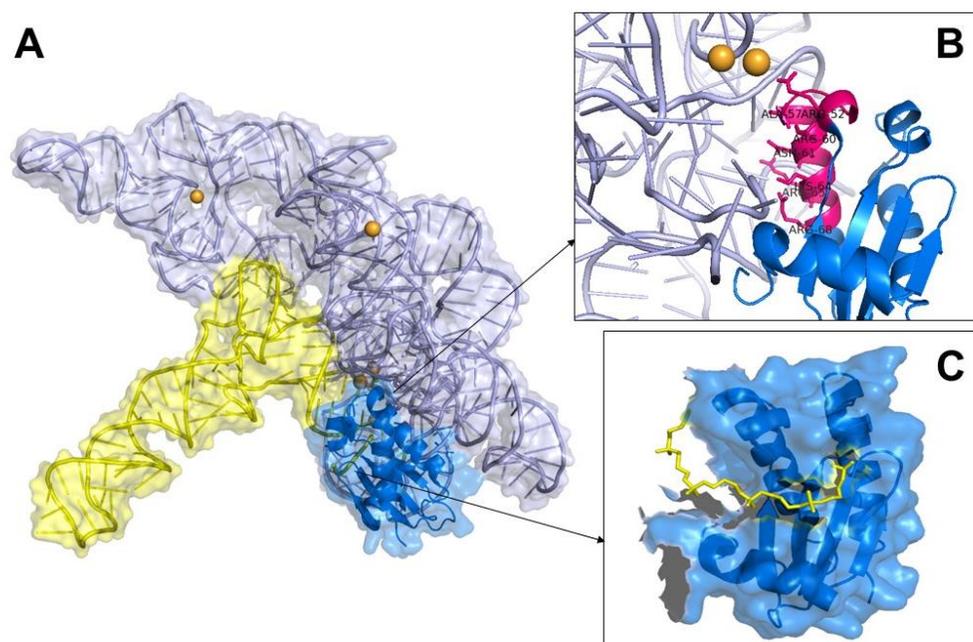
Figure 1.2 Bacterial Ribonuclease P holoenzyme. A) The overview of Ribonuclease P holoenzyme from *Thermotaga maritima* (PDB ID:3Q1R)[35]. P RNA (light blue), RnpA (marine), tRNA and soaked 5' leader sequence (yellow), Magnesium ions (orange). B) Close-up for the P RNA-RnpA binding region. The "RNR" box is shown in pink, and the reserved residues within this box are shown in sticks. C) Close-up for the RnpA-5' leader sequence binding region, RnpA "central cleft". Phosphate backbone (yellow sticks).

1.3   RnpA as a part of the messenger RNA degradosome

RNase P was postulated to have originated from a single catalytic RNA, and protein(s) were gradually recruited and started to play increasingly important roles within the complex as the system evolved[22,39]. Before 2008, it was reckoned that P RNA is ubiquitously essential and associated with the catalytic activity of RNase P in all three domains of life and that RnpA played

a role only in this pathway[22]. However, compelling studies then showed that in several species ribonuclease proteins were essential for the endonuclease activity of other RNA species. The Rossmanith group first revealed that human mitochondrial RNase P is composed of three protein subunits (mitochondrial RNase P proteins:MRPP1, MRPP2, and MRPP3) and no RNA subunit by proteomics, and the catalytic activity was successfully reconstructed *in vitro*[40]. 'Proteinaceous RNase P' (PRORP) was subsequently found in *Trypanosoma brucei*[41] and Arabidopsis[42]. More relevantly, the RnpA subunit of *Mycobacterium tuberculosis* (*M. tuberculosis*) has demonstrated RNase P activity in the *absence* of P RNA *in vitro*. The authors suggest that RnpA is able to achieve a different conformation and be imparted with the endonuclease activity of RNase P[43]. While bacterial RnpA exhibiting RNase P activity was the first time to be observed, it was actually not the first time that bacterial RnpA was found to have RNA processing capabilities. Increasing amounts of studies have shown that RnpA on its own was able to cleave messenger RNA (mRNA) and act as an essential component of a newly discovered mRNA degradosome. The first inclination came from the study by the Altman group, which showed aberrant RNase P activity in an *Escherichia coli* temperature-sensitive mutant strain A49 caused accumulation of at least 49 RNA species including 29 ptRNAs and 20 mRNAs at a lethal temperature of 43˚C[44,45]. It was shown that RNase P processed non-coding regions of polycistronic mRNA, and the generated mRNA cleavage products were subsequently degraded by nucleases in an *E. coli* mRNA degradosome, suggesting a new function for RNase P in mRNA degradation. Additionally, the study showed that some cleavage sites of the mRNAs were not the same as predicted RNase P cleavage sites, and the cleavage rate was slower for mRNA processing compared to ptRNA cleavage[45]. This raised the possibility that RnpA has endonuclease activity for mRNA cleavage and may require assistance from other proteins to get catalytic efficiency. This hypothesis was supported by studies from the

Dunman group. Repetitive experiments have demonstrated that the *S. aureus* RnpA was able to cleave mRNA *in vitro* in the absence of P RNA. *In cellulo* assays indicated that RnpA cellular depletion corresponded to increased mRNA accumulation in addition to ptRNA accumulation, as well as a loss of growth phenotype[46,47]. A group of German researcher has previously discovered an mRNA degradosome complex in *B. subtillis* and glycolytic enzymes were recognized as components in the RNA processing complex using bacteria two-hybrid analysis[48]. Inspired by this study, the Dunman group identified *S. aureus* RnpA as a central component of the *S. aureus* mRNA degradosome by bacterial two-hybrid assays, and the complete protein collective of *S. aureus* mRNA degradosome was proposed (**Fig. 1.3**)[49]. Dunman's work was corroborated with results from the Zang group's tandem affinity purification (TAP)/pull-down experiments in which they co-immunoprecipitated the complex in *S. aureus* and confirmed them by mass spectrometry[50]. These results suggested that some Gram-positive bacteria have an mRNA degradosome, and RnpA may function as an endonuclease[51] and/or as a nuclease regulator in the degradosome complex[50]. Bacterial bulk mRNA degradation is an essential cellular process that keeps the RNA system in homeostasis, regulates the expression of virulence factors, and controls the formation of biofilms[52]. This suggests the important role RnpA plays in the mRNA degradosome contributes to this protein's essentiality in the pathogen. Therefore, targeting bacterial RnpA-mediated mRNA degradation pathway may be a promising antimicrobial production strategy.
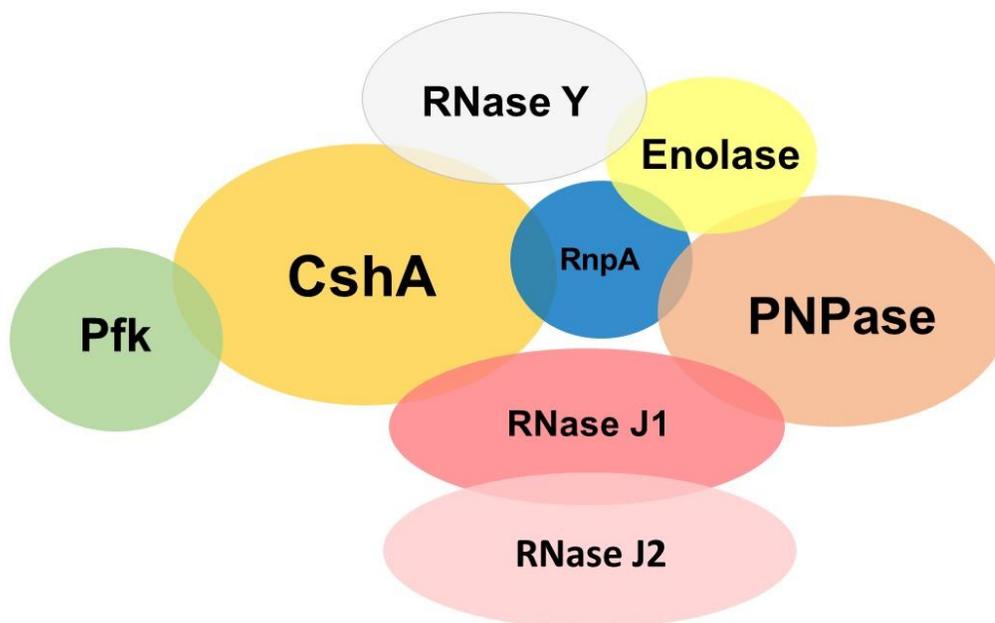
Figure 1.3 Proposed *S. aureus* mRNA degradosome complex. Overlaps indicate proposed interaction between the protein components, while size of overlaps don't necessary indicate the strength of interactions. CshA: DEAD box RNA helicase; Pfk: phosphofructokinase; PNPase: polynucleotide phosphorylase.

1.4 Previous study on bacterial RnpA as a drug target

Provided with continuous and substantial findings from biochemical and structural studies, bacterial RNase P has been of interest as a drug target. A brief review of previous drug discovery attempts to target the bacterial RNase P protein subunit (RnpA) might provide useful insights for future drug development.

Previous drug discovery with RnpA can be divided into 2 phases. Before 2011, bacterial RnpA was recognized merely as the non-catalytic subunit of RNase P, so early drug discovery regarding RnpA focused on interrupting functions of the RNase P complex as a whole. It was found that the amino-nucleoside puromycin (**Fig. 1.4A**) can inhibit RNase P, although its primary intracellular target is the ribosome[53]. There are two main classes of RNase P inhibitors under development: antisense oligonucleotides[54,55] and aminoglycosides[56-58]. Antisense oligonucleotides inhibit RNase

P by causing misfolding of the P RNA. These antisense oligonucleotides have targeted the 5'-m(cagccuacccgg)[54] sequence, the catalytic center sequence, or other essential sequences such as 5'- caagcagccuaccc[59]. Alternatively, antisense oligonucleotides combined with peptide nucleic acid like PNA-G-peptide (H₂N-KFFKFFKFFK-G-caagcagcctacc-CONH₂, peptide region underlined and nucleotide region not underlined)[60] induce degradation of the P RNA. Researchers are still working on resolving the issues of length optimization and position precision for these antisense inhibitors[61]. Aminoglycosides (and their arginine derivatives) function by binding with P RNA and interrupting P RNA-substrate/metal ion interactions that are essential for the catalytic activity. The antibiotic Neomycin B is an RNase P inhibitor (K_i 35 µM)[56], while its arginine derivative NeoR6 showed improved inhibitory activity (IC_{50} 0.1 µM)[57] (**Fig. 1.4B**). Same as puromycin, Neomycin B's primary target is the ribosomal complex and is not being optimized for inhibiting RNase P activity[61,62]. Aminoglycoside arginine derivatives are under development for higher selectivity and higher drug uptake by bacteria in order for them to become effective drug candidates[61]. In addition to the two classes of RNase P inhibitor being studied, some guanylhydrazone compounds, represented by MES 10608 (**Fig. 1.4C**) were discovered to inhibit *Neisseria gonorrhoeae* RNase P in a high-throughput screening performed by Message Pharmaceuticals, but no mechanism of actions or further research were published[61,63].

While the above examples identified P RNA binders that interrupt RNase P functions, there has been little activity to identify novel inhibitors of RnpA function. Researchers from Smith Kline Beecham (now GlaxoSmithKline) determined the NMR structure of *S. aureus* RnpA expecting it to be useful in antibiotic development against *S.aureus*[23], but no further studies have become known. Virtual screening with commercial ZINC libraries against *B. subtilis* RnpA crystal

structure identified a piperazine-like scaffold, but no more evaluation of this scaffold in biochemical/biological assays have been published[61].

The second phase of RnpA drug discovery was initiated in 2011 by the studies that suggested *S. aureus* RnpA's nuclease activity and that it is a component of mRNA degradosome. Three structurally distinct RnpA inhibitors, RNPA 1000, RNPA 2000 and RNPA 3000 (**Fig. 1.4D**) were identified from a high-throughput screen by the Dunman group. All three molecules inhibited RnpA's processing of mRNA in biochemical assays, while RNPA 2000 also demonstrated inhibition of tRNA maturation[46,47,64]. All three inhibitors showed moderate antimicrobial activity against drug-resistant *S. aureus* in broth microdilution assays with micro-molar efficacy[46,47]. The mRNA turnover within *S. aureus* was suppressed by all three inhibitors, indicating that the antimicrobial ability of these molecules was dependent on their inhibitory activity against RnpA[46,47,64]. Additionally, RNPA2000 and RNPA3000 were shown to confer no human cell cytotoxicity using MTT assay[64]. Further development of RNPA2000 has been made recently, as analogues were synthesized and structure-activity relationship (SAR) was analyzed in the effort to replace an undesirable furan moiety[65]. Although most of the compounds were found to be less potent than RNPA2000 and more toxic to human cells, two of them exhibited synergistic activity when used in combination with mupirocin, an anti-staphylococcal drug that targets isoleucyl-tRNA synthase that is on the same bio-processing pathway for tRNA maturation[66]. Because two inhibitors targeting two individual steps of the same pathway can have cooperative effects, the RNPA2000 analogues being synergistic with mupirocin suggests that they inhibit RNase P within *S.aureus*[65,67].

RnpA's essentiality in two bacterial RNA molecular pathways, combined with on target inhibition that effectively kills *S. aureus* suggest the potential of RnpA as an antimicrobial drug target.

Additionally, there has previously been interest from other groups and industry in pursuing this target. These taken together suggest further research with biophysical study for the drug discovery of this protein are warranted.
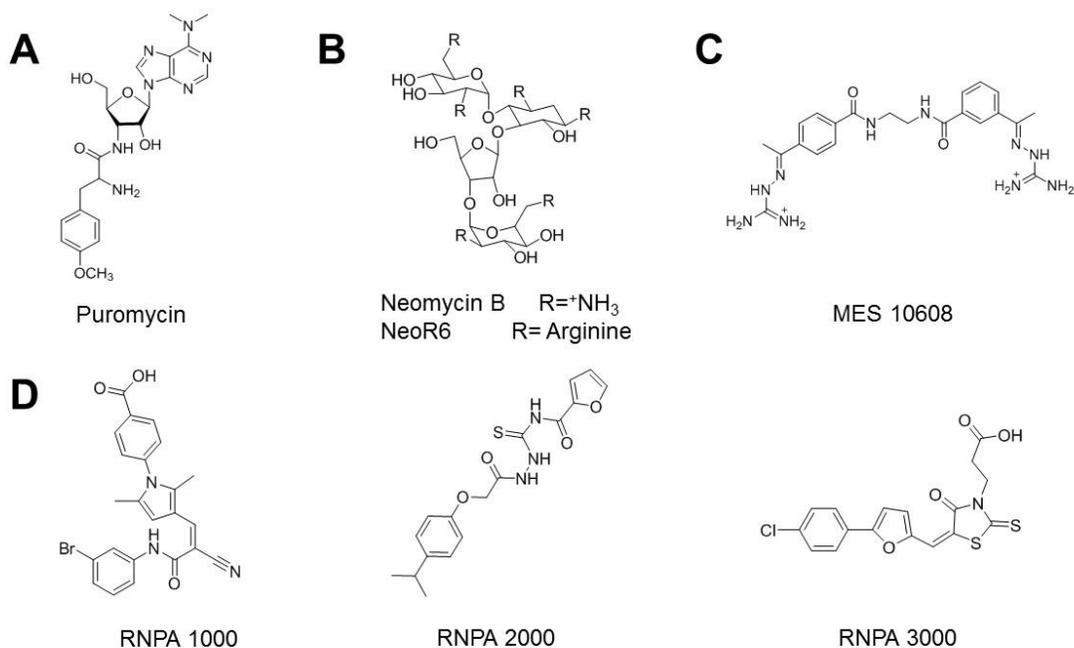
Figure 1.4 RNase P and RnpA inhibitors. A) Puromycin is an amino-nucleoside RNase P inhibitor. B) Neomycin is an aminoglycoside RNase P inhibitor; NeoR6 is an arginine derivative of aminoglycoside RNase P inhibitor[56-58]. C) MES 10608 is a guanylhydrazones RNase P inhibitor[61,63]. D) RnpA inhibitors[46,47,64,65].

1.5 Project summary

The prevalence of resistant *S. aureus* is becoming an increasingly concerning public health threat[9,68]. Existing drug targets and mechanism of actions are rendered less effective as the pathogen has acquired ways to circumvent the inhibitory strategies. The morbidity and mortality of MRSA and VRSA are increasing because of the lack of novel antimicrobial stratregies[15,69]. New antibiotic development is, therefore, in tremendous need for the intervention of *S. aureus* infections. Bacterial RnpA has elicited interest as a drug target based on the facts: 1) it is an essential protein for bacterial viability[24], 2) it is vastly different from eukaryotic RNase P proteins both structurally

and functionally[70], and 3) specifically for *S. aureus* RnpA, it was shown to play important roles in two different RNA processing machineries[27,47,49]. Inhibitor development for *S. aureus* RnpA targeting either the ptRNA maturation or the RnpA-mediated mRNA degradation process or both pathways simultaneously could lead to new staphylococcal treatment options and alleviate the medical burden. RnpA is an essential protein, and the tertiary structures and surface electrostatic elements of RnpA from phylogenetically diverse bacteria are remarkably conserved despite of low sequence identity[21,22,33], allowing the protein to be interchangeable among bacterial species[38]. Therefore, we hypothesize successful development of *S. aureus* RnpA inhibitors may have broad-spectrum potential for other bacterial pathogens of immediate or emerging healthcare concern.

In this thesis, *S. aureus* RnpA is further evaluated as a drug target by biophysical studies. Supported by gel-based enzymatic and protein alanine scanning studies performed by our collaborators, we were able to demonstrate that RnpA is a promising antimicrobial drug target. In the following chapters I will present my work on the biophysical characterization of wild-type *S. aureus* RnpA and an mRNA degradation defective mutant RnpA^{P89A} and describe the methodology I used to obtain the results presented herein.

References

1   Tong, S. Y. *et al*. Staphylococcus aureus infections: epidemiology, pathophysiology, clinical manifestations, and management. *Clin Microbiol Rev* **28**, 603-661, doi:10.1128/CMR.00134-14 (2015).

2   Sievert, D. M. *et al.* Antimicrobial-resistant pathogens associated with healthcare-associated infections: summary of data reported to the National Healthcare Safety Network at the Centers for Disease Control and Prevention, 2009-2010. *Infect Control Hosp Epidemiol* **34**, 1-14, doi:10.1086/668770 (2013).

3   Weiner, L. M. *et al.* Antimicrobial-Resistant Pathogens Associated With Healthcare-Associated Infections: Summary of Data Reported to the National Healthcare Safety Network at the Centers for Disease Control and Prevention, 2011-2014. *Infect Control Hosp Epidemiol* **37**, 1288-1301, doi:10.1017/ice.2016.174 (2016).

4   van Hal, S. J. *et al.* Predictors of mortality in Staphylococcus aureus Bacteremia. *Clin Microbiol Rev* **25**, 362-386, doi:10.1128/CMR.05022-11 (2012).

5   Chen, C. J. & Huang, Y. C. New epidemiology of Staphylococcus aureus infection in Asia. *Clin Microbiol Infect* **20**, 605-623, doi:10.1111/1469-0691.12705 (2014).

6   Rasigade, J. P., Dumitrescu, O. & Lina, G. New epidemiology of Staphylococcus aureus infections. *Clin Microbiol Infect* **20**, 587-588, doi:10.1111/1469-0691.12718 (2014).

7   Schaumburg, F., Alabi, A. S., Peters, G. & Becker, K. New epidemiology of Staphylococcus aureus infection in Africa. *Clin Microbiol Infect* **20**, 589-596, doi:10.1111/1469-0691.12690 (2014).

8   Tokajian, S. New epidemiology of Staphylococcus aureus infections in the Middle East. *Clin Microbiol Infect* **20**, 624-628, doi:10.1111/1469-0691.12691 (2014).

9   Fair, R. J. & Tor, Y. Antibiotics and bacterial resistance in the 21st century. *Perspect Medicin Chem* **6**, 25-64, doi:10.4137/PMC.S14459 (2014).

10   Boucher, H. W. *et al.* Bad bugs, no drugs: no ESKAPE! An update from the Infectious Diseases Society of America. *Clin Infect Dis* **48**, 1-12, doi:10.1086/595011 (2009).

11   Santajit, S. & Indrawattana, N. Mechanisms of Antimicrobial Resistance in ESKAPE Pathogens. *Biomed Res Int* **2016**, 2475067, doi:10.1155/2016/2475067 (2016).

12   Rice, L. B. Federal funding for the study of antimicrobial resistance in nosocomial pathogens: no ESKAPE. *J Infect Dis* **197**, 1079-1081, doi:10.1086/533452 (2008).

13   Chang, S. *et al.* Infection with vancomycin-resistant Staphylococcus aureus containing the vanA resistance gene. **348**, 1342-1347 (2003).

14   ANTIBIOTIC RESISTANCE THREATS in the United States, 2013. (U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, 2013).

15   Lodise, T. P., Jr. & McKinnon, P. S. Burden of methicillin-resistant Staphylococcus aureus: focus on clinical and economic outcomes. *Pharmacotherapy* **27**, 1001-1012, doi:10.1592/phco.27.7.1001 (2007).

16   Sader, H. S., Mendes, R. E., Jones, R. N. & Flamm, R. K. Antimicrobial susceptibility patterns of community- and hospital-acquired methicillin-resistant Staphylococcus aureus from United States Hospitals: results from the AWARE Ceftaroline Surveillance Program (2012-2014). *Diagn Microbiol Infect Dis* **86**, 76-79, doi:10.1016/j.diagmicrobio.2016.06.017 (2016).

17   Williamson, D. A., Coombs, G. W. & Nimmo, G. R. Staphylococcus aureus 'Down Under': contemporary epidemiology of S. aureus in Australia, New Zealand, and the South West Pacific. *Clin Microbiol Infect* **20**, 597-604, doi:10.1111/1469-0691.12702 (2014).

18   Rodvold, K. A. & McConeghy, K. W. Methicillin-resistant Staphylococcus aureus therapy: past, present, and future. *Clin Infect Dis* **58 Suppl 1**, S20-27, doi:10.1093/cid/cit614 (2014).

19   Robertson, H. D., Altman, S. & Smith, J. D. Purification and properties of a specific Escherichia coli ribonuclease which cleaves a tyrosine transfer ribonucleic acid presursor. *The Journal of biological chemistry* **247**, 5243-5251 (1972).

20   Guerriertakada, C., Gardiner, K., Marsh, T., Pace, N. & Altman, S. The Rna Moiety of Ribonuclease-P Is the Catalytic Subunit of the Enzyme. *Cell* **35**, 849-857 (1983).

21   Stams, T., Niranjanakumari, S., Fierke, C. A. & Christianson, D. W. Ribonuclease P protein structure: evolutionary origins in the translational apparatus. *Science* **280**, 752-755 (1998).

22   Evans, D., Marquez, S. M. & Pace, N. R. RNase P: interface of the RNA and protein worlds. *Trends Biochem Sci* **31**, 333-341, doi:10.1016/j.tibs.2006.04.007 (2006).

23   Spitzfaden, C. *et al.* The structure of ribonuclease P protein from Staphylococcus aureus reveals a unique binding site for single-stranded RNA. *J Mol Biol* **295**, 105-115, doi:10.1006/jmbi.1999.3341 (2000).

24   Kazantsev, A. V. & Pace, N. R. Bacterial RNase P: a new view of an ancient enzyme. *Nat Rev Microbiol* **4**, 729-740, doi:10.1038/nrmicro1491 (2006).

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

26   Buck, A. H., Dalby, A. B., Poole, A. W., Kazantsev, A. V. & Pace, N. R. Protein activation of a ribozyme: the role of bacterial RNase P protein. *EMBO J* **24**, 3360-3368, doi:10.1038/sj.emboj.7600805 (2005).

27   Reich, C., Olsen, G. J., Pace, B. & Pace, N. R. Role of the protein moiety of ribonuclease P, a ribonucleoprotein enzyme. *Science* **239**, 178-181 (1988).

28   Kim, J. J., Kilani, A. F., Zhan, X., Altman, S. & Liu, F. The protein cofactor allows the sequence of an RNase P ribozyme to diversify by maintaining the catalytically active structure of the enzyme. *RNA* **3**, 613-623 (1997).

29   Koutmou, K. S. *et al.* Protein-precursor tRNA contact leads to sequence-specific recognition of 5' leaders by bacterial ribonuclease P. *J Mol Biol* **396**, 195-208, doi:10.1016/j.jmb.2009.11.039 (2010).

30   Kurz, J. C., Niranjanakumari, S. & Fierke, C. A. Protein component of Bacillus subtilis RNase P specifically enhances the affinity for precursor-tRNAAsp. *Biochemistry* **37**, 2393-2400, doi:10.1021/bi972530m (1998).

31   Crary, S. M., Niranjanakumari, S. & Fierke, C. A. The protein component of Bacillus subtilis ribonuclease P increases catalytic efficiency by enhancing interactions with the 5' leader sequence of pre-tRNAAsp. *Biochemistry* **37**, 9409-9416, doi:10.1021/bi980613c (1998).

32   Kurz, J. C. & Fierke, C. A. The affinity of magnesium binding sites in the Bacillus subtilis RNase P x pre-tRNA complex is enhanced by the protein subunit. *Biochemistry* **41**, 9545-9558 (2002).

33   Kazantsev, A. V. *et al.* High-resolution structure of RNase P protein from Thermotoga maritima. *Proc Natl Acad Sci U S A* **100**, 7497-7502, doi:10.1073/pnas.0932597100 (2003).

34   Zeng, D., Brown, B. P., Voehler, M. W., Cai, S. & Reiter, N. J. NMR resonance assignments of RNase P protein from Thermotoga maritima. *Biomol NMR Assign* **12**, 183-187, doi:10.1007/s12104-018-9806-7 (2018).

35   Reiter, N. J. *et al.* Structure of a bacterial ribonuclease P holoenzyme in complex with tRNA. *Nature* **468**, 784-789, doi:10.1038/nature09516 (2010).

36   Beebe, J. A. & Fierke, C. A. A kinetic mechanism for cleavage of precursor tRNA(Asp) catalyzed by the RNA component of Bacillus subtilis ribonuclease P. *Biochemistry* **33**, 10294-10304 (1994).

37   Gosringer, M. & Hartmann, R. K. Function of heterologous and truncated RNase P proteins in Bacillus subtilis. *Mol Microbiol* **66**, 801-813, doi:10.1111/j.1365-2958.2007.05962.x (2007).

38   Turrini, P. C., Loveland, J. L. & Dorit, R. L. By any other name: heterologous replacement of the Escherichia coli RNase P protein subunit has in vivo fitness consequences. *PLoS One* **7**, e32456, doi:10.1371/journal.pone.0032456 (2012).

39   Hartmann, E. & Hartmann, R. K. The enigma of ribonuclease P evolution. *Trends Genet* **19**, 561-569, doi:10.1016/j.tig.2003.08.007 (2003).

40   Holzmann, J. *et al.* RNase P without RNA: Identification and Functional Reconstitution of the Human Mitochondrial tRNA Processing Enzyme. *Cell* **135**, 462-474 (2008).

41   Taschner, A. *et al.* Nuclear RNase P of Trypanosoma brucei: a single protein in place of the multicomponent RNA-protein complex. *Cell Rep* **2**, 19-25, doi:10.1016/j.celrep.2012.05.021 (2012).

42   Gobert, A. *et al.* A single Arabidopsis organellar protein has RNase P activity. *Nat Struct Mol Biol* **17**, 740-U113 (2010).

43   Singh, A., Ubaid-ullah, S., Ramteke, A. K. & Batra, J. K. Influence of Conformation of M. tuberculosis RNase P Protein Subunit on Its Function. *PLoS One* **11**, e0153798, doi:10.1371/journal.pone.0153798 (2016).

44   Li, Y., Cole, K. & Altman, S. The effect of a single, temperature-sensitive mutation on global gene expression in Escherichia coli. *RNA* **9**, 518-532 (2003).

45   Li, Y. & Altman, S. A specific endoribonuclease, RNase P, affects gene expression of polycistronic operon mRNAs. *Proc Natl Acad Sci U S A* **100**, 13213-13218, doi:10.1073/pnas.2235589100 (2003).

46   Eidem, T. M. *et al.* Small-molecule inhibitors of Staphylococcus aureus RnpA-mediated RNA turnover and tRNA processing. *Antimicrob Agents Chemother* **59**, 2016-2028, doi:10.1128/AAC.04352-14 (2015).

47   Olson, P. D. *et al.* Small molecule inhibitors of Staphylococcus aureus RnpA alter cellular mRNA turnover, exhibit antimicrobial activity, and attenuate pathogenesis. *PLoS Pathog* **7**, e1001287, doi:10.1371/journal.ppat.1001287 (2011).

48   Commichau, F. M. *et al.* Novel activities of glycolytic enzymes in Bacillus subtilis: interactions with essential proteins involved in mRNA processing. *Mol Cell Proteomics* **8**, 1350-1360, doi:10.1074/mcp.M800546-MCP200 (2009).

49   Roux, C. M., DeMuth, J. P. & Dunman, P. M. Characterization of components of the Staphylococcus aureus mRNA degradosome holoenzyme-like complex. *J Bacteriol* **193**, 5520-5526, doi:10.1128/JB.05485-11 (2011).

50   Wang, X. *et al.* Enolase binds to RnpA in competition with PNPase in Staphylococcus aureus. *FEBS Lett* **591**, 3523-3535, doi:10.1002/1873-3468.12859 (2017).

51   Shin, E. *et al.* Identification of amino acid residues in the catalytic domain of RNase E essential for survival of Escherichia coli: functional analysis of DNase I subdomain. *Genetics* **179**, 1871-1879, doi:10.1534/genetics.108.088492 (2008).

52   Gorna, M. W., Carpousis, A. J. & Luisi, B. F. From conformational chaos to robust regulation: the structure and function of the multi-enzyme RNA degradosome. *Q Rev Biophys* **45**, 105-145, doi:10.1017/S003358351100014X (2012).

53   Vioque, A. Protein synthesis inhibitors and catalytic RNA. Effect of puromycin on tRNA precursor processing by the RNA component of Escherichia coli RNase P. *FEBS Lett* **246**, 137-139 (1989).

54   Childs, J. L., Poole, A. W. & Turner, D. H. Inhibition of Escherichia coli RNase P by oligonucleotide directed misfolding of RNA. *RNA* **9**, 1437-1445 (2003).

55   Willkomm, D. K. *et al.* Evaluation of bacterial RNase P RNA as a drug target. *Chembiochem* **4**, 1041-1048, doi:10.1002/cbic.200300674 (2003).

56   Mikkelsen, N. E., Brannvall, M., Virtanen, A. & Kirsebom, L. A. Inhibition of RNase P RNA cleavage by aminoglycosides. *Proc Natl Acad Sci U S A* **96**, 6155-6160 (1999).

57   Eubank, T. D. *et al.* Inhibition of bacterial RNase P by aminoglycoside-arginine conjugates. *FEBS Lett* **511**, 107-112 (2002).

58   Berchanski, A. & Lapidot, A. Bacterial RNase P RNA is a drug target for aminoglycoside-arginine conjugates. *Bioconjug Chem* **19**, 1896-1906, doi:10.1021/bc800191u (2008).

59   Gruegelsiepe, H., Willkomm, D. K., Goudinakis, O. & Hartmann, R. K. Antisense inhibition of Escherichia coli RNase P RNA: mechanistic aspects. *Chembiochem* **4**, 1049-1056, doi:10.1002/cbic.200300675 (2003).

60   Gruegelsiepe, H., Brandt, O. & Hartmann, R. K. Antisense inhibition of RNase P: mechanistic aspects and application to live bacteria. *The Journal of biological chemistry* **281**, 30613-30620, doi:10.1074/jbc.M603346200 (2006).

61   Willkomm, D. K. P., P.; Reuter, K.;Klebe, G.; Hartmann, R.K. in *Ribonuclease P Protein Reviews* (ed Sidney Altman Fengyong Liu) Ch. 13, 235-256 (Springer-Verlag New York, 2010).

62   Bosscha, M. I., van Dissel, J. T., Kuijper, E. J., Swart, W. & Jager, M. J. The efficacy and safety of topical polymyxin B, neomycin and gramicidin for treatment of presumed bacterial corneal ulceration. *Br J Ophthalmol* **88**, 25-28 (2004).

63   Giordano, T. S., M.A.; Rao , S.J. Inhibitors of RNase P proteins as antibacterial compounds. (2006).

64   Dunman, P. M. O., P.D.; Childers, W. Use of the RNase Inhibitor RNPA-3000 for Treating Staphylococcus Infections. (2017).

65   Lounsbury, N. *et al.* Novel inhibitors of Staphylococcus aureus RnpA that synergize with mupirocin. *Bioorg Med Chem Lett* **28**, 1127-1131, doi:10.1016/j.bmcl.2018.01.022 (2018).

66   Hughes, J. & Mellows, G. Interaction of pseudomonic acid A with Escherichia coli B isoleucyl-tRNA synthetase. *Biochem J* **191**, 209-219 (1980).

67   Potter, V. R. & Simonson, H. Sequential blocking of metabolic pathways in vivo. *Proc Soc Exp Biol Med* **76**, 41-46 (1951).

68   McKenna, M. Antibiotic resistance: the last resort. *Nature* **499**, 394-396, doi:10.1038/499394a (2013).

69   Staphylococcus aureus resistant to vancomycin--United States, 2002. Report No. 0149-2195 (Print) 0149-2195 (Linking), 565-567 (Centers for Disease Control Prevention, 2002).

70   *Ribonuclease P*. 1 edn, (Springer-Verlag New York, 2010).

CHAPTER 2.   CRYSTALLIZATION OF *STAPHYLOCOCCUS AUREUS* WILD-TYPE RNPA

2.1   Introduction

Most of bacterial Ribonuclease P holoenzymes (RNase P) are constructed as ribonucleoprotein complexes. Each complex contains a catalytic RNA ribozyme component (P RNA) and an auxiliary, but essential, protein subunit (RnpA)[1]. It has been demonstrated in some bacterial species, including *Staphylococcus aureus*, that RnpA is involved in two RNA processing pathways: tRNA maturation and mRNA degradation. The tRNA maturation pathway has been well-studied since the discovery of RNase P. This holoenzyme cleaves the 5'-leader sequence of precursor-tRNA (ptRNA) to generate a mature 5' end of tRNA[1-3]. Whereas, RnpA's role in the mRNA degradosome was discovered recently[4]. Because of the essentiality of RnpA in RNA processing pathways of *S. aureus* and the significant differences between eukaryotic and prokaryotic RNA processing machinery, it is hypothesized that RnpA represents a novel antimicrobial therapeutic drug target[5-10].

Structural studies of wild-type RnpA from two non-pathogenic bacterial species have been previously reported. The X-ray crystal structure of RnpA from *Bacillus subtilis* (PDB ID: 1A6F) first established the fundamental understanding of bacterial RnpA[11]. The crystal structures of RnpA (PDB ID: 1NZ0)[12] and RNase P holoenzyme (PDB IDs: 3Q1Q, 3Q1R)[13], and also the recently reported NMR structure of *Thermotoga maritima* RnpA[14], provided more information on RnpA interactions with RNAs. Regarding *S. aureus* RnpA, the NMR structure had been reported by a research group from Smith Kline Beecham (now GlaxoSmithKline) in 2000 and deposited into the protein databank (PDB ID: 1D6T)[15]. However, the authors failed to deposit either the ^{15}N-^{1}H heteronuclear single quantum coherence (HSQC) 2-dimensional (2D) NMR spectra, or the

protein NMR backbone assignments for this structure in any bio-NMR data repository. The lack of 2D NMR data associated with the *S. aureus* RnpA structure significantly limits the ability to leverage the structure for drug-discovery purposes. Experiments to recollect the spectra and recalculation of NOE for structure determination will be required for the purpose of using this structure to map ligand-binding sites via 2D NMR chemical shift perturbation. Additionally, the NMR is an ensemble of 20 conformers selected out of 100 calculated conformers that best satisfy the through-bond and NOE-based restraints obtained by 2D and 3D spectroscopies. Atoms of each conformers showed displacement from the atomic position of the average structure, especially for the more dynamic C- and N-termini and loop regions. While atomic displacements reflect important solution-state dynamics, it also brings higher uncertainties in atomic positions of the average structure. Therefore, we sought to determine the *S. aureus* RnpA crystal structure in support of structure-based drug discovery.

Reported herein are the crystallization studies of wild-type *S. aureus* RnpA. Along this structural characterization process several strategies were undertaken to successfully obtain RnpA crystals. First, a maltose-binding protein (MBP)-tagged RnpA construct, in which the MBP tag is engineered to drive crystallization, was utilized. However, due to difficulties in purification of MBP-RnpA, this construct failed to provide reproducible RnpA crystals. I then moved onto the 6x-Histidine (His)-tagged RnpA construct. But because of the large size of the tag, which accounted for ¼ of the full length of the fusion protein, the crystals obtained from this construct diffracted to a resolution of 3.0 Å and were also difficult to reproduce. Seeking to increase resolution and reproducibility I engineered a cleavable His-tag into the RnpA construct to produce a tag-free version of RnpA. Crystals from this construct diffracted to a resolution of 2.0 Å and were reproducible in the original crystallization conditions. The crystal structure was deposited in

the PDB (PDB ID: 6D1R)[16] providing a reliable method for obtaining structural data in support of rational design of novel RnpA inhibitors.

2.2   Materials and Methods

2.2.1   Vector constructs and cloning

The recombinant constructs for MBP RnpA (pETXM1::*rnpA*) and His-RnpA (pET30 Ek-LIC::*rnpA*) were designed by the Das group and the Dunman group, respectively. Both vectors were produced and provided to us by Jennifer Colquhoun from the Dunman group. The recombinant vector for tag-free RnpA was designed and cloned in our lab with the help of Prof. Nicholas Noinaj in the steps described as follows.

Primer Design: The *S. aureus rnpA* gene sequence coding for residues 4–117 was input into NEBcutter V2.0[17], and restriction enzymes NcoI (C˅CATG˄G) and XhoI (C˅TCGA˄G) were selected from the 0 cutter list, which ensured that the restriction enzymes would not cut within the *rnpA* gene sequence. Forward primer was designed as AAA<u>CCATGG</u>GATGGGAAAAAGCTTACCGA containing the NcoI restriction site (underlined). Reverse primer was designed as TTT<u>CTCGAG</u>TTACTTAATCTTTTT containing the XhoI restriction site (underlined) with the help of online tool Reverse Compliment[18].

Target Gene Amplification: The template plasmid pET30 Ek-LIC::*rnpA* was prepared with a Hi-Speed Mini Plasmid Kit (IBI Scientific). The *rnpA* gene was then amplified by polymerase chain reaction (PCR) using the designed primers, Phusion® High-Fidelity DNA Polymerase (New England BioLabs, NEB), Deoxynucleotide (dNTP) Solution Mix (NEB) and required buffers. PCR process consists of a 30 second-initial denaturation at 98˚C, 40 cycles of 10 second-denaturation

at 98°C, 30 second-annealing at 55°C, 15 second-extension at 72°C, and a 5 minute-final extension at 72°C.

Vector Construction: The PCR product was purified with a 1% Agarose gel, extracted with a Gel/PCR DNA Fragments Extraction Kit (IBI Scientific) and digested with restriction enzymes NcoI and XhoI. A solution of pHis-Parallel2 vector was also digested with the same enzymes. The digested *rnpA* gene and target vector were purified by gel, extracted and ligated with T4 DNA ligase (NEB) to yield the vector pHis-Parallel2*::rnpA* encoding a cleavable 6x-His fusion tag (underlined) and a tobacco etch virus (TEV) protease-cleavage site (#) (MSYY<u>HHHHHH</u>DYDIPTTENLYFQ#GAMGW-RnpA). The sequence was verified by sequencing analysis at the Purdue Genomics Core Facility.

2.2.2 Protein production

The MBP-RnpA was produced with recombinant plasmid pETXM1::*rnpA*, the His-RnpA was produced with recombinant plasmid pET30 Ek-LIC::*rnpA*, and the tag-free RnpA was produced with recombinant plasmid pHis2-Parallel::*rnpA*. General expression and purification protocols were the same for all three proteins, while the tag-free RnpA required a TEV cleavage and purification step before final purification.

Protein Expression: The plasmid DNA was transformed into competent *E. coli* BL21 (DE3) Rosetta cells (Novagen) by 45 second heat shock at 42 °C. Cells were subsequently spread on an agar plate with 50 µg/mL kanamycin (for MBP-RnpA and His-RnpA) or 100 µg/mL ampicillin (for tag-free RnpA) and incubated at 37 °C overnight. A single colony was taken from the plate and cultured overnight in 5 mL of Luria Broth (LB) medium containing the respective antibiotic with shaking at 225 rev/min. Overnight culture was pelleted the next morning by centrifugation at 3724 g for 15 min with a benchtop Allegra X-12R Beckman Coulter centrifuge then re-suspended

with 6 mL of fresh LB. 1 mL of bacteria LB solution was aliquoted into each of the six-1 L amp-LB media or kan-LB and incubated at 37˚C with shaking at 225 rev/min until OD_{600} reached 0.6. The inoculation was then induced with 300 mM isopropyl-*D*-1-thiogalactopyranoside (IPTG) and incubated at 37˚C for 4 h. Cells were pelleted by centrifugation at 3724g for 15 min and stored at -80 ˚C. Specifically, LB media used to prepare MBP-RnpA (pETXM1*::rnpA*) contains 0.2% D-Glucose (w/v).

Protein Initial Purification: For MBP-RnpA, cell pellet was re-suspended in buffer A (50 mM sodium phosphate buffer, 300 mM sodium chloride, pH 7.4) and lysed by high pressure using a French Press. Lysate was centrifuged at 100,000 rpm for 1 h on a Beckman Coulter ultracentrifuge to pellet cell debris, and supernatant was diluted with equivalent volume of column buffer (20 mM Tris-HCl, 200 mM NaCl and 1 mM EDTA). The supernatant was then purified on an amylose column (NEB). Proteins eluted with 10mM D-maltose in column buffer.

For His-RnpA and the final tag-free RnpA (fused with a His-tag at this step), cell pellet was re-suspended in buffer A (50 mM sodium phosphate buffer, 300 mM sodium chloride, pH 7.4) and lysed by sonication. Lysate was centrifuged at 14,000 g for 1 h to pellet cell debris, and supernatant with soluble protein of interest was purified on a 5 mL Nickel-nitrilotriacetic acid (Ni^{2+}-NTA) affinity column (GE Healthcare Bio-Sciences) with an ÄKTA-fast protein liquid chromatography system (FPLC) (GE Healthcare Bio-Sciences). Proteins eluted in a single peak with a linear imidazole gradient (0–500 mM) provided by mixing buffer A and buffer B (1 M imidazole, 50 mM sodium phosphate buffer, 300 mM sodium chloride, pH 7.4). Fractions under the peak were analyzed by Coomassie-stained SDS–PAGE before being pooled together.

TEV Cleavage and Purification: Specifically, for the tag-free RnpA, the initial purified protein bears a cleavable 6x-His-tag. Complete removal of the tag was achieved by incubating the protein

with TEV (200:1, v:v), 0.5 mM EDTA and 1 mM DTT for 48 h at 4°C in the elution buffer. The cleavage product was then purified on the Ni^{2+}-NTA affinity column in the same manner as the supernatant.

Final Purification: Protein product after initial purification or cleavage, was further purified by size-exclusion chromatography (SEC) with a HiPrep Sephacryl S200 column (GE Healthcare Bio-Sciences) on the ÄKTA-FPLC system to remove aggregates. The MBP-RnpA was eluted with 50 mM Tris pH 7.4, 50 mM NaCl, 1 mM DTT and 5 mM Maltose, while the His-RnpA and tag-free RnpA were eluted in buffer C (20 mM Tris pH 7.4, 200 mM sodium chloride). Fractions were analyzed by Coomassie stained SDS–PAGE, and those with >95% purity were pooled, concentrated to 1 mg/mL, and stored at 4 K for short-term use or flushed froze with liquid nitrogen and stored -80°C for long-term reservation. The macromolecule information was summarized in **Table 2.1**.

2.2.3   Crystallization

Proteins were thawed and kept on ice for over 2 hrs to equilibrate. Pre-crystallization tests were then performed with PCT™ kit (Hampton Research) to determine the broad screening protein concentrations.

MBP-RnpA: The purified protein sample was concentrated by centrifugation at 14,000g in an Amicon Ultra centrifugal filter (cutoff MW 3 kDa) to 10 mg/mL. Broad screening was set up using commercial crystallization screens available to our laboratory, Index (Hampton Research), ComPAS Suite (Qiagen), MemGold™ HT-96, MemGold2™ HT-96, PACT premier™ HT-96 (Molecular Dimensions)[19], on a TTP LabTech Mosquito crystallization robot using the hanging-drop vapor-diffusion method at room temperature. Once initial hit conditions were identified optimization trays of crystallization conditions from broad screen were set up with variation in

concentrations of salts and precipitants on a TTP LabTech dragonfly discovery liquid handling robot and the Mosquito crystallization robot using the hanging-drop vapor-diffusion method at room temperature. Stocks of reservoir solution components were all purchased from Hampton Research except that a 50% polyethylene glycol 4,000 solution was made manually from solids (Millipore). Seven 24-well crystallization trays were subsequently set up for crystallization conditions from optimization trays with finer variation in concentrations of salts and precipitant. The crystals from four conditions were collected, immerged in 20% ethylene glycol solution, and flash-cooled by plunging into liquid nitrogen. Detailed crystallization information for broad screening, optimization screening and manual optimization tray was summarized in **Table 2.2**.

His-RnpA: The purified protein sample was concentrated by centrifugation at 14,000g in an Amicon Ultra centrifugal filter (cutoff MW 3 kDa) to 22 mg/mL. Broad screening was set up using commercial crystallization screens available to our laboratory, PEG/Ion HT (Hampton Research), MCSG-1, MCSG-2, MCSG-3 (Anatrace)[19], on a TTP LabTech Mosquito Crystal crystallization robot using the hanging-drop vapor-diffusion method at room temperature. Once initial crystal screening hit conditions were identified, optimization trays of three crystallization conditions from broad screen were set up manually with variation in concentrations of salt and precipitants using the hanging-drop vapor-diffusion method at room temperature. Stock solutions of reservoir solution components were all purchased from Hampton Research. A further optimization tray was subsequently set up for one crystallization condition with variation in ratio of protein and reservoir solution in the drop. Crystals from one condition were flash-cooled by direct plunging into liquid nitrogen with no cryo-protectant. Detailed crystallization information was summarized in **Table 2.2**.

Tag-free RnpA: The purified recombinant protein sample was concentrated by centrifugation at 14, 000 g in an Amicon Ultra centrifugal filter (cutoff MW 3 kDa) to 12 mg/mL, and screened using commercial crystallization screens available to our laboratory, MemGold[TM] HT-96 and MemGold2[TM] HT-96 (Molecular Dimensions)[19], on a TTP LabTech Mosquito Crystal crystallization robot using the hanging-drop vapor-diffusion method at room temperature. Crystals produced from the screens were flash-cooled by direct plunging into liquid nitrogen with no extra cryo-protectants. Crystallization information is summarized in **Table 2.2**.

2.2.4   Data collection and structure determination

All the X-ray Crystallography data were collected at the Argonne National Laboratory (Chicago, IL). Data processing was done with HKL-2000[20]. All three *S. aureus* RnpA structures were solved by molecular replacement with Phaser[21] in PHENIX using the average NMR structure of RnpA as the search model (PDB ID: 1D6T). Structural refinement was performed with PHENIX[22] and Coot[23]. The coordinate of the tag free version of RnpA has been deposited and detailed data collection and refinement information were provided in **Table 2.3**. Structure figures were prepared using the PyMOL molecular graphics system (Schrodinger, LLC) and Coot.

2.2.5   Structure and sequence alignments

Sequence alignment of *S. aureus* RnpA with RnpAs from eight other bacteria (Gram-positive: *B. subtillis, T.maritima, Enterococcus faecalis, Streptococcus pneumoniae*; Gram-negative: *E. coli, Acinetobacter baumannii, Pseudomonas aeruginosa, Klebsiella pneumoniae*) was performed using CLUSTALW[24]. The resulting alignment file and *S. aureus* RnpA crystal structure (PDB ID: 6D1R) were utilized to prepare the alignment figure (**Fig. 2.9**) using ESPript 3.0[25]. 3-Dimential (3D) structures of different RnpA constructs were aligned using PyMol.

2.3   Result and Discussion

2.3.1   Macromolecule production

MBP-RnpA: E. coli MBP is a protein tag commonly used in protein expression and purification for the purpose of improving expression and solubility[26,27]. In protein crystallography, MBP is also used as a carrier protein fused to the target with short linkers to solve recalcitrant protein crystallization issues[28-32]. Moon et al further engineered MBP with the surface entropy reduction (SER) technique by mutating flexible hydrophilic surface side chains, and these engineered MBP tags have been used to solve several problematic protein structures and multiprotein complexes[33-36]. In the first attempt of RnpA crystallization, one of the engineered crystallization-prone MBP tags was selected. This MBP tag consists of the 27-392 residues of E. coli MBP with six flexible surface residues (D83, K84, K240, E360, K363, D364) mutated to alanine. The MBP tag was fused to the N-terminus of RnpA with a short triple alanine linker (underlined in **Table 2.1**).

MBP-RnpA was expressed and purified following the instruction manual of pMAL™ Protein Fusion & Purification System provided by New England Biolabs Inc. (NEB). Different from most protein expression and purification with LB media, the expression of a protein fused with an MBP tag requires the addition of 0.2 % (w/v) *D*-glucose. Glucose is able to inhibit the expression of *mal* genes, one of which encodes an amylase that degrades the amylose of the affinity column used in the following purification step[37,38]. Expression of MBP-RnpA with 6 L of media yielded 8 mg of pure protein. The purified protein was immediately concentrated, flash frozen with liquid nitrogen and stored at -80˚C to preserve its quality.

His-RnpA: The pET-30 Ek/ LIC vector is a commercially available plasmid that allows the fusion of a hexa-histidine tag to the N-terminus of the target protein. Our collaborator cloned RnpA into this vector and used in previous studies[5]. This construct was therefore utilized for crystallography

attempts after the failure of MBP-RnpA crystals reproduction. Poly-histidine affinity tag in combination with immobilized metal-affinity chromatography (IMAC) is widely used in protein purification[39]. The imidazole group of the histidine sidechain forms coordination bonds with the transition metal ion (options are Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+}) immobilized on the column matrix[39], so the protein of interest fused with the poly-histidine tag is able to be extracted out of the cell lysates. Ni^{2+}-NTA was used in the expression and purification of his-RnpA, as it was readily available in our lab. Poly-histidine tagged protein can be eluted with free imidazole solution, and his-RnpA was eluted with 150mM-500mM of imidazole solution. 3 mg of pure protein His-RnpA was yielded with 6 L of media (**Table 2.1**). Because of the instability of His-RnpA after a freeze-thaw cycle, the prepared protein was stored at 4°C and used within a week.

Tag-free RnpA: The cloned pHis2-parallel*::rnpA* vector encodes a hexa-histidine tag and a TEV cleavage site to the N terminal of RnpA, allowing the hexa-histidine tag to be removed by TEV protease and providing a tag-free version of RnpA. In the amino acid sequence of tag-free RnpA, a glycine (GGA) was introduced after the start codon (ATG) of the *rnpA* gene to make up for the frame shift caused by the restriction enzyme cut. An aromatic residue tryptophan (TGG) was introduced after the glycine for the convenience of monitoring protein purification with a UV detector.

The primary expression and purification steps of tag-free RnpA is similar to that of his-RnpA. Yet the before the final SEC purification, tag-free RnpA requires the cleavage of its affinity tag and Ni^{2+}-NTA IMAC purification to separate the free RnpA from the cleaved tag and/or any remain un-cleaved hexa-histidine fused RnpA. SDS-Page electrophoresis study has shown that the TEV cleavage is usually achieved with 100% completion. Separate SDS-Page electrophoresis study showed that the free RnpA has short retention time on the Ni^{2+}-NTA column during IMAC

purification which requires about 150mM-200mM of Imidazole solution to elute. The retention of RnpA on column may be the result of surface exposed histidine residues (His23, His42, His73, His76 and His104), especially the two spatially proximate His73 and His76 (bolded and underlined in **Table 2.1**), interacting with the Ni^{2+} ions of the column. For 6 L of media, expression and purification of tag-free RnpA yielded 1-1.3 mg of pure protein. The prepared protein was stored at 4°C and used within a week.

Table 2.1 Macromolecule production of wild-type *S. aureus* RnpA

	MBP-RnpA	His-RnpA	Tag-free RnpA
Source Organism	*S. aureus* UAMS-1	*S. aureus* UAMS-1	*S. aureus* UAMS-1
Expression vector	pETXM1	pET30 Ek-LIC	pHis2-parallel
Expression host	E. coli BL21(DE3)	E. coli BL21(DE3)	E. coli BL21(DE3)
Molecular weight (Da)	53757.81	17465.44	13797.33
Sequence	MKIEEGKLVIWINGDKG YNGLAEVGKKFEKDTGI KVTVEHPDKLEEKFPQV AATGDGPDIIFWAHDRFG GYAQSGLLAEITPAAAFQ DKLYPFTWDAVRYNGKL IAYPIAVEALSLIYNKDLL PNPPKTWEEIPALDKELK AKGKSALMFNLQEPYFT WPLIAADGGYAFKYENG KYDIKDVGVDNAGAKA GLTFLVDLIKNKHMNAD TDYSIAEAAFNKGETAM TINGPWAWSNIDTSAVN YGVTVLPTFKGQPSKPFV GVLSAGINAASPNKELAK EFLENYLLTDEGLEAVNK DKPLGAVALKSYEEELA KDPRIAATMENAQKGEI MPNIPQMSAFWYAVRTA	MHHHHHHSSGL VPRGSGMKETA AAKFERQHMDS PDLMEKAYRIKK NADFQRIYKKG HSVANRQFVVY TCNNKEIDHFRL GISVSKKLGNAV LRNKIKRAIREN FKVHKSHILAKD IIVIARQPAKDM TTLQIQNSLEHV LKIAKVFNKKIK	GAMGWEKAYRI KKNADFQRIYKK GHSVANRQFVV YTCNNKEIDHFR LGISVSKKLGNA VLRNKIKRAIRE NFKV**H**KS**H**ILAK DIIVIARQPAKD MTTLQIQNSLEH VLKIAKVFNKKI K

Table 2.1 continued

Sequence continued	VINAASGRQTVDAALAA AQTNAAAMEKAYRIKKN ADFQRIYKKGHSVANRQ FVVYTCNNKEIDHFRLGI SVSKKLGNAVLRNKIKR AIRENFKVHKSHILAKDII VIARQPAKDMTTLQIQNS LEHVLKIAKVFNKKIK	n/a	n/a

2.3.2   Crystallization and data collection

MBP-RnpA: Crystallization of MBP-RnpA went through broad screening and 2 steps of optimization. The efficiency of broad screening and first step optimization was significantly improved with liquid handling robots, while the second step manually set-up optimization trays allowed higher volume drops and yielded larger crystals. Vapor diffusion method was adopted for all crystallization experiments as it is the most common method used in crystallography and its effectiveness was demonstrated over years[40]. The best diffracted crystals of MBP-RnpA were produced in 4M NH_4OAc, 1M HPEPS pH 7.5, 50% PEG 3,350. They exhibited spindle morphology (**Fig. 2.1**), and were preserved with 20% ethylene glycol solution before flash-frozen with liquid nitrogen. The best diffraction dataset was cut off at 1.98 Å resolution, and it showed that the protein crystallized in space group C2. Detailed crystallization and data collection data for MBP-RnpA is reported in **Table 2.2** and **Table 2.3**.

However, the reproduction of these crystals had encountered some difficulty as it was hard to prepare protein with high purity. It was then recognized that MBP fusion protein purification with amylose affinity chromatography may have an issue of periplasmic protein contamination[41]. Additionally, the large size of MBP compared to RnpA, is expected to become a disadvantage in future crystal soaking experiments, as it may shield the ligands from accessing RnpA. Both issues

of MBP-RnpA prompted us to withdraw from further study on this construct.

His-RnpA: Broad screening and one-step optimization were performed for His-RnpA in the same manner as crystallization of MBP-RnpA. The produced His-RnpA crystals were tiny and potentially in multi-lattice morphology. The multi-lattice morphology indicates that the crystal was packed in many different crystallographic orientations with smaller crystals, which will result in severe overlaps in the X-ray diffraction patterns and thus difficulty in indexing data points from one single crystal for data processing and structure construction. One way to address the issue is to produce larger crystals, so a ratio screen was carried. 1 µL, 1.5 µL and 2 µL of protein solution were separately mixed with 1 µL of well solution in the hanging drop for the experiment. The best result came out from a condition of 0.8M LiCl, 0.1M Tris-HCl pH8.5, 37% PEG 4,000 with a ratio of 2 µL protein solution to 1 µL well solution in the hanging drop. However, the quality of the crystal did not improve (**Fig. 2.1**). Two 3 Å datasets were collected and processed for His-RnpA, which showed that it was crystallized in space group P21. Detailed crystallization and data collection data for His-RnpA is reported in **Table 2.2** and **Table 2.3.**

Despite the unfavorable crystal and data quality of His-RnpA, the reproduction of its crystals had also encountered some difficulty. It was hypothesized that the tag containing section of this construct was potentially reducing the quality of the crystals and datasets. The tag section accounts for 36 amino acids out of the total of 151 residues, which is about 24% of the total length. The flexible tendency of the long tag impairs the possibility of crystallization and the crystal quality. Therefore, a construct with a readily cleavable affinity tag would be a promising candidate for RnpA crystallization.

Tag-free RnpA: Sparse matrix screens MemGold™ HT-96 and MemGold2™ HT-96 were used for the broad screen of tag-free RnpA because of their availability at the time of the experiment.

These two screens were primarily designed based on data mining with the crystallization conditions of 121 membrane proteins and were to improve success rates for membrane protein crystallization[42]. However, the diversified salts, buffers, pHs, PEGs and additives covered in these screens still validate their usage as non-membrane protein crystallization screens. Two tag-free RnpA crystals were produced from one single condition with 0.15 M Phosphate buffer, 3.3 M $(NH_4)_2SO_4$ at pH 6.5 in the broad screen of MemGold[TM] HT-96. The crystal was in a hexagonal prism morphology (**Fig. 2.1**). Because ammonium sulfate was recognized as a cryo-salt that prevents ice formation and protects protein crystal from freezing damage[43], the 3.3 M ammonium sulfate in the well solution allowed us to freeze the tag-free RnpA crystals without other cryo-protectants. The crystals diffracted to sub-2 Å resolution, and the best dataset was processed to 2.0 Å resolution to retain good data quality. Data showed that tag-free RnpA crystallized in space group P321. Crystal of this RnpA construct was later proven to be reproducible with the same quality, and will therefore be further utilized in ligand soaking experiments for RnpA drug discovery.

MBP-RnpA   His-RnpA   Tag-free RnpA
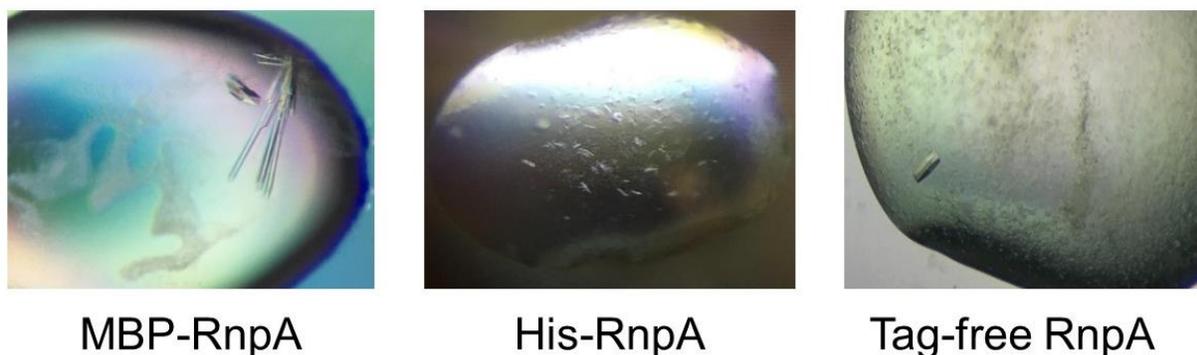
Figure 2.1 *S. aureus* wild-type RnpA crystals prior to harvest.

Table 2.2 Crystallization of wild-type *S. aureus* RnpA

		MBP-RnpA	His-RnpA	Tag-free RnpA
General Information	**Method**	Hanging-drop vapor diffusion		
	Temperature (K)	298.15		
	Protein concentration (mg/mL)	10	22	12
	Protein buffer	50 mM Tris pH 7.4, 50 mM NaCl, 1 mM DTT, 5 mM Maltose	20 mM Tris pH 7.4, 200 mM NaCl	20 mM Tris pH 7.4, 200 mM NaCl
Broad Screen	**Plate type**	Greiner 96-well polypropylene opaque microplate	Greiner 96-well polypropylene opaque microplate	Greiner 96-well polypropylene opaque microplate
	Volume and ratio of drop (μl)	0.2 : 0.2	0.2 : 0.2	0.2 : 0.2
	Volume of reservoir (μl)	250	250	250
	Initial hits	1) 0.2M NH_4OAc, 0.1 M HEPES pH7.5, 25% w/v PEG 3,350 2) 0.1 M $Mg(OAc)_2 \cdot 4H_2O$, 0.1 M MES pH6.0, 22% w/v PEG 4,000	1) 0.8 M LiCl, 0.1 M Tris-HCl, pH 8.5, 32 % (w/v) PEG 4,000 2) 0.2 M Ammonium Citrate Dibasic, 20 % (w/v) PEG 3,350 3) 0.1 M Imidazole: HCl, pH 8 1 M, sodium citrate tribasic dihydrate	0.15 M Phosphate buffer, 3.3 M $(NH_4)_2SO_4$ pH 6.5

Table 2.2 Continued

		Greiner 96-well polypropylene opaque microplate	VDXm™ 24-well Plate with sealant	
First optimization	**Plate type**	Greiner 96-well polypropylene opaque microplate	VDXm™ 24-well Plate with sealant	
	Volume and ratio of drop (µl)	0.2 : 0.2	1 : 1	
	Volume of reservoir (µl)	250	1250	
	Optimization conditions	1) 0-700 mM NH_4OAc (Vertical), 12-36% w/v PEG 4,000 (Horizontal) 1 M HEPES pH 7.5[*] 2) 0-700 mM NH_4OAc (Vertical), 12-36% w/v PEG 3,350 (Horizontal), 1 M HEPES pH 7.5 3) 0-350 mM $Mg(OAc)_2 \cdot 4H_2O$ (Vertical), 10-35% w/v PEG 4,000 (Horizontal), 0.1 M MES pH6.0	1) 0.8 M LiCl, 0.1 M Tris: HCl, pH 8.5, 26-40 % (w/v) PEG 4,000 (in 8 wells) 2) 0.2 M Ammonium Citrate Dibasic, 14-26 % (w/v) PEG 3,350 (in 8 wells) 3) 0.1 M Imidazole-HCl pH 8, 700-1400 mM Sodium Citrate sodium citrate tribasic dihydrate	n/a
Second optimization	**Plate type**	VDXm™ 24-well Plate with sealant	VDXm™ 24-well Plate with sealant	
	Volume and ratio of drop (well solution: protein solution) (µl)	1 : 1	1: 1 1 : 1.5 1 : 2	n/a
	Volume of reservoir	1250 µl	1250 µl	

Table 2.2 Continued

		1) 360-480 mM NH$_4$OAc , 28-33% w/v PEG 4,000, 1 M HEPES pH 7.5 2) 360-480 mM NH$_4$OAc, 28-33% w/v PEG 3,350, 1 M HEPES pH 7.5	0.8M LiCl, 0.1M Tris-HCl pH8.5, 37% PEG 4,000	n/a
Second optimization	**Optimization conditions**			
Condition of the best diffracted crystal		4M NH$_4$OAc, 1M HPEPS pH 7.5, 50% PEG 3,350	0.8M LiCl, 0.1M Tris-HCl pH8.5, 37% PEG 4,000 (1 µl : 2 µl)	0.15 M Phosphate buffer, 3.3 M (NH$_4$)$_2$SO$_4$ pH 6.5

* PEG 4,000 was added as a component instead of PEG 3,350 by accident

Table 2.3 Data collection and processing of wild-type *S. aureus* RnpA

	MBP-RnpA	His-RnpA	Tag-free RnpA
Diffraction source	Beamline 23-ID-B, APS	Beamline 23-ID-D, APS	Beamline 23-ID-B, APS
Wavelength (Å)	1.0	1.3	1.0
Temperature (K)	100	100	100
Detector	Dectris Eiger X 16M	Dectris Pilatus3 6M	Dectris Eiger X 16M
Crystal-detector distance (mm)	199	499	250
Rotation range per image (°)	1	1	1
Total rotation (°)	180	150	180
Exposure time per image (s)	0.5	1	0.5
Space group	C2	P21	*P*321
a, b, c (Å)	128.73, 44.74, 87.23	50.00, 137.96, 55.24	72.16, 72.16, 58.96
α, β, γ (°)	90, 105.21, 90	90, 116.85, 90	90, 90, 120
Mosaicity (°)	0.6-1.0	0.4-1.2	0.4-1.0
Resolution range (Å)	50.000-1.980 (2.010-1.980)	50.000-3.000 (3.110-3.000)	50.00-2.00 (2.07-2.00)
Total No. of reflections	104414	35804	111835 ($\geq$11411)
No. of unique reflections	33718	13434	12313 (1214)
Completeness (%)	97.000 (96.900)	95.100 (92.600)	100 (100)
Redundancy	3.200 (3.200)	2.800 (2.800)	9.1 (9.4)
$\langle I/\sigma (I)\rangle$	6.900 (1.909)[*]	2.900 (1.333)[*]	31.3 (2.34)
$CC_{1/2}$	0.998 (0.851)	0.972 (0.455) †	0.999 (0.625)
$R_{r.i.m.}$	0.066 (0.475)	0.320 (1.066)	0.090 (1.3)
$R_{p.i.m.}$	0.036 (0.333)	0.184 (0.609)	0.030 (1.2)
Overall B factor form Wilson plot (Å^2)	26.750	52.080	39

* $CC_{1/2} \geq 0.5$ was chosen as a data cutoff indicator[44].

†The $CC_{1/2}$ is slightly lower than 0.5. Because of the low quality of the dataset, a lower cutoff was selected to keep more data.

2.3.3   Protein structures

All three structures of RnpA were solved using the RnpA NMR structure (PDB ID: 1D6T) as a

search model, while MBP-RnpA also required the MBP structure (PDB ID: 3PY7) as a co-search

model. The structures were refined until no major geometry issues regarding bond lengths, bond

angles, atomic clashes and symmetry related clashes were identified, and until the best R-values

were achieved (**Table 2.4**). The 3D structure alignments of different RnpA constructs showed that

the three structures share high similarities (RMSD of MBP-RnpA to RnpA = 0.497 Å; RMSD of

His-RnpA to RnpA = 0.694 Å, **Fig. 2.2**). Therefore, in the following sections, the overall structures

of MBP-RnpA and His-RnpA will be generally described, while the tag-free RnpA structure will

be used for more detailed discussion.

Table 2.4 Structure solution and refinement of wild-type *S. aureus* RnpA

		MBP-RnpA	**His-RnpA**	**Tag-free RnpA**
PDB entry		n/a	n/a	6D1R
Resolution range (Å)		44.6880–1.9770 (2.0268–1.9774)	49.2870-2.9970 (3.1171-2.9971)	31.25-2.00 (2.10-2.00)
Completeness (%)		96.3	95.0	99.0 (100)
σ cutoff		$F > 1.380\sigma(F)$	$F > 1.380\sigma(F)$	$F > 1,360\sigma(F)$
No. of reflections, working set		30692 (1978)	11463 (1219)	11086 (1125)
No. of reflections, test set		1999 (128)	1286 (139)	1223 (123)
Final R_{cryst}		0.166 (0.2197)	0.267 (0.3438)	0.214 (0.324)
Final R_{free}		0.213 (0.3082)	0.331 (0.4137)	0.234 (0.343)
No. of non-H atoms	**Protein**	3723	3533	868
	Ligand	34	0	0
	Solvent	310	0	36
	Total	4067	7231	904

Table 2.4 continued

| R.m.s. deviations	Bonds (Å)	0.010	0.02	0.004
	Angles (°)	1.223	0.504	0.640
Average B factors (Å²)	Protein	43	67	60
	Solvent	39	n/a	49
Ramachandran plot	Most favored (%)	97	91.88	99.07
	Allowed (%)	3	7.42	0.93
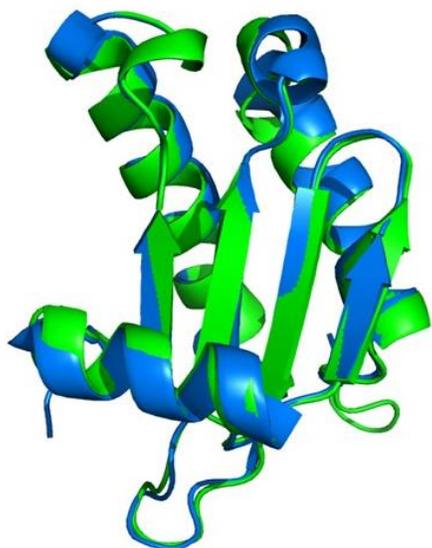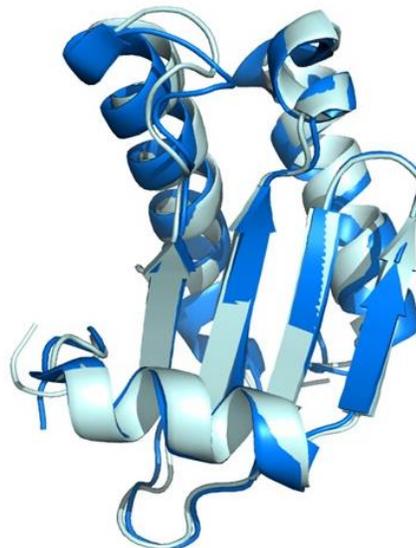
Figure 2.2 Tag-free RnpA crystal structure alignment with two other constructs of RnpAs. A) Tag-free *S. aureus* RnpA aligned with the MBP-RnpA (green) RMSD = 0.497 Å. B) Tag-free *S. aureus* RnpA aligned with the His-RnpA (pale blue) RMSD=0.694 Å.

MBP-RnpA: MBP-RnpA has a single monomer in the asymmetric unit, and its 486 residues of the protein were modelled. As reported in **Table 2.4**, there were no Ramachandran outliers, and 97 % of the residues are in the favored region and 3 % of residues in the allowed region. Side chains of residues Gln 99, Lys 114, Lys 115 were removed because of the lack of side chain density. The refined structure showed that the MBP tag is at the N-terminus of RnpA (**Fig. 2.3A**). Because of the close proximity between the tag and RnpA, the RnpA C-terminus and residues 41-42 in the flexible loop showed non-specific interactions with the MBP tag (**Fig. 2.3B**). Besides, a maltotriose was found in the maltose binding site of MBP, and this ligand was believed to be introduced as an impurity of the maltose powder used to make the elution buffer.
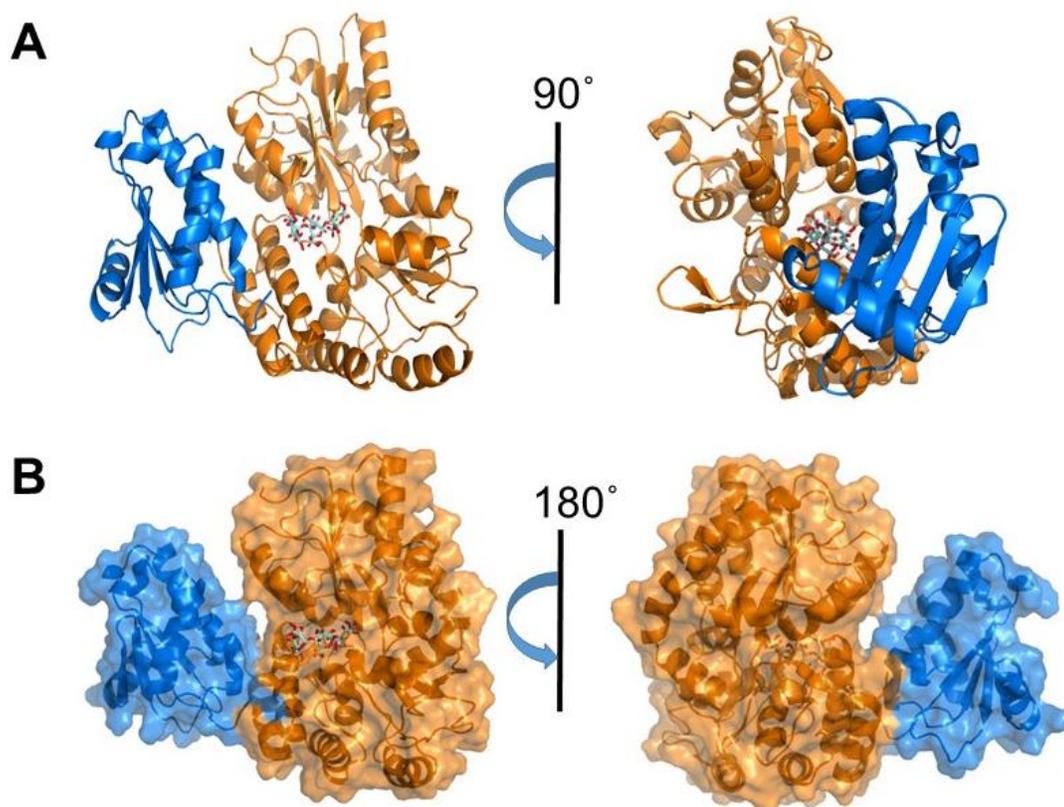

Figure 2.3 MBP-RnpA crystal structure at 2.0 Å resolution. A) Cartoon representation of MBP-RnpA's overall architecture (RnpA, marine. MBP tag, orange). B) Surface representation of MBP-RnpA presents interactions between RnpA and MBP-tag (transparency: 40%). A maltotriose ligand is present in each representation (Backbone, cyan. Hydrogens, red).

His-RnpA: His-RnpA has four monomers in the asymmetric unit. In the structure model, both Chain A and Chain B consist of residues 5-114, Chain C contains residues 5-112 and Chain D includes residues 4-114. The missing electron density and modelled residue of the His-tag, N- and C-termini of RnpA were largely due to the high flexibility of these parts. The four monomers interact with each other differently: Chain A monomer interacts with an inverted chain B monomer at the edge of the β sheet. Chain D monomer interacts with chain A monomer in a 90-degree manner. Chain C monomer, however, does not interact with any other monomers (**Fig. 2.4**). Because 3.0 Å resolution structure only shows the basic contours the protein chain instead of exact atomic positions, interactions between residue side chains are not further discussed. Detailed structure determination and refinement information is reported in **Table 2.4**.
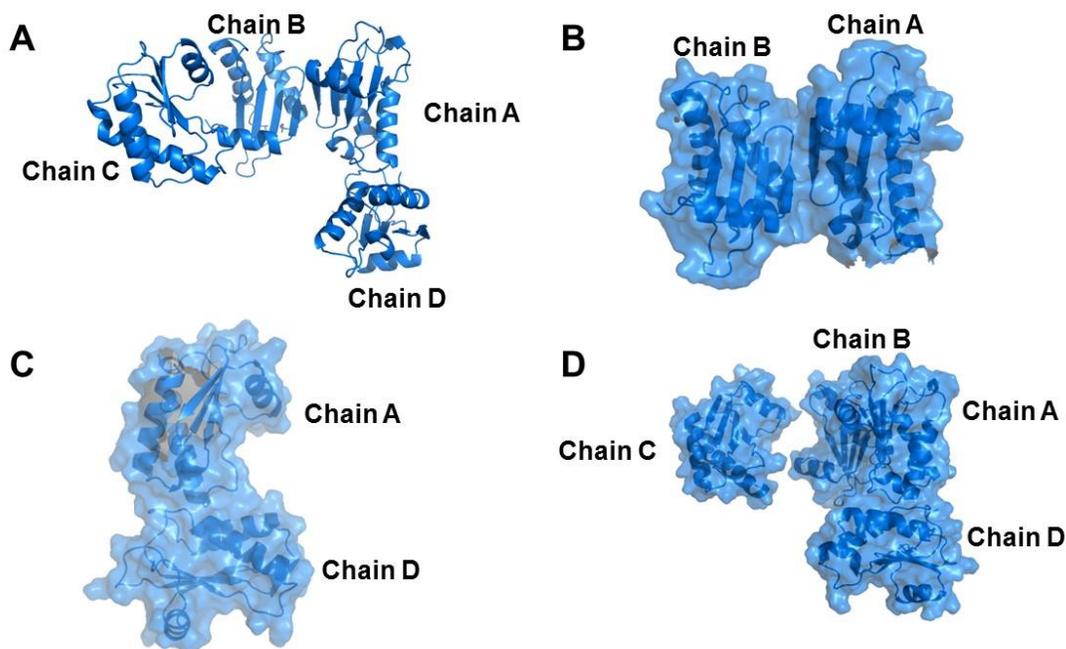

Figure 2.4 His-RnpA crystal structure at 3.0 Å resolution. Cartoon representation of His-RnpAs in an asymmetric unit. B), C) and D) surface representation of different his-RnpA monomers (transparency: 40%).

Tag-free RnpA: This protein model was refined to R_{crys}/R_{free}= 0.214/0.234, where R is the statistical residual value. R_{crys} was computed with the data used to refine the protein structure, while R_{free} was determined with a small part of 10% of the dataset that is not used in the refinement[45]. The smaller the R values, the better agreement between the experimental data and the calculated structure[46]. The validation report of tag-free RnpA from the Protein Data Bank (PDB) showed that the R_{free} value is better than the average R_{free} values of X-ray structures of similar resolution and of all X-ray structures.

In the protein model, a single monomer of tag-free RnpA molecule is shown in the asymmetric unit. The structure has residues 5-113 modelled, and 99% of residues are in the Ramachandran favored region and 1% in the allowed region. The electron density of six N-terminal residues (GAMGWE) and four C-terminal residues (KKIK) were missing because of their high flexibility. Real-space R-value Z-score (RSRZ) outliers lies in regions 12–16, 37–41 and 74–79, which do not have intact electron density for residue sidechains (**Fig. 2.5**).
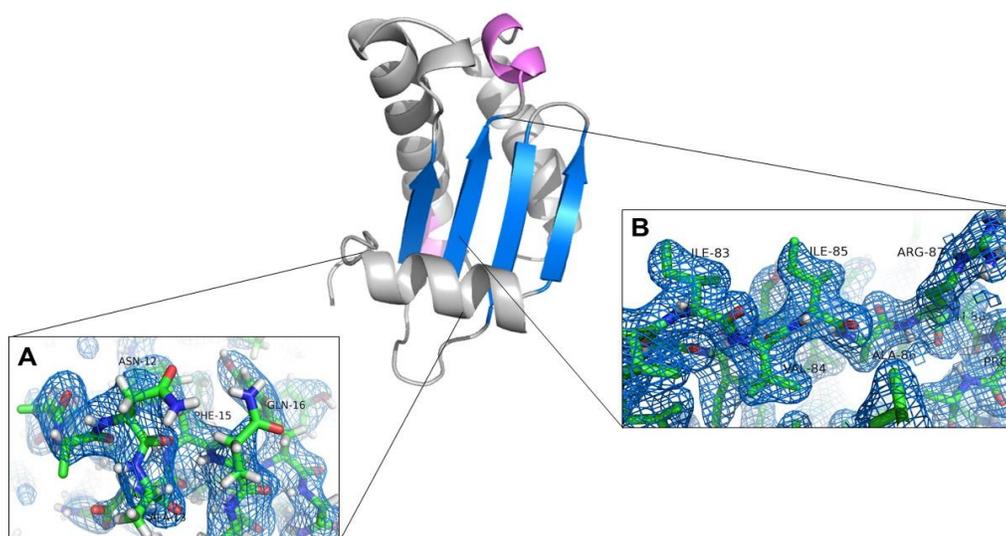

Figure 2.5 Density maps of different regions in tag-free RnpA. A) Density map of residues 12-16 showed intact density around residue sidechains. B) Density map of residues 83-87 showed intact density around residues including their sidechains.

This crystal structure showed that *S. aureus* RnpA is a compact, globular protein similar to that of previously known RnpA structures. It has secondary structures of α_1-β_1-β_2-β_3-α_2-η_1-β_4- η_2-α_3 along the protein sequence from the N- to the C-termini (**Fig. 2.6**). The four β-strands form a sheet arranged in the order of β_1-β_2-β_4-β_3 from right to left same as observed previously (**Fig. 2.6**, marine)[15]. Strands β_1-β_2 lay antiparallel to each other, while β_4-β_3 orient in parallel and β_3-α_2-η_1-β_4 formed an unusual left-handed topology[11,47]. On the front side of this β-sheet, helix α_1 (residues 12–21) locates at the bottom running perpendicular to the strands. Helices α_2 (residues 56–73) and α_3 (residues 95–109) packed against the back side of the β-sheet with hydrophobic interactions. The 3_{10}-helix η_1 (residues 74-76, **Fig. 2.6**, violet) appears commonly as an extension of a helix (α_2 in RnpA) where the polypeptide chain starts to change directions[48]. 3_{10}-helix η_2 (residues 88–92, **Fig. 2.6**, violet), however, is not connected to a helix, but it mediates the direction change of the polypeptide chain from the winding-up strand β_4 to the going downward helix α_3 and was only 2 residues away from the beginning of helix α_3.
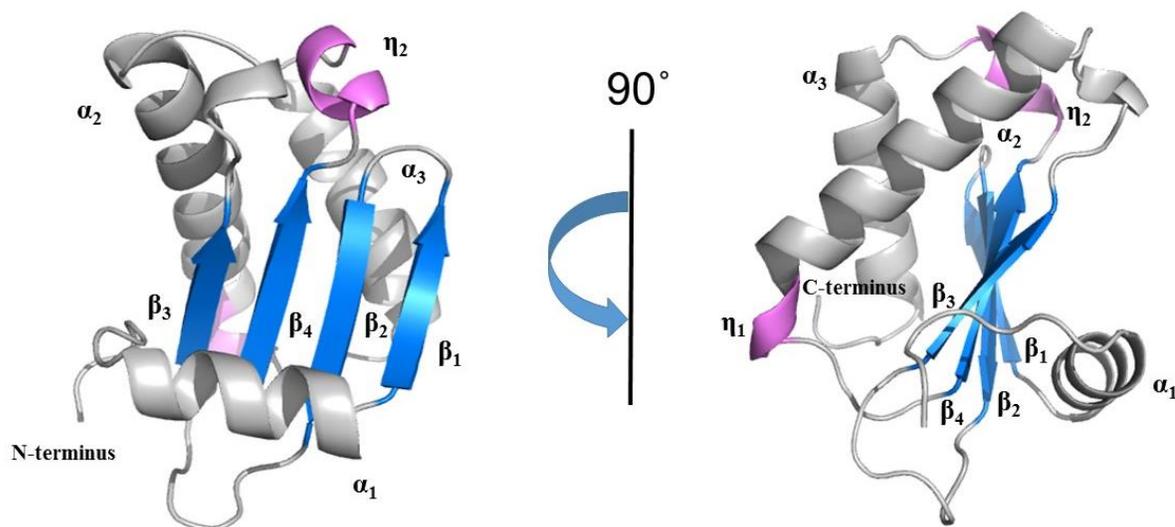

Figure 2.6 Tag-free RnpA crystal structure at 2.0 Å resolution. The β-strands is in blue. The 3_{10} helices is in violet. The rest of the protein is in grey.

This crystal structure of *S. aureus* RnpA includes structural motifs for protein-RNA interactions as described in previous studies[11,15]. The P-RNA binding domain is a highly phylogenetically conserved segment called the "RNR box" motif. The consensus sequence of this motif is K-X_{4-5}-A-X_2-R-N-X_2-K/R-R-X_2-R/K[11,49-51], which corresponds to residues 52–67 (KLGNAVLRNKIKRAIR; **Fig. 2.7A**, magenta sticks, and **Fig. 2.11**, marked with stars). This arginine-rich segment is a typical RNA binding domain, as the positive-charged sequence facilitates electrostatic contacts with the phosphate backbones of the catalytic P RNA. The other RNA binding site of *S. aureus* RnpA is a groove formed by the β-sheet and helix $α_1$, and is recognized as the binding site for the leader sequence of ptRNA substrates[15]. This groove is solvent-exposed, and on its surface reside a group of aromatic and/or hydroxyl-containing residues (Phe15, Tyr19, Ser24, Tyr33 and Ser48; **Fig. 2.7B**, pink sticks). The aromatic residues are likely to stack with the RNA bases, while the hydroxyl-containing residues can form hydrogen bonds with donors and acceptors within the single strand nucleotide chain[13,15,52]. In addition, the conserved Lys51 and Lys52 (**Fig. 2.7B**, green sticks) flanking the opening of this groove were reckoned as electrostatic stabilizers for interaction with RNA[13,52]. Lys52, however, was considered to be involved in different RNA interactions according to different studies, and it will require further structural information of substrate-bound *S. aureus* RNase P holoenzyme or specific biochemical assays to draw correct conclusions. The 5'-leader sequence from the substrate-bound *T.maritima* RNase P holoenzyme crystal structure (PDB ID: 3Q1R) was aligned to the *S. aureus* RnpA structure. The surface representation showed that the backbone of the 5'-leader sequence (yellow sticks) fits into the groove (**Fig. 2.8**). In crystallography, B-factor depicts the scale to which the electron density of one atom extends from its mean position. An atom with a large B-factor therefore is suggested to be more flexible than one with a low B-factor[53]. The B-factor map

of *S. aureus* RnpA showed that helix α₁ of this 5'-leader sequence binding site has higher flexibility than the rest of the secondary structures, and it is comparable to the flexibility of a free loop (residues 37-42). (**Fig. 2.9**, increased red color and thicker cylinder indicates higher B factor). This indicated that the substrate binding of this groove is likely an induced-fit process, and it also provided a possible explanation of the substrate diversity of RNase P[54].
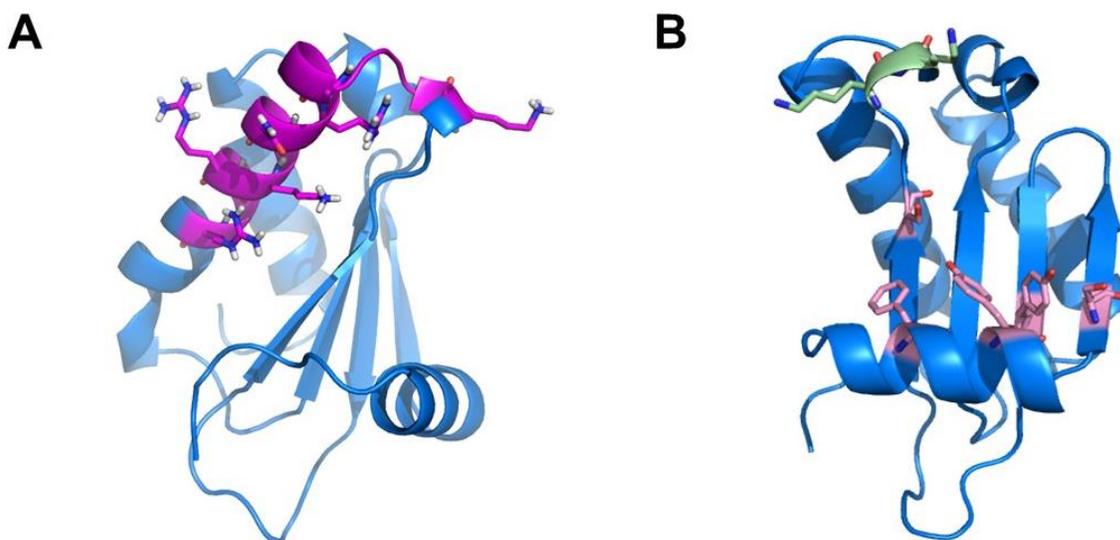

Figure 2.7 RnpA residues involved in protein-RNA interactions. A) P-RNA-binding RNR box, magenta. The conserved residues are in stick. B) The ptRNA 5' leader sequence binding site and important residues. The aromatic and/or hydroxyl-containing residues, pink. Lys 51 and Lys 52, green.
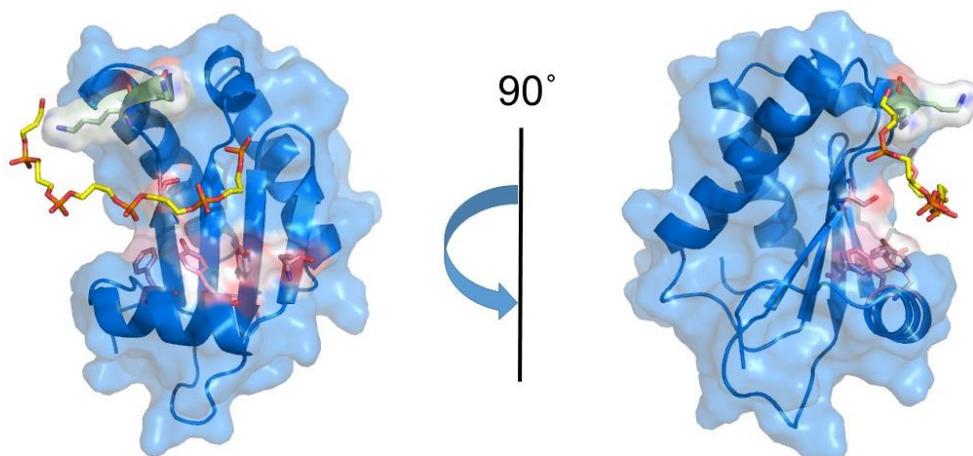
Figure 2.8. *S. aureus* RnpA surface representation overlaid with the phosphate backbone of 5'-leader sequence of ptRNA substrate. The transparency is 50%. The ptRNA substrate is in yellow sticks (PDB entry 3q1r).
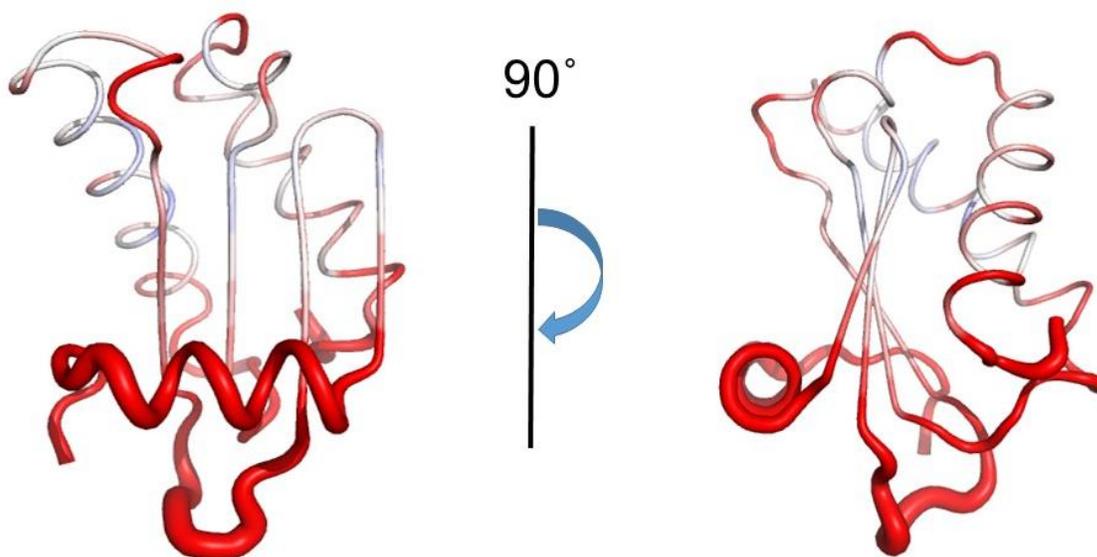

Figure 2.9 B-factor map of *S. aureus* RnpA crystal structure. Red and thick cylinder indicates region of protein with highest B-factor values, blue and thin cylinder indicates area of the protein with lowest B-factor values.

In regards to the previously solved *S. aureus* RnpA NMR structure, the crystal structure was aligned with the average structure of its 20 conformers, and an RMSD of 1.2 Å was found between them (**Fig. 2.10**). The two structures are similar, but the crystal structure has better defined secondary structures. In particular, β_1 and β_2 were more ordered in the crystal structure, and 3_{10} helix η_1 was not identified in the NMR structure.
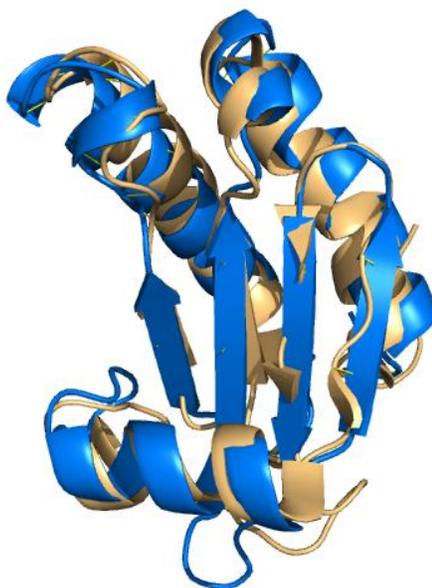
Figure 2.10 The *S. aureus* RnpA crystal structure overlaid onto the average of its 20 NMR conformers. The RnpA crystal structure is in marine. *S. aureus* RnpA average structure of 20 NMR conformers is in light orange (PDB: 1D6T). The alignment of these two structures has an RMSD = 1.2 Å.

2.3.4   Comparison of RnpAs from different species

S. aureus RnpA sequence was aligned with eight other bacterial RnpAs, and the results depict the essentiality of conserved P-RNA-binding RNR-box and the aromatic residues in the 5'-leader sequence binding domain (**Fig. 2.11**). These residues of the RNR-box consensus sequence are 100% conserved across all species and the aromatic residues are over 75% conserved. These observations corroborate with the fact that RnpA complexes with the P RNA to form the RNase P holoenzyme and function as an RNA processor in the bacterial cells.
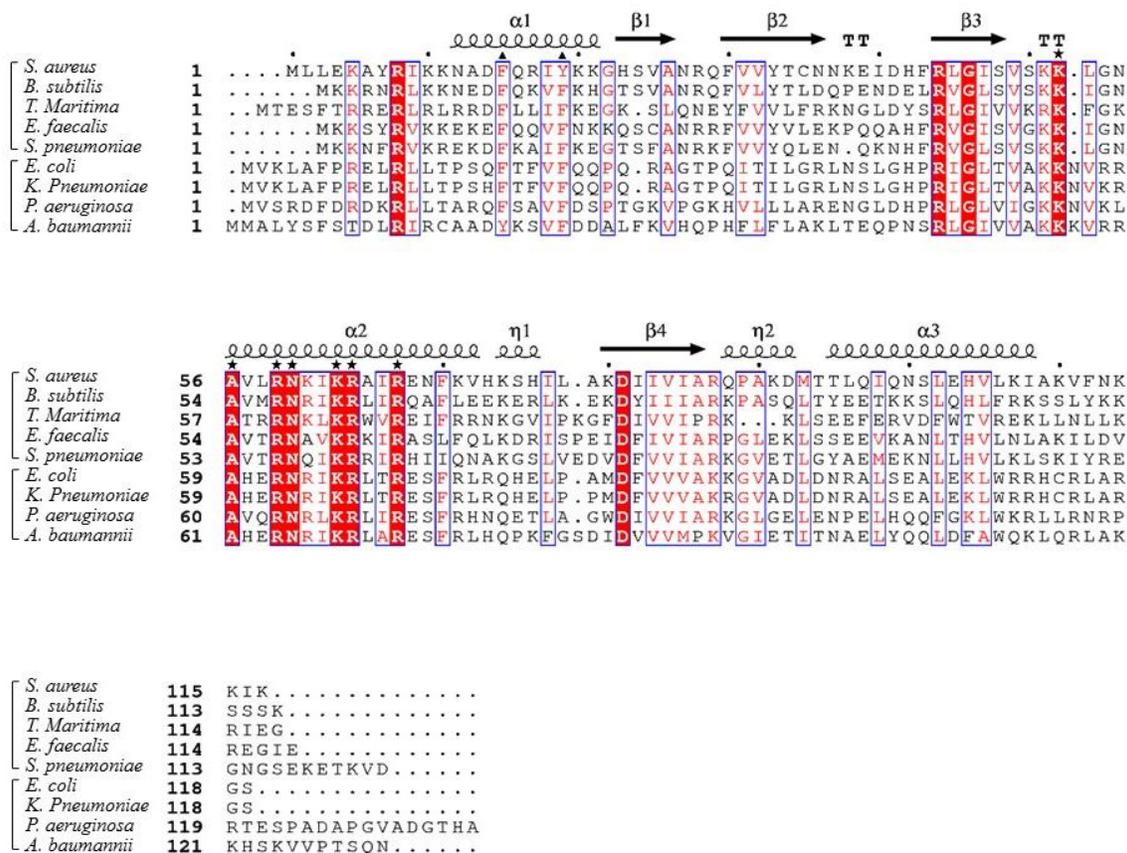
Figure 2.11 Sequence alignment of *S. aureus* RnpA with eight other bacterial RnpA. Eight bacterial RnpA includes four Gram-positive (top bracket) and four Gram-negative (bottom bracket). Sequence similarity depicted by shaded boxes: Red box–white character (100% identity), red character (>70% residue similarity), blue box (>70% residue similarity across group). Secondary structure sequences: helices (curly line) and sheets (arrows) strict β-turns (TT). RNR-box conserved residues (star). Reserved aromatic residues in ptRNA 5'leader sequence binding domain (triangle).

In addition, the *S. aureus* RnpA crystal structure was aligned with RnpAs from *B. subtilis* (PDB ID 1A6F) and *T. maritima* (PDB ID 1NZ0), respectively. (**Fig. 2.12**) *S. aureus* RnpA and *B. subtilis* RnpA alignment has an RMSD of 0.882 Å, while their sequence identity is 48%. *S. aureus* RnpA and *T. maritima* RnpA alignment has an RMSD of 0.916 Å, while their sequence identity is 46%. Taken together, the high conservation of 3D architecture and secondary structural components of RnpAs despite of the low sequence identity, explains the ability for RnpA subunits of bacterial species to be used interchangeably with P RNA of other bacterial species[55].
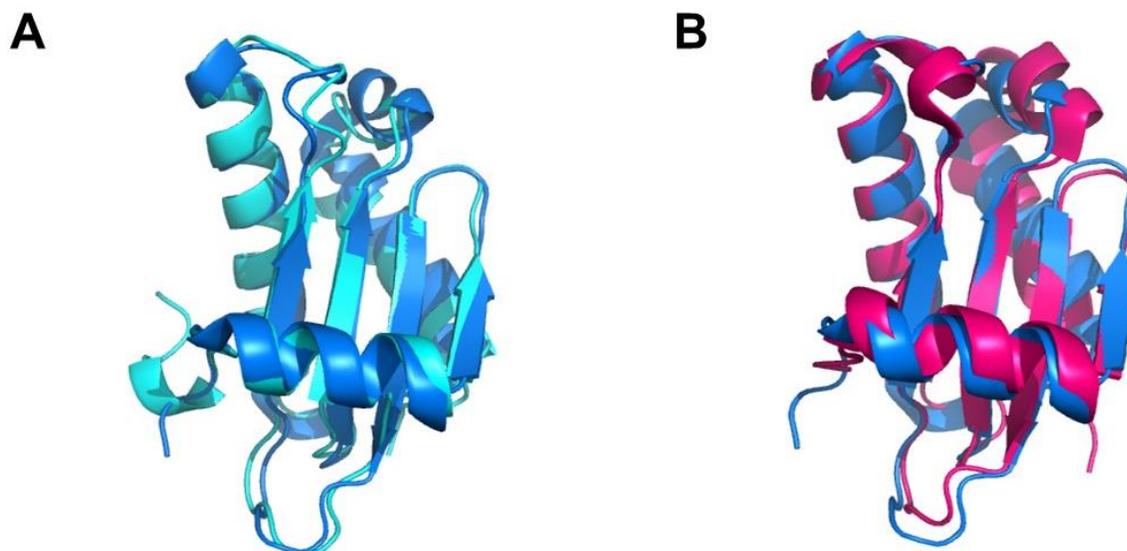
Figure 2.12 The *S. aureus* RnpA crystal structure overlaid onto RnpAs from different species. A) The *S. aureus* RnpA (marine, PDB ID: 6D1R) aligned with the *B. subtilis* RnpA (cyan, PDB ID: 1A6F) RMSD = 0.882 Å. B) The *S. aureus* RnpA (marine, PDB ID: 6D1R) aligned with the *T. maritima* RnpA (pink, PDB ID: 1NZ0). RMSD=0.916 Å

2.4   Conclusion

RnpA is an essential component of two RNA processing macromolecular machines in *S. aureus*: the RNase P holoenzyme for tRNA maturation and mRNA degradosome for mRNA degradation. In this chapter, I report my attempts in crystallization of the *S. aureus* RnpA protein subunit. Three different constructs were utilized, and crystals structures were solved for each of them. It was shown that the structures are highly similar to each other. The pHis2-parallel*::rnpA* construct was selected for future usage because of the high quality of crystals and diffraction data it produced and its reproducibility. This construct produces the tag-free RnpA, which crystallized with high quality crystals in the first-step broad screening. The final structure of tag-free RnpA has a resolution of 2.0 Å, and the model has R_{crys}/R_{free}= 0.214/0.234. This crystal structure of *S. aureus* RnpA shares high similarity with its previously solved solution state NMR structure and RnpA structures from other species. The conservation of the P-RNA-binding RNR-box and the aromatic

residues in the 5'-leader sequence binding domain addressed the essentiality to RnpA functions. The conservation of 3D architecture and secondary structural components of bacterial RnpAs despite of the low sequence identity, accounts for the ability of RnpA subunits from different bacterial species to be used interchangeably.

This work is instrumental as an alternative approach to obtain ligand-bound structural data in support of future structure-based drug discovery for multi-drug resistant *S. aureus* treatment. This structure may also begin to shed light on the newly hypothesized role of *S. aureus* RnpA in RNA degradation in addition to ptRNA maturation.

References

1   Hartmann, E. & Hartmann, R. K. The enigma of ribonuclease P evolution. *Trends Genet* **19**, 561-569, doi:10.1016/j.tig.2003.08.007 (2003).
2   Evans, D., Marquez, S. M. & Pace, N. R. RNase P: interface of the RNA and protein worlds. *Trends Biochem Sci* **31**, 333-341, doi:10.1016/j.tibs.2006.04.007 (2006).
3   Schedl, P. & Primakoff, P. Mutants of Escherichia coli thermosensitive for the synthesis of transfer RNA. *Proc Natl Acad Sci U S A* **70**, 2091-2095 (1973).
4   Roux, C. M., DeMuth, J. P. & Dunman, P. M. Characterization of components of the Staphylococcus aureus mRNA degradosome holoenzyme-like complex. *J Bacteriol* **193**, 5520-5526, doi:10.1128/JB.05485-11 (2011).
5   Eidem, T. M. *et al.* Small-molecule inhibitors of Staphylococcus aureus RnpA-mediated RNA turnover and tRNA processing. *Antimicrob Agents Chemother* **59**, 2016-2028, doi:10.1128/AAC.04352-14 (2015).
6   Olson, P. D. *et al.* Small molecule inhibitors of Staphylococcus aureus RnpA alter cellular mRNA turnover, exhibit antimicrobial activity, and attenuate pathogenesis. *PLoS Pathog* **7**, e1001287, doi:10.1371/journal.ppat.1001287 (2011).
7   *Ribonuclease P*. 1 edn, (Springer-Verlag New York, 2010).
8   Willkomm, D. K. *et al.* Evaluation of bacterial RNase P RNA as a drug target. *Chembiochem* **4**, 1041-1048 (2003).
9   Berchanski, A. & Lapidot, A. Bacterial RNase P RNA is a drug target for aminoglycoside-arginine conjugates. *Bioconjugate Chem* **19**, 1896-1906 (2008).
10   Drainas, D. Antibiotics and RNase P. *Antibiotics (Basel)* **5**, doi:10.3390/antibiotics5020015 (2016).
11   Stams, T., Niranjanakumari, S., Fierke, C. A. & Christianson, D. W. Ribonuclease P protein structure: evolutionary origins in the translational apparatus. *Science* **280**, 752-755 (1998).
12   Kazantsev, A. V. *et al.* High-resolution structure of RNase P protein from Thermotoga maritima. *Proc Natl Acad Sci U S A* **100**, 7497-7502, doi:10.1073/pnas.0932597100 (2003).
13   Reiter, N. J. *et al.* Structure of a bacterial ribonuclease P holoenzyme in complex with tRNA. *Nature* **468**, 784-789, doi:10.1038/nature09516 (2010).
14   Zeng, D., Brown, B. P., Voehler, M. W., Cai, S. & Reiter, N. J. NMR resonance assignments of RNase P protein from Thermotoga maritima. *Biomol NMR Assign* **12**, 183-187, doi:10.1007/s12104-018-9806-7 (2018).
15   Spitzfaden, C. *et al.* The structure of ribonuclease P protein from Staphylococcus aureus reveals a unique binding site for single-stranded RNA. *J Mol Biol* **295**, 105-115, doi:10.1006/jmbi.1999.3341 (2000).
16   Ha, L. *et al.* Crystal structure of the ribonuclease-P-protein subunit from Staphylococcus aureus. *Acta Crystallogr F Struct Biol Commun* **74**, 632-637, doi:10.1107/S2053230X18011512 (2018).
17   Vincze, T., Posfai, J. & Roberts, R. J. NEBcutter: A program to cleave DNA with restriction enzymes. *Nucleic Acids Res* **31**, 3688-3691 (2003).
18   Stothard, P. The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. *Biotechniques* **28**, 1102, 1104, doi:10.2144/00286ir01 (2000).
19   Jancarik, J. & Kim, S. H. Sparse-Matrix Sampling - a Screening Method for Crystallization of Proteins. *Journal of Applied Crystallography* **24**, 409-411 (1991).

20   Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276**, 307-326 (1997).

21   McCoy, A. J. *et al.* Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674, doi:10.1107/S0021889807021206 (2007).

22   Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221, doi:10.1107/S0907444909052925 (2010).

23   Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501, doi:10.1107/S0907444910007493 (2010).

24   Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-2948, doi:10.1093/bioinformatics/btm404 (2007).

25   Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* **42**, W320-324, doi:10.1093/nar/gku316 (2014).

26   Kapust, R. B. & Waugh, D. S. Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci* **8**, 1668-1674, doi:10.50/ps.8.8.1668 (1999).

27   Sachdev, D. & Chirgwin, J. M. Solubility of proteins isolated from inclusion bodies is enhanced by fusion to maltose-binding protein or thioredoxin. *Protein Expr Purif* **12**, 122-132, doi:10.1006/prep.1997.0826 (1998).

28   Ke, A. & Wolberger, C. Insights into binding cooperativity of MATa1/MATalpha2 from the crystal structure of a MATa1 homeodomain-maltose binding protein chimera. *Protein Sci* **12**, 306-312, doi:10.1110/ps.0219103 (2003).

29   Liu, Y. *et al.* Crystal structure of the SarR protein from Staphylococcus aureus. *Proc Natl Acad Sci U S A* **98**, 6877-6882, doi:10.1073/pnas.121013398 (2001).

30   Smyth, D. R., Mrozkiewicz, M. K., McGrath, W. J., Listwan, P. & Kobe, B. Crystal structures of fusion proteins with large-affinity tags. *Protein Sci* **12**, 1313-1322, doi:10.1110/ps.0243403 (2003).

31   Kobe, B., Center, R. J., Kemp, B. E. & Poumbourios, P. Crystal structure of human T cell leukemia virus type 1 gp21 ectodomain crystallized as a maltose-binding protein chimera reveals structural evolution of retroviral transmembrane proteins. *P Natl Acad Sci USA* **96**, 4319-4324 (1999).

32   Center, R. J. *et al.* Crystallization of a trimeric human T cell leukemia virus type 1 gp21 ectodomain fragment as a chimera with maltose-binding protein. *Protein Sci* **7**, 1612-1619, doi:10.1002/pro.5560070715 (1998).

33   Martinez-Zapien, D. *et al.* Structure of the E6/E6AP/p53 complex required for HPV-mediated degradation of p53. *Nature* **529**, 541-+ (2016).

34   Ullah, H. *et al.* Structure of a signal transduction regulator, RACK1, from Arabidopsis thaliana. *Protein Sci* **17**, 1771-1780, doi:10.1110/ps.035121.108 (2008).

35   Bethea, H. N., Xu, D., Liu, J. & Pedersen, L. C. Redirecting the substrate specificity of heparan sulfate 2-O-sulfotransferase by structurally guided mutagenesis. *Proc Natl Acad Sci U S A* **105**, 18724-18729, doi:10.1073/pnas.0806975105 (2008).

36   Zanier, K. *et al.* Structural basis for hijacking of cellular LxxLL motifs by papillomavirus E6 oncoproteins. *Science* **339**, 694-698, doi:10.1126/science.1229934 (2013).

37   Lengsfeld, C., Schonert, S., Dippel, R. & Boos, W. Glucose- and glucokinase-controlled mal gene expression in Escherichia coli. *J Bacteriol* **191**, 701-712, doi:10.1128/JB.00767-08 (2009).

38   Freundlieb, S. & Boos, W. Alpha-amylase of Escherichia coli, mapping and cloning of the structural gene, malS, and identification of its product as a periplasmic protein. *The Journal of biological chemistry* **261**, 2946-2953 (1986).

39   Bornhorst, J. A. & Falke, J. J. Purification of proteins using polyhistidine affinity tags. *Method Enzymol* **326**, 245-254 (2000).

40   Stevens, R. C. High-throughput protein crystallization. *Curr Opin Struct Biol* **10**, 558-563 (2000).

41   New England Biolabs, I.

42   Newstead, S., Ferrandon, S. & Iwata, S. Rationalizing alpha-helical membrane protein crystallization. *Protein Sci* **17**, 466-472, doi:10.1110/ps.073263108 (2008).

43   Rubinson, K. A., Ladner, J. E., Tordova, M. & Gilliland, G. L. Cryosalts: suppression of ice formation in macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **56**, 996-1001 (2000).

44   Karplus, P. A. & Diederichs, K. Assessing and maximizing data quality in macromolecular crystallography. *Curr Opin Struct Biol* **34**, 60-68, doi:10.1016/j.sbi.2015.07.003 (2015).

45   Brunger, A. T. Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472-475 (1992).

46   Kleywegt, G. J. & Brunger, A. T. Checking your imagination: applications of the free R value. *Structure* **4**, 897-904 (1996).

47   Richardson, J. S. Describing patterns of protein tertiary structure. *Methods Enzymol* **115**, 341-358 (1985).

48   Kendrew, J. C. *et al.* A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature* **181**, 662-666 (1958).

49   Pace, N. R. & Brown, J. W. Evolutionary perspective on the structure and function of ribonuclease P, a ribozyme. *J Bacteriol* **177**, 1919-1928 (1995).

50   Brown, J. W. The ribonuclease P database. *Nucleic Acids Res* **26**, 351-352 (1998).

51   Altman, S. Ribonuclease P: an enzyme with a catalytic RNA subunit. *Adv Enzymol Relat Areas Mol Biol* **62**, 1-36 (1989).

52   Reiter, N. J., Osterman, A. K. & Mondragon, A. The bacterial ribonuclease P holoenzyme requires specific, conserved residues for efficient catalysis and substrate positioning. *Nucleic Acids Res* **40**, 10384-10393, doi:10.1093/nar/gks744 (2012).

53   Bahar, I., Atilgan, A. R. & Erman, B. Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. *Fold Des* **2**, 173-181 (1997).

54   Frank, D. N. & Pace, N. R. Ribonuclease P: unity and diversity in a tRNA processing ribozyme. *Annu Rev Biochem* **67**, 153-180, doi:10.1146/annurev.biochem.67.1.153 (1998).

55   Guerriertakada, C., Gardiner, K., Marsh, T., Pace, N. & Altman, S. The Rna Moiety of Ribonuclease-P Is the Catalytic Subunit of the Enzyme. *Cell* **35**, 849-857 (1983).

CHAPTER 3. BIOPHYSICAL ANALYSIS OF *STAPHYLOCOCCUS AUREUS* RNPA[P89A]

3.1 Introduction

S. aureus RnpA is a dual-functional protein subunit that evolved to be essential in two cellular processes: ptRNA processing and mRNA degradation. Information of the protein's essential residue(s) for ptRNA processing has been well-studied, while there is little information regarding the mechanism for which RnpA participates in mRNA degradation. Yet such information will equip us with the knowledge of 1) potential drug targeting sites 2) mechanism of action for inhibitors 3) hit optimization strategies and 4) potential interaction region(s) between RnpA protein subunit and its binding partners within the mRNA degradosome complex.

Previous studies on individual residue contribution to RnpA ptRNA processing activity mainly focused on RnpA's RNA binding sites, the "RNR box" (P RNA binding site) and the central cleft (ptRNA 5' leader sequence binding site). Results showed that when applied to recombinant RnpA, site-directed mutagenesis of conserved residues (R59, N60, K63, R64. *S. aureus* numbering) within the "RNR box" to cysteines in *B. subtilis* had no effect on the *in vitro* ptRNA processing activity of RNase P[1]. Alanine mutants to the same region in *E. coli* RnpA yielded reduced initial velocity (v_{int} = 49-81% comparing to wild-type normalized to 100%)[2] . A similar effect was seen when individual conserved hydrophobic residues (F18, v_{int} =36%; F22, v_{int} = 42%. *E. coli* numbering) in the RnpA central cleft was mutated into alanine in *E. coli*, and double mutation F18A/F22A caused dramatic decrease in ptRNA hydrolysis rate (v_{int} =7%)[2]. In an *in cellulo* complementation study, a chromosomal *rnpA* gene mutation (R45H) caused temperature-sensitive *E. coli* strain was used, and it was shown that all the individual RnpA mutants mentioned above were able partially rescue bacteria growth of this strain after heat shock, and double mutant failed

to rescue the cells[2]. While these studies showed the importance of some conserved basic and hydrophobic residues for RnpA-mediated ptRNA processing activity, no systematic interrogation of other RnpA essential amino acids, including those needed for mRNA degradation has been performed. Therefore, our team, in collaboration with the Dunman group, set off to fill in this niche. An alanine scanning library (alanine codon: GCA) with every individual residue being mutated (except for residue Met1) was generated (Note: alternative alanine codons at wild-type alanine positions A6, A13, A26, A65, A79, A86, A90, and A109 have also been changed to GCA). Each mutant was inserted into a *pML100* plasmid downstream to an anhydrotetracycline (ATc)-inducible promoter. Each plasmid was transduced into a wild-type *S. aureus* strain RN4220. Bacteria viability of *S. aureus* with over-expressed RnpA mutants were closely monitored by colony-forming unit determination. RnpA mutants that impaired bacteria viability were identified and further investigated for their effect on ptRNA processing and mRNA degradation both *in vitro and in cellulo.* So far 108 RnpA mutants have been screened for their effect on *S. aureus* viability, 16 of which have been tested for their effects on RNA processing pathways *in vitro*, and 7 out of the 16 mutants have been further investigated for their cellular RNA processing abilities. Among the 7 mutants that have been through the whole testing process, RnpA^{P89A} has drawn our attention.

In the bacteria viability study of RnpA^{P89A}, RnpA^{A79A} which is a wild-type RnpA phenotype was used as the control. Both cultures of plasmid bearing bacteria were grown to exponential phase, and then overexpression of RnpA^{P89A} and RnpA^{A79A} were induced by the addition of ATc (final concentration=312 ng/ml). Samples were taken at every hour from -1hr – 4hr (induction = 0 hr) and plated promptly to determine CFU/ml at each time point. As previously determined by western blotting using a polyclonal antibody generated by Dunman's lab, RnpA production reaches climax (9.8-fold increase in cellular RnpA) at 1.5 hr post-induction, and returns to baseline ~ 4 hr post-

induction presumably due to inducer depletion. As shown in **Fig. 3.1**, RnpA[P89A] displayed a 0.18

log reduction in CFU/ml at 1 hr post-induction compared to that at -1 hr before induction, while

the control group exhibited bacteria growth by 0.36 log CFU/ml. The accumulation of RnpA[P89A]

slowed down the growth of *S. aureus* in comparison to cells overexpressing wild-type protein,

indicating that residue proline89 is important for bacteria viability.
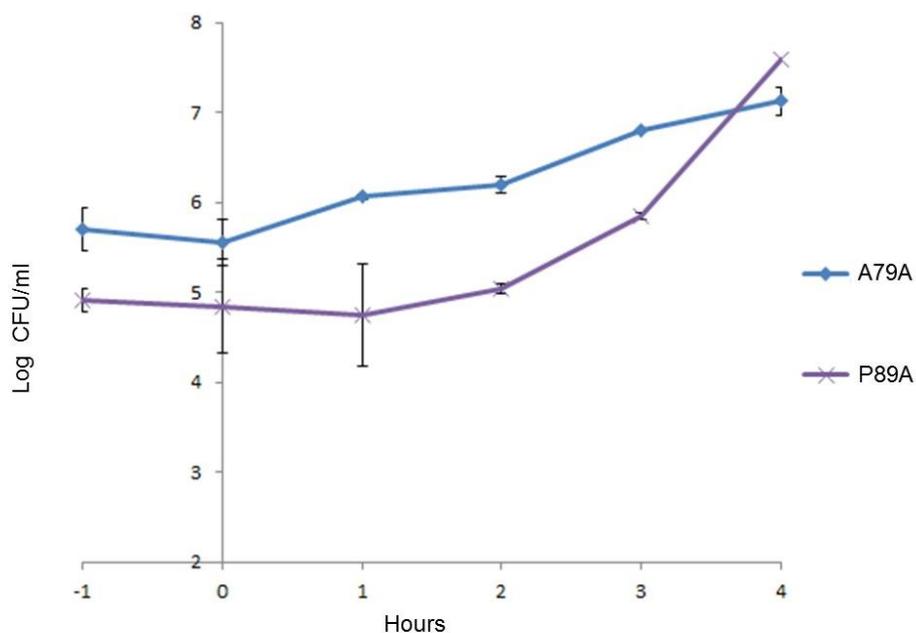
Figure 3.1 Bacteria viability study with alanine scanning library members, RnpA[A79A] and RnpA[P89A]. RnpA[A79A] (A79A) is a wild-type phenotype control. RnpA[P89A] (P89A) caused growth defect. Error bar was shown in black on the curve at each time point.

To investigate the *in vitro* RNase P ptRNA processing and mRNA degradation abilities of

RnpA[P89A], recombinant his-tagged (pET30 Ek-LIC) wild-type RnpA[A79A] and RnpA[P89A] proteins

were expressed and purified as described in the previous chapter. RNase P tyrosine ptRNA

(ptRNA[Tyr]) processing and degradation of *spa* mRNA were monitored for each mutant as

previously described[3,4]. RnpA[A79A] (A79A) was used as a positive control, which exhibited efficient

activities in both assays **(Fig. 3.2)**. The inhibitor RNPA2000 was used as a negative control, which

exhibited sufficient inhibition of both RNA processing activities of wild-type RnpA. The assay results of RnpA[P89A] showed that this mutant has no effect on ptRNA processing activity compared to the wild-type RnpA, but there was a dramatic *loss* in its mRNA degradation activity (**Fig. 3.2**). It suggested that residue proline 89 is essential for RnpA's mRNA degradation activity *in vitro*. This is the first time a site-directed mutant of RnpA has been shown to maintain wild-type level ptRNA processing activity but reduced mRNA degradation activity.
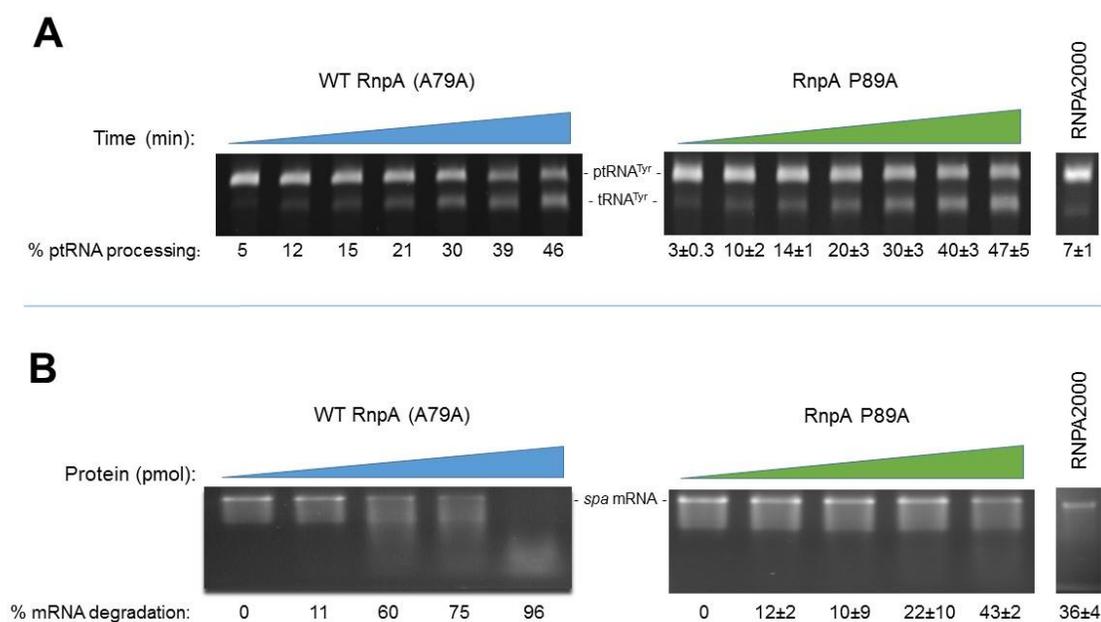
Figure 3.2 *In vitro* activity profiles of wild-type RnpA[A79A] and RnpA[P89A]. RnpA[A79A] (A79A) is a wild-type phenotype control. RnpA[P89A] (P89A) lost *in vitro* mRNA degradation activity. A. Time course measurements (0, 2.5, 5, 10, 20, 40 and 60 min) of percentage ptRNA conversion to mature tRNA. B. Protein titration assays (0, 0.625, 1.25, 2.5 and 5 pmol) to determine percentage spa mRNA degradation. The small molecule RnpA inhibitor, RNPA2000 used as negative control.

RnpA[P89A]'s effects on ptRNA processing and mRNA turnover were assessed *in cellulo*. RnpA proteins were over-expressed within cells under control of ATc as described above for bacteria viability. Cellular ptRNA[Tyr] and tRNA[Tyr] levels at different time points post-induction were evaluated by northern blotting as previously described[3]. Cellular mRNA turnover profiles were

produced by transcription arrest and subsequent *S. aureus* microarrays as previously described[3-7].

Results of cellular ptRNA processing showed no differences in the ptRNA^{Tyr} and tRNA^{Tyr} levels between cells over-expressing RnpA^{A79A} and RnpA^{P89A}, whereas RNAP2000 treated wild-type cells showed poly-ptRNA^{Tyr} accumulation (**Fig. 3.3A**). On the other hand, results of cellular mRNA turnover profiles showed significant accumulation of mRNA species in RnpA^{P89A} over-expressing cells comparing to those over-expressed RnpA^{A79A} (**Fig. 3.3B**). The retention of ptRNA processing but loss of cellular mRNA degradation ability in RnpA^{P89A} over-expressing cells indicated that residue proline 89 is essential for RnpA's mRNA degradation activity but not ptRNA processing in cells.
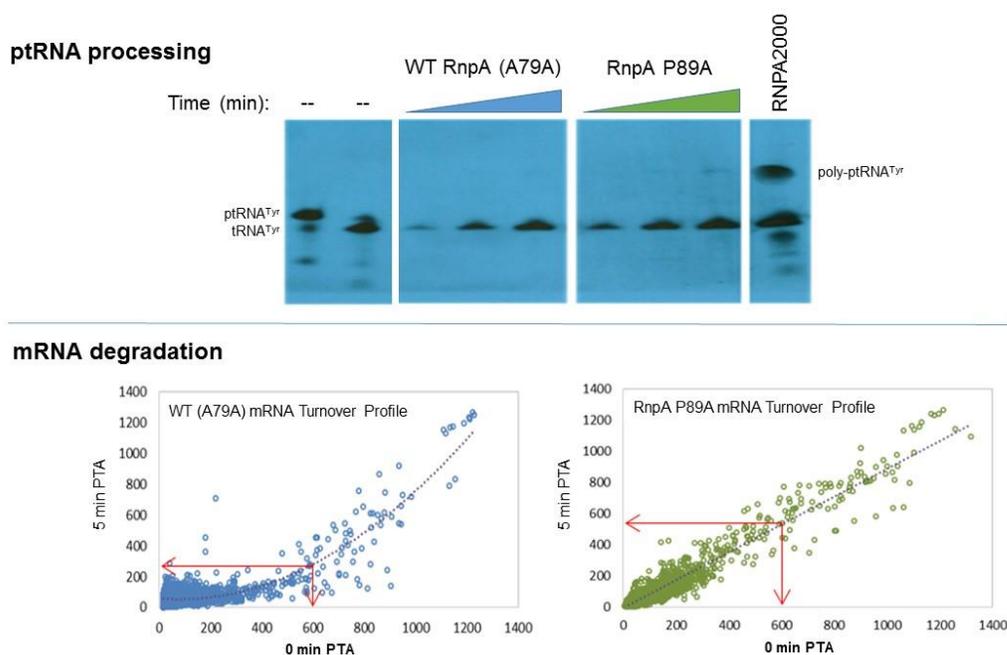
Figure 3.3 Cellular activity of wild-type RnpA^{A79A} and RnpA^{P89A} over-expressing cells. RnpA^{A79A} (A79A) is a wild-type phenotype control. RnpA^{P89A} (P89A) lost *in vivo* mRNA degradation activity. A) The ptRNA^{Tyr} and tRNA^{Tyr} levels at 0, 2, and 4hr post induction by northern blotting. Mobility of synthesized ptRNA^{Tyr} and tRNA^{Tyr} was shown; RNAP2000 was used as a control. B) The mRNA turnover profiles by microarray. Red arrow indicates titer of mRNA species at 0min (x-axis) and 5 min (y-axis) PTA.

Taken together, the assays suggested that residue proline 89 is indispensable for mRNA degradation and cell viability. The structural and biophysical basis of this observation remains unclear, yet this information may provide strategies to design and optimize small molecules that interrupt RnpA-mediated mRNA degradation and suppress *S. aureus*. Therefore, in this chapter, I report thermal shift, crystallization, and circular dichroism studies of RnpA^{P89A}.

3.2   Materials and Methods

3.2.1   Tag-free RnpA^{P89A} mutant production

Tag-free RnpA^{P89A} vector (pHis2-parallel::rnpA^{P89A}) was generated by site-directed mutagenesis with the wild-type tag-free RnpA vector (pHis2-parallel::rnpA). Primers were designed to replace the proline codon CCA at position 89 with the alanine codon GCA used in the alanine scanning library.   Forward   primer   sequence   was ATTGTAATAGCAAGACAG<u>GCA</u>GCTAAAGATATGACCA, and Reverse Primer sequence was TGGTCATATCTTTAGC<u>TGC</u>CTGTCTTGCTATTACAAT. The primers were ordered from Integrated DNA Technologies (IDT). Tag-free RnpA^{P89A} vector was then generated by PCR with the same parameters described in Section *2.2.1*. The cloned vector sequence was verified by sequencing analysis at the Purdue Genomics Core Facility. Plasmid DNA was transduced into competent *Escherichia coli* BL21 (DE3) Rosetta cells (Novagen). Expression and purification protocols were the same as that of tag-free wild-type RnpA (Refer to Section *2.2.2*). Detailed macromolecule information was summarized in **Table 3.1**. Purified proteins were used in protein stability and secondary structure studies.

Table 3.1 Macromolecule information of RnpA[P89A]

Source Organism	S. aureus UAMS-1	
Expression vector	pET30 Ek-LIC	pHis2-parallel
Expression host	E. coli BL21(DE3)	E. coli BL21(DE3)
Molecular weight (Da)	15868.51	13771.29
Protein Sequence	MSGSHHHHHHGSSGENLYFQSL MEKAYRIKKNADFQRIYKKGHS VANRQFVVYTCNNKEIDHFRLGI SVSKKLGNAVLRNKIKRAIRENF KVHKSHILAKDIIVIARQAAKDM TTLQIQNSLEHVLKIAKVFNKKIK	GAMGWEKAYRIKKNADFQRIYK KGHSVANRQFVVYTCNNKEIDH FRLGISVSKKLGNAVLRNKIKRAI RENFKVHKSHILAKDIIV IARQAAKDMTTLQIQNSLEHVLK IAKVFNKKIK

3.2.2 Protein stability analysis by differential scanning fluorimetry (DSF)

Thermal melt curves were determined using a modified method from Mashalidis et al[8]. Concentrations of both purified tag-free wild-type RnpA and tag-free RnpA[P89A] protein samples were determined with a NanoDrop™ One Microvolume UV-Vis Spectrophotometer and used in a final concentration of 10 µM in wells. Buffer conditions with final concentration of 200 mM sodium chloride, and 150 mM sodium phosphate buffer at pH 7.4 were used. 2% DMSO was added for the purpose of future comparison with ligand-protein complexes screening, and 10x SYPRO Orange dye (Invitrogen) was used as the fluorescent reagent. The reactions were prepared in 50-µL scale and each condition was assayed in triplicate. Fluorescence intensity was measured and derivative was calculated as the protein samples were heated from 25 °C to 95 °C at a ramp rate of +0.5 °C/s in continuous mode with a StepOnePlus™ Real-Time PCR System (Applied Biosystems™). The melting curves were subsequently analysed using the Protein Thermal Shift™

software v1.3. Boltzman T_m values were used to calculate the average T_m of the triplicates for each protein.

3.2.3   Crystallization of RnpA^{P89A}

Purified his-tagged RnpA^{P89A} (pET30 Ek-LIC::*rnpA*^{P89A}) was provided by Jennifer Colquhoun from the Dunman group. The protein sample was concentrated by centrifugation at 14,000 g in an Amicon Ultra centrifugal filter (cutoff MW 3k Da) to 10 mg/mL. A 24-well screen plate was set up manually in the same conditions of wild-type his-RnpA first optimization screen with hanging-drop vapor-diffusion method at room temperature (**Table 2.2** and **Table 3.2**). The crystal produced from the screen was flash-cooled by direct plunging into liquid nitrogen with no cryo-protectant. Detailed macromolecule and crystallization information were summarized in **Table 3.1** and **Table 3.2**.

Data collection and structure determination were performed with the same resources and technology as described in the previous chapter (Section *2.2.4*). The coordinate of his-RnpA^{P89A} has been deposited and detailed data collection and refinement information were provided in **Table 3.3**. Wild-type/mutant structural alignment and figures and were prepared using the PyMOL molecular graphics system (Schrodinger, LLC) and Coot. The electrostatic surface representation of proteins was prepared by the APBS plugin in PyMOL

.

Table 3.2 Crystallization of His-RnpA^{P89A}

	His-RnpA
Method	Hanging-drop vapor diffusion
Temperature (K)	298.15
Protein concentration (mg/ml)	22
Protein buffer	20 mM Tris pH 7.4, 200 mM NaCl
Plate type	VDXm™ 24-well Plate with sealant
Volume and ratio of drop (µl)	1 : 1
Volume of reservoir (µl)	1250
Plate type	VDXm™ 24-well Plate with sealant
Volume of reservoir	1250 µl
Screen Conditions	1)0.8 M LiCl, 0.1 M Tris: HCl, pH 8.5, 26-40 % (w/v) PEG 4,000 (in 8 wells) 2).2 M Ammonium Citrate Dibasic, 14-28 % (w/v) PEG 3,350 (in 8 wells) 3)0.1 M Imidazole-HCl pH 8, 700-1400 mM Sodium Citrate tribasic dehydrate (in 8 wells)
Condition of the best diffracted crystal	0.1 M Imidazole-HCl, pH8, 1.3 M Sodium Citrate tribasic dehydrate

Table 3.3 Data collection and processing of His-RnpA[P89A]

Diffraction source	Beamline 23-ID-D, APS
Wavelength (Å)	1.3
Temperature (K)	100
Detector	Dectris Pilatus3 6M
Crystal-detector distance (mm)	300
Rotation range per image (°)	1
Total rotation (°)	180
Exposure time per image (s)	1
Space group	P1
a, b, c (Å)	40.39, 43.47, 45.84
α, β, γ (°)	91.11, 114.83, 110.95
Mosaicity (°)	0.4-0.6
Resolution range (Å)	50.000–1.660 (1.720–1.660)
Total No. of reflections	52284
No. of unique reflections	30678
Completeness (%)	95.200 (92.200)
Redundancy	1.800 (1.800)
$\langle I/\sigma (I) \rangle$	6.800
$CC_{1/2}$	0.996 (0.573) †
$R_{\text{r.i.m.}}$	0.059 (0.725)
$R_{\text{p.i.m.}}$	0.042 (0.513)
Overall B factor form Wilson plot (Å^2)	26.520

3.2.4   Protein secondary structure study by circular dichroism

Purified tag-free wild-type RnpA and tag-free RnpA[P89A] protein samples were dialyzed into a CD

buffer of 50 mM sodium phosphates, 200 mM potassium fluoride at pH 7.4. Protein concentrations

were determined with a NanoDrop™ One Microvolume UV-Vis Spectrophotometer, adjusted with

CD buffer and used at a final concentration of 12.5 µM. The experiment was in 300-µL scale and

each protein was analyzed in triplicate at 298.15 K in a 1 mm path length quartz cell (Starna Cells, Inc) using a Jasco J-1500 Circular Dichroism Spectrophotometer. Blank CD buffer was used to measure baseline signals, and each spectra was obtained with baseline subtracted. Spectra were acquired from 185 to 260 nm with data pitch of 0.5nm and integration time of 2s[9]. Accumulation was set at 3, and integrated spectra were smoothed with Savitsky-Golaywith smoothing method. Raw CD data (in millidegree) were converted to mean residue ellipticity, [θ] x 10^{-3} (degree x cm^2/dmole), and analyzed with online server BestSel for protein secondary structure compositions[10,11].

3.3   Result and Discussion

3.3.1   Protein thermal stability analysis by differential scanning fluorimetry (DSF)

Thermal stability is fundamental for the folding and function of a protein[12], and mutation that causes amino acid substitution has been found to affect protein thermal stability[13-16]. Differential scanning fluorimetry (DSF) is a fast and efficient fluorescence-based protein thermal shift assay that measures thermal stability of a protein[17]. It uses a fluorescent dye to monitor the denaturation of the target protein, and the subsequently determined melting temperature (T_m) represents the protein thermal stability[17].

To investigate the activity change of RnpA[P89A], its thermal stability was first assessed in comparison with wild-type RnpA (WT) using DSF. Tag-free versions of both WT and RnpA[P89A] were prepared with the same construct (pHis2-parallel) and identical expression and purification methods to rule out variants that can cause differences in experiment result. SYPRO-orange dye which fluoresces (ex: 470 nm /em: 570 nm) upon binding the hydrophobic regions of a protein was used[18]. Conditions were optimized with final concentrations of 10 µM of protein, and buffer with 150 mM sodium phosphate, 200 mM sodium chloride at pH 7.4 was used. The temperature

range of 25 °C to 95 °C was recommended by literature[17,18]. The fluorescence emission was measured every 0.3 °C to yield the protein melting curves.

T_m value of a protein was determined with the Protein thermal shift[TM] software, which fits the fluorescent data to a Boltzmann Sigmoidal curve. The Boltzmann T_m of WT was 52.77 ± 0.07 °C, and T_m of RnpA[P89A] was 50.82 ± 0.16 °C (**Fig. 3.4**). The substitution of proline 89 to alanine destabilized the protein by ~2.0 °C (ΔT_m). It was hypothesized that the destabilization was caused by alterations in protein structure, which will also explain the activity change of RnpA[P89A]. Structural studies are therefore required to test this hypothesis.
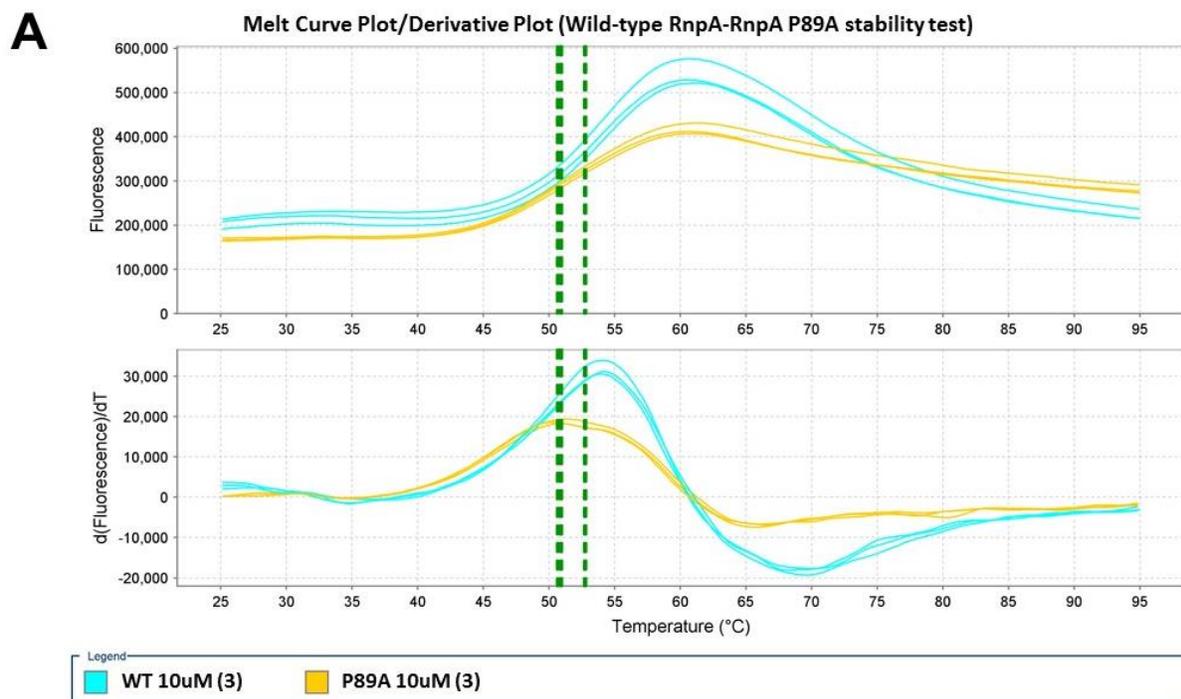

Figure 3.4 Differential scanning fluorimetry experiment for *S. aureus* wild-type RnpA and RnpA[P89A]. *S. aureus* wild-type RnpA (WT) is more stable than RnpA[P89A] (P89A) by ~2.0 °C. A) Fluorescence-temperature plot/ Melting curve (Upper panel) and fluorescence derivative-temperature plot (lower panel) generated by protein thermal shift software. Green dash lines indicated the determined T_ms based on plot analysis. B) Melting curve with error bars. Black dash lines indicates T_ms based on plot analysis.

Figure 3.4 continued

B

WT $T_m = 52.80 \pm 0.04\ °C$

P89A $T_m = 50.82 \pm 0.09\ °C$
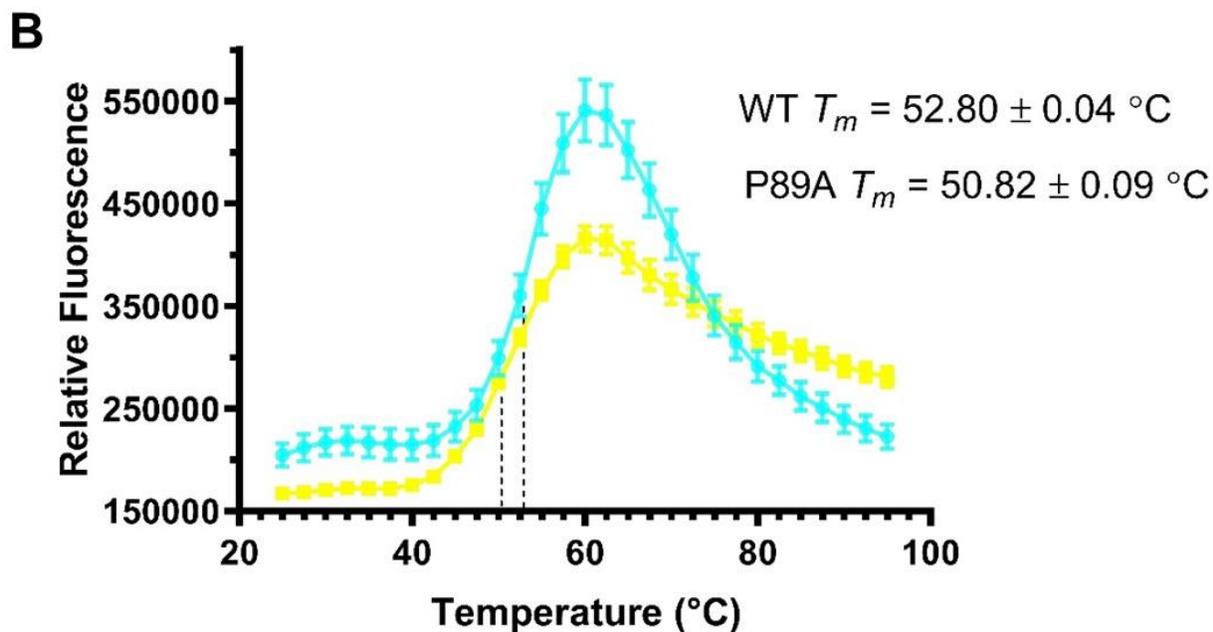
3.3.2 Crystallization of RnpA[P89A]

Crystallization is a common way to examine the structure of a protein. Based on the experience of

crystallizing the wild-type his-RnpA (His-WT), crystallization of his-tagged (His) RnpA[P89A] was

attempted. As both proline in the wild-type and alanine in RnpA[P89A] are surface hydrophobic

residues, 3D structures of both proteins were not expected to be dramatically different even though

there were differences in activities and thermal stability. Therefore, broad screen condition(s) that

yielded the His-wild-type crystals was used with variation for initial crystallization attempt for

His-RnpA[P89A], which produced an initial screen plate for His- RnpA[P89A] with the same layout as

the optimization screen plate of the wild-type. There were three initial hit conditions for the wild-

type: Condition 1: 0.8 M LiCl, 0.1 M Tris-HCl, pH 8.5, 32 % (w/v) PEG 4,000. Condition 2: 0.2

M Ammonium Citrate Dibasic, 20 % (w/v) PEG 3,350. Condition 3: 0.1 M Imidazole: HCl, pH 8,

1 M sodium citrate tribasic dihydrate (**Table 2.2**). The 24-well plate was divided into 3 sections,

and each section has 2 columns and 4 rows, in total of 8 wells. The first section has Condition1 in

variation with % (w/v) PEG 4,000 ranging 26-40 in steps of 2%. The second section has Condition 2 in variation with % (w/v) PEG 3,350 ranging 14-28 in steps of 2%. The third section has Condition 3 in variation with sodium citrate tribasic dihydrate concentration ranging 700-1400 mM in steps of 200 mM. There were two crystals found in the well with 0.1 M Imidazole-HCl pH 8, 1.3 M Sodium Citrate tribasic dehydrate (**Table 3.2**). The crystal was in a morphology of rhombus plates clustered on a thick rod (**Fig. 3.5**). Because of the clustering, the edges of single crystal plates had to be broken off for harvest and diffraction in order to avoid severe anisotropy which impedes data processing and structure refinement[19]. The crystal X-ray diffraction data was processed to 1.66 Å resolution which still retained good data quality. It was shown that His-RnpA^{P89A} crystallized in space group P1. Detailed information of data collection and processing was reported in **Table 3.3**.

Figure 3.5 His-tagged RnpA^{P89A} crystals prior to harvest.

The solved His-RnpA^{P89A} model showed two monomers in the asymmetric unit. Chain A monomer interacts with an inverted chain B monomer at the edge of the β sheet (**Fig. 3.6**), similar to that of the wild-type his-RnpA chain A and B. This observation does not necessarily suggest that His-RnpA^{P89A} forms dimers, as no dimerization was reported during protein production by our collaborator. Nonetheless, previous studies did show wild-type RnpAs from *B. subtilis* and *T.*

maritima had tendencies to dimerize[20,21], specific native gel and/or SEC may need to be performed in the future if relevant. The high quality of the His-RnpA[P89A] crystal and its data allowed this protein model to show more residues and secondary structures. Residues -5 to 112 of Chain A and Residues -5-117 of Chain B were modelled. Distinct from the wild-type crystal structure, an α-helix (α_0) initiated by the tag residue Leu-3 and extended to the protein N-terminal residue Tyr7 was presented the RnpA[P89A] structural model in addition to the secondary structure sequence of α_1-β_1-β_2-β_3-α_2-η_1-β_4-η_2-α_3 that was found in both the wild-type and RnpA[P89A] (**Fig. 3.6**). Side chain of a few surface residues (ChainA: K38, K51, K52, R64, K71, K74, K107; ChainB: R8, K20, K38, K51, K52, R64, K71, K114) were truncated because of missing electron densities. 99% of residues are in the Ramachandran favored region and 1% in the allowed region.
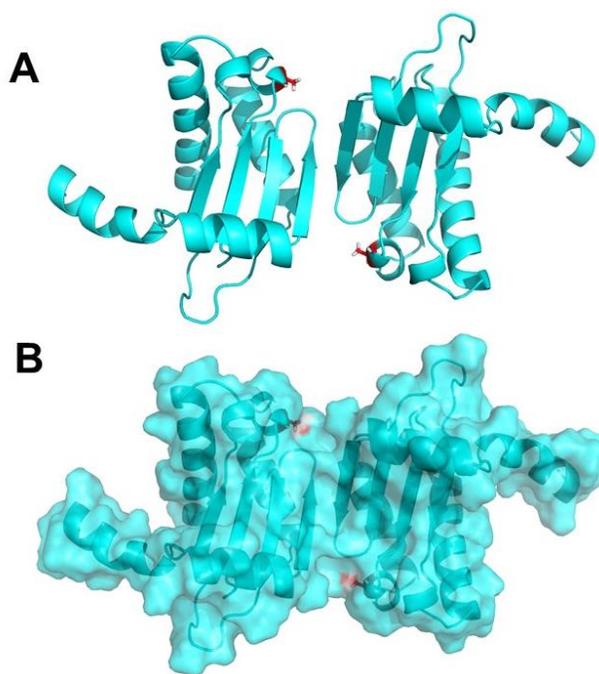

Figure 3.6 His-tagged RnpA[P89A] crystal structure at 1.66 Å resolution. A. Cartoon representation of RnpA[P89A] monomers in an asymmetric unit. B. Surface representation of RnpA[P89A] monomers in an asymmetric unit.

The tertiary structures of RnpA[P89A] were highly similar to that of the wild-type. Crystal structure of RnpA[P89A] (Chain B residues 6-117) was aligned with the tag-free wild-type RnpA crystal structure (PDB ID 6D1R), and an RMSD of 0.404 Å was found between them (**Fig. 3.7**). The two structures were nearly superimposable, and only slight variations were shown for residues K38 in the loop region, L53-V57 at the turn, H76 at the end of 3_{10} helix η_1, K91-L95 the end of 3_{10} helix η_2 and the loop connected to the start of α_3, and F112 and N113 at the C terminal. The Cα backbone of proline 89 overlaid perfectly with that of the mutant alanine 89, which indicate that mutation at proline 89 did not cause any significant changes in the 3D conformation of RnpA. This result was also supported by the structural study of *B. subtilis* RnpA F107W/P39A/P90A mutant, which showed high similarity to the wild-type protein[22]. This study suggested that the differences in protein activities and thermal stability between the wild-type and RnpA[P89A] were not resulted from protein structural changes.
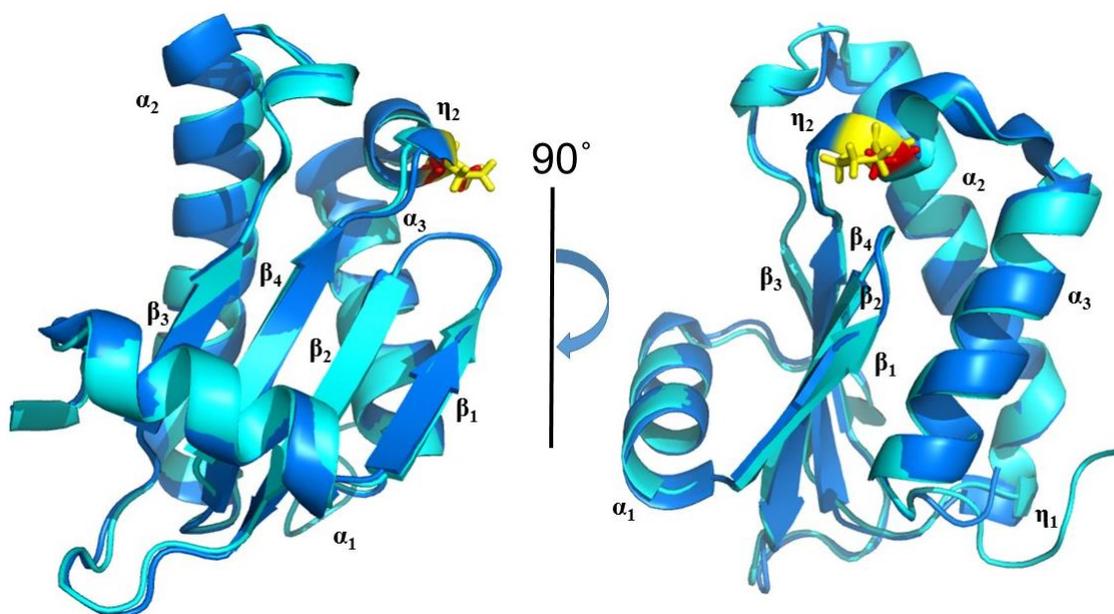

Figure 3.7 His-tagged *S. aureus* RnpA[P89A] overlaid onto the wild-type *S. aureus* RnpA structure. RnpA[P89A] (Chain B, residues 6-117, cyan) overlaid onto the wild-type *S. aureus* RnpA structure (PDB: 1D6T, marine) with an RMSD = 0.404 Å. Residue proline 89 (yellow) in the wild-type aligned well with the mutated residue alanine (red) in RnpA[P89A].

Table 3.4 Structure solution and refinement of His-RnpA^{P89A}

PDB entry		6OV1
Resolution range (Å)		40.7610–1.6570 (1.6985–1.6571)
Completeness (%)		94.9
σ cutoff		$F > 1.970\sigma(F)$
No. of reflections, working set		27203 (1779)
No. of reflections, test set		2003 (135)
Final R_{cryst}		0.189 (0.3100)
Final R_{free}		0.233 (0.3111)
No. of non-H atoms	Protein	1898
	Ligand	0
	Solvent	127
	Total	2025
R.m.s. deviations	Bonds (Å)	0.006
	Angles (˚)	0.837
Average B factors (Å^2)	Protein	40.43
	Solvent	45
Ramachandran plot	Most favored (%)	99.16
	Allowed (%)	0.84

3.3.3   Protein secondary structure study by circular dichroism

High-resolution crystal structures that are snapshots of the protein in a thermodynamically favored state are deemed more reliable in terms of protein backbone atomic positions, and are in better agreement with the experimental data when compared to solution state NMR structures[23,24]. Yet

NMR structures can provide additional information regarding proteins in a more physiologically relevant state[23]. In our study, even though crystallography showed no difference between the structures of WT and RnpA[P89A], this may be a function of the lowest energy stable state and packing of the crystal. Using circular dichroism (CD) we propose to analyze the secondary structure components of the protein and compare the WT and RnpA[P89A].

A CD buffer of 50 mM sodium phosphate, 200 mM potassium fluoride at pH 7.4 was chosen in the consideration of: 1) keeping consistency with the pH and high ionic concentrations used in protein activity and thermal shift assays (150 mM sodium phosphate, 200 mM sodium chloride at pH 7.4) 2) staying away from Cl^- that absorbs in circular dichroism below 195 nm. Protein absorption from 190 nm to 260 nm was measured for both protein and data was processed as previously described[9,11]. The processed data were subsequently submitted to online server BeStSel for secondary structure composition analysis[10]. The ellipticity-wavelength curves of WT and RnpA[P89A] showed similar absorption profile as extrema of both proteins appeared at 193 nm, 208 nm, and 223 nm, yet the ellipticities of RnpA[P89A] at the extrema were relatively greater than those of WT (**Fig. 3.8A**). According to the analysis by BeStSel, α helices represented 21.4% of the wild-type and 22.9% of RnpA[P89A], which suggested that RnpA[P89A] has about 1.7 more residues in the α-helices. The β strands represented 19.9% of the wild-type and 19.0% of RnpA[P89A], which indicated that WT has 1 more residue in β strands (**Fig. 3.8B**). However, these differences in α helices and β strands components were within errors of CD detection accuracy for alpha-beta mix protein (0.97 for α helices and 0.85 for β strand)[25,26]. Therefore, it is reasonable to conclude that the secondary structures of RnpA[P89A] were highly similar to that of the wild-type. This study corroborated the result of the crystallography study, and indicated that the differences in protein

activities and thermal stability between the wild-type and RnpA[P89A] were not resulted from protein structural changes.
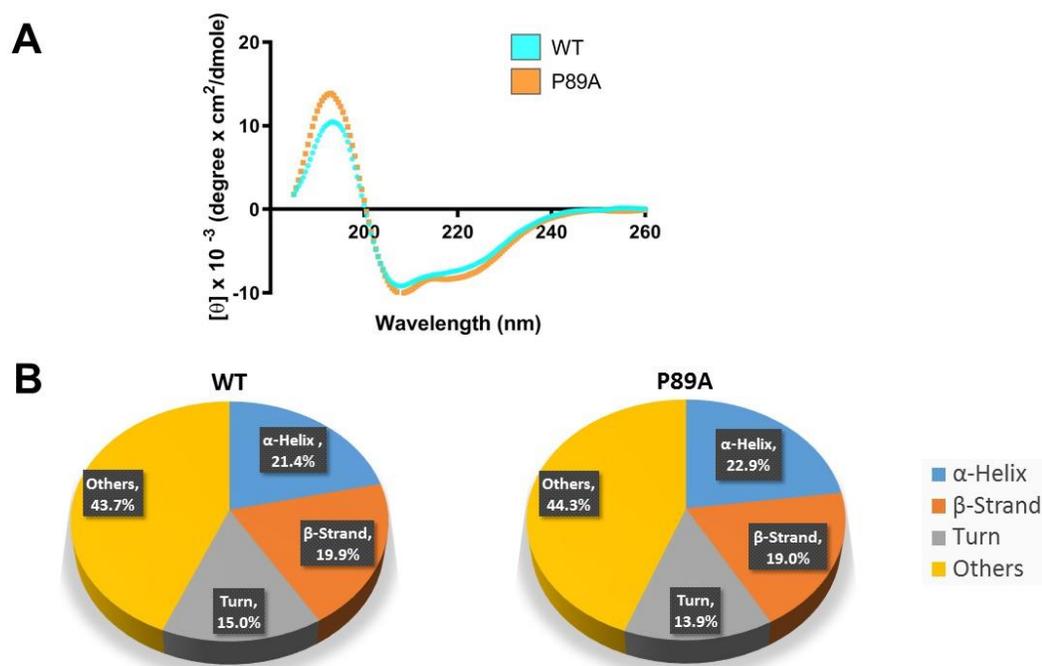

Figure 3.8 Secondary structure analysis for *S. aureus* wild-type RnpA and RnpA[P89A]. A) Elipticity-wavelength plot. B) Pie charts of secondary structure compositions.

3.3.4   Hypothesis for RnpA[P89A] activity and stability

Biological and biochemical studies by our collaborator showed that RnpA[P89A] has equivalent ptRNA processing activity as wild-type RnpA, but no mRNA degradation activity. Thermal shift assay done in our lab showed that RnpA[P89A] is ~ 2 ˚C less stable than the wild-type RnpA. Results of crystallography and CD studies indicated no secondary or 3D conformational structure differences were present between WT and RnpA[P89A]. A possible explanation for the observed change in thermal stability may be due to changes in the proteins free energy in unfolding. Protein free energy change of unfolding (ΔG) is the determinant of protein thermal stability. The larger and more positive ΔG is the more stable the protein is. Entropy of unfolding (ΔS) is a contributor

to ΔG, and ΔG is inversely related to ΔS. As observed in many cases, substituting any residue with a proline had a stabilizing effect on the protein[27-30]. Studies on these observations suggested that because proline's restrained conformation, it reduces protein backbone flexibility in solution, and thus lowers the configurational ΔS. The decrease of ΔS yields an increase of ΔG, which indicates the stabilization of a protein[29-31]. Consistent with this trend, RnpA as the WT protein that contains the proline is more thermally stable than the P89A counterpart. This theory can be tested using differential scanning calorimetry (DSC) which provides information of thermodynamic parameters of protein unfolding.

As for observed loss in mRNA degradation of RnpA[P89A], it is possibly caused by the differences in electrostatic potential surfaces of the two proteins. In the crystal structure of WT, side chain of Gln 88 at the beginning of η_2 extends across a cleft and forms a hydrogen bond (H-bond) with the backbone carbonyl of Lys 52 located on loop preceding α_2 (**Fig. 3.9,** marine sticks). Whereas upon substitution of Pro 89 with Ala, the side chain of Gln 88 no longer bridged the cleft to form an H-bond to the Lys 52 backbone carbonyl. Rather, it oriented itself back toward the η_2 and formed an H-bond with the side chain of Lys91 (**Fig. 3.9**, cyan sticks). These changes in side chain orientation and H-bonding introduced two main differences on the electrostatic potential surfaces around the mutation (residues 88-92). As shown in **Fig 3.10**, the WT has a smooth positively charged (colored in blue) surface over the hydrogen bonded Lys 52-Gln 88 region, whereas in RnpA[P89A] there is a partially negatively charged (colored in red) groove in-between. Additionally, in WT the surface over Lys 91 backbone and Asp 92 is nearly neutral (colored in white), whereas in the same region of RnpA[P89A], a negatively charged surface is found. This negatively charged surface is created potentially because higher electron density localized around the Asp 92 carbonyl, while its salt-bridge forming partner Lys 91 engages hydrogen bonding with the flipped over Gln 88. These

alterations in surface potential may interrupt RnpA[P89A]'s interaction with other protein partners in the mRNA degradosome or its substrate(s), thus may be the reason for its loss of mRNA degradation activity of RnpA.
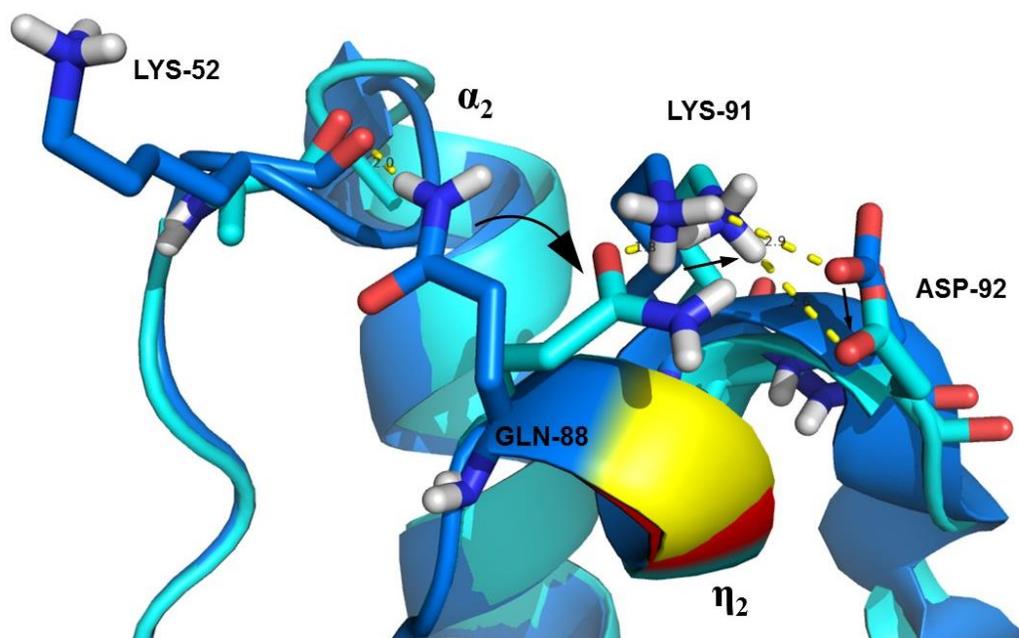

Figure 3.9 Side chain movements of relevant residues around mutation residue 89. WT (marine) and RnpA[P89A] (cyan) crystal structures were aligned, and relevant residues Lys 52, Gln 88, Lys 91 and Asp 92 were shown in sticks. Black arrows shows the side chain movement from their WT positions to RnpA[P89A] positions. Pro 89 in WT was colored in yellow, and Ala 89 in RnpA[P89A] was colored in red.
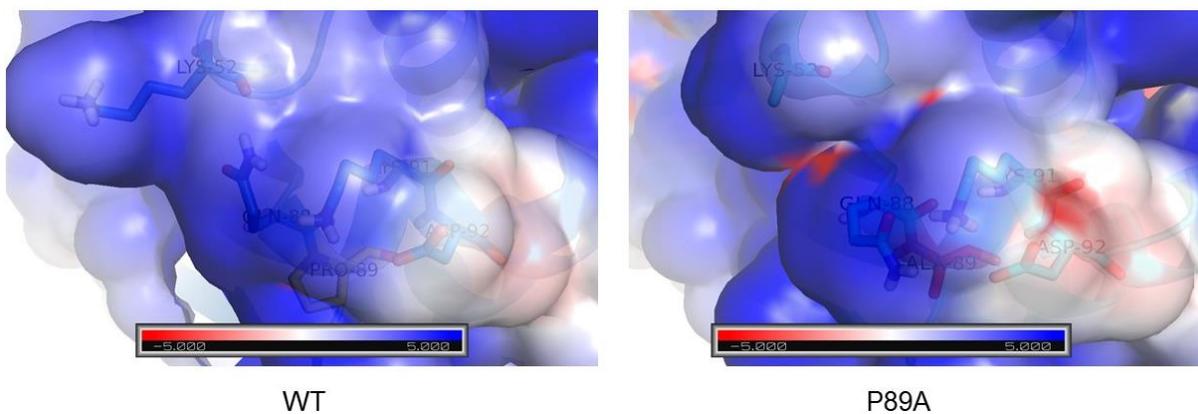

Figure 3.10 Close-ups of electrostatic potential surface around mutation residue 89 of wild-type RnpA and RnpA[P89A].Blue indicates positive potential, red indicates negative potential, and white indicates neutral.

3.4   Conclusion

Our collaborator has demonstrated that over expression of mutant RnpA^{P89A} reduces cell viability, and both in *vitro* and *in cellulo* studies suggested that RnpA^{P89A} retains ptRNA processing activity but loses mRNA degradation activity. To better understand the biophysical effects of this mutation we carried out thermal stability, crystallography, and CD studies for RnpA^{P89A}. Results showed thermal destabilization of RnpA^{P89A} compared to wild-type RnpA, but no secondary structural or 3D conformational differences were found between the two proteins. Therefore, it was concluded that the mutant's change of activity was not resulted from secondary or 3D conformational alteration. However, upon close structural examination, we propose that the activity change may be a result of electrostatic surface potential alteration of the protein based on a rearrangement of residue Gln88. The thermal destabilization is likely due to the loss of a proline, which increases entropy of unfolding, and lowers free energy.

Although the mechanism behind activity and thermal stability change of RnpA^{P89A} from wild-type RnpA requires further investigation, the identification of RnpA^{P89A} mutant and its effect was vital for future anti-*S. aureus* drug discovery targeting RnpA. It highlighted the essentiality of proline 89 to RnpA's mRNA degradation activity and *S. aureus* cell viability, which provided foundation for identifying small molecules that suppress *S. aureus* by interrupting RnpA-mediated mRNA degradation pathway.

References

1   Niranjanakumari, S. *et al.* Probing the architecture of the B. subtilis RNase P holoenzyme active site by cross-linking and affinity cleavage. *RNA* **13**, 521-535, doi:10.1261/rna.308707 (2007).

2   Gopalan, V., Baxevanis, A. D., Landsman, D. & Altman, S. Analysis of the functional role of conserved residues in the protein subunit of ribonuclease P from Escherichia coli. *J Mol Biol* **267**, 818-829, doi:10.1006/jmbi.1997.0906 (1997).

3   Eidem, T. M. *et al.* Small-molecule inhibitors of Staphylococcus aureus RnpA-mediated RNA turnover and tRNA processing. *Antimicrob Agents Chemother* **59**, 2016-2028, doi:10.1128/AAC.04352-14 (2015).

4   Olson, P. D. *et al.* Small molecule inhibitors of Staphylococcus aureus RnpA alter cellular mRNA turnover, exhibit antimicrobial activity, and attenuate pathogenesis. *PLoS Pathog* **7**, e1001287, doi:10.1371/journal.ppat.1001287 (2011).

5   Anderson, K. L. *et al.* Characterization of the Staphylococcus aureus heat shock, cold shock, stringent, and SOS responses and their effects on log-phase mRNA turnover. *Journal of Bacteriology* **188**, 6739-6756 (2006).

6   Anderson, K. L. *et al.* Characterizing the effects of inorganic acid and alkaline shock on the Staphylococcus aureus transcriptome and messenger RNA turnover. *Fems Immunol Med Mic* **60**, 208-250 (2010).

7   Roberts, C. *et al.* Characterizing the effect of the Staphylococcus aureus virulence factor regulator, SarA, on log-phase mRNA half-lives. *Journal of Bacteriology* **188**, 2593-2603 (2006).

8   Mashalidis, E. H., Sledz, P., Lang, S. & Abell, C. A three-stage biophysical screening cascade for fragment-based drug discovery. *Nat Protoc* **8**, 2309-2324, doi:10.1038/nprot.2013.130 (2013).

9   Greenfield, N. J. Using circular dichroism spectra to estimate protein secondary structure. *Nat Protoc* **1**, 2876-2890 (2006).

10   Micsonai, A. *et al.* BeStSel: a web server for accurate protein secondary structure prediction and fold recognition from the circular dichroism spectra. *Nucleic Acids Res* **46**, W315-W322, doi:10.1093/nar/gky497 (2018).

11   Micsonai, A. *et al.* Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy. *Proc Natl Acad Sci U S A* **112**, E3095-3103, doi:10.1073/pnas.1500851112 (2015).

12   Becktel, W. J. & Schellman, J. A. Protein stability curves. *Biopolymers* **26**, 1859-1877, doi:10.1002/bip.360261104 (1987).

13   Schaefer, C. & Rost, B. Predict impact of single amino acid change upon protein structure. *Bmc Genomics* **13** (2012).

14   De Cristofaro, R. *et al.* The natural mutation by deletion of Lys9 in the thrombin A-chain affects the pKa value of catalytic residues, the overall enzyme's stability and conformational transitions linked to Na+ binding. *FEBS J* **273**, 159-169, doi:10.1111/j.1742-4658.2005.05052.x (2006).

15   Lorch, M., Mason, J. M., Sessions, R. B. & Clarke, A. R. Effects of mutations on the thermodynamics of a protein folding reaction: implications for the mechanism of formation of the intermediate and transition states. *Biochemistry* **39**, 3480-3485 (2000).

16   Reva, B., Antipin, Y. & Sander, C. Predicting the functional impact of protein mutations: application to cancer genomics. *Nucleic Acids Res* **39**, e118, doi:10.1093/nar/gkr407 (2011).

17   Huynh, K. & Partch, C. L. Analysis of protein stability and ligand interactions by thermal shift assay. *Curr Protoc Protein Sci* **79**, 28 29 21-14, doi:10.1002/0471140864.ps2809s79 (2015).

18   Niesen, F. H., Berglund, H. & Vedadi, M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat Protoc* **2**, 2212-2221, doi:10.1038/nprot.2007.321 (2007).

19   Strong, M. *et al.* Toward the structural genomics of complexes: crystal structure of a PE/PPE protein complex from Mycobacterium tuberculosis. *Proc Natl Acad Sci U S A* **103**, 8060-8065, doi:10.1073/pnas.0602606103 (2006).

20   Buck, A. H., Dalby, A. B., Poole, A. W., Kazantsev, A. V. & Pace, N. R. Protein activation of a ribozyme: the role of bacterial RNase P protein. *EMBO J* **24**, 3360-3368, doi:10.1038/sj.emboj.7600805 (2005).

21   Kazantsev, A. V. *et al.* High-resolution structure of RNase P protein from Thermotoga maritima. *Proc Natl Acad Sci U S A* **100**, 7497-7502, doi:10.1073/pnas.0932597100 (2003).

22   Daniels, K. G. *et al.* Ligand concentration regulates the pathways of coupled protein folding and binding. *J Am Chem Soc* **136**, 822-825, doi:10.1021/ja4086726 (2014).

23   Sikic, K., Tomic, S. & Carugo, O. Systematic comparison of crystal and NMR protein structures deposited in the protein data bank. *Open Biochem J* **4**, 83-95, doi:10.2174/1874091X01004010083 (2010).

24   Clore, G. M. & Gronenborn, A. M. New methods of structure refinement for macromolecular structure determination by NMR. *Proc Natl Acad Sci U S A* **95**, 5891-5898, doi:10.1073/pnas.95.11.5891 (1998).

25   Manavalan, P. & Johnson, W. C., Jr. Variable selection method improves the prediction of protein secondary structure from circular dichroism spectra. *Anal Biochem* **167**, 76-85 (1987).

26   Venyaminov, S. & Vassilenko, K. S. Determination of protein tertiary structure class from circular dichroism spectra. *Anal Biochem* **222**, 176-184 (1994).

27   Hardy, F. *et al.* Stabilization of Bacillus stearothermophilus neutral protease by introduction of prolines. *FEBS Lett* **317**, 89-92 (1993).

28   Nicholson, H., Tronrud, D. E., Becktel, W. J. & Matthews, B. W. Analysis of the effectiveness of proline substitutions and glycine replacements in increasing the stability of phage T4 lysozyme. *Biopolymers* **32**, 1431-1441, doi:10.1002/bip.360321103 (1992).

29   Watanabe, K., Chishiro, K., Kitamura, K. & Suzuki, Y. Proline residues responsible for thermostability occur with high frequency in the loop regions of an extremely thermostable oligo-1,6-glucosidase from Bacillus thermoglucosidasius KP1006. *The Journal of biological chemistry* **266**, 24287-24294 (1991).

30   Bogin, O. *et al.* Enhanced thermal stability of Clostridium beijerinckii alcohol dehydrogenase after strategic substitution of amino acid residues with prolines from the homologous thermophilic Thermoanaerobacter brockii alcohol dehydrogenase. *Protein Sci* **7**, 1156-1163, doi:10.1002/pro.5560070509 (1998).

31   Watanabe, K., Masuda, T., Ohashi, H., Mihara, H. & Suzuki, Y. Multiple proline substitutions cumulatively thermostabilize Bacillus cereus ATCC7064 oligo-1,6-glucosidase. Irrefragable proof supporting the proline rule. *Eur J Biochem* **226**, 277-283 (1994).

FINAL REMARKS

This thesis highlighted with the structural studies of wild-type *S. aureus* RnpA and mutant RnpA^{P89A} that has only partial activity can be useful tools in structure-based drug discovery for multi-drug resistant S. aureus treatment. Such drug discovery has been initiated in our lab and our collaborator Dr. Dunman's lab with both a fragment-based drug design screening paradigm and a whole-cell synergistic screening approach. The screening hits will be validated and prioritized using methods including crystallography, and the yielded binding information between hits and RnpA will be further used for hit optimization. It is expected that leads and drug candidates will be discovered in future studies following this general process.

PUBLICATIONS

1. **Ha, L.**, Colquhoun, J., Noinaj, N., Das, C., Dunman, P. M., and Flaherty, D. P.. Crystal structure of the ribonuclease-P-protein subunit from Staphylococcus aureus. *Acta Crystallogr F Struct Biol Commun* **74**, 632-637, doi:10.1107/S2053230X18011512 (2018).

2. Colquhoun, J. M., **Ha, L.**, Beckley, A., Meyers, B., Flaherty, D. P., Dunman, P. M.. Identification of Small Molecule Inhibitors of Staphylococcus aureus RnpA. *Antibiotics (Basel)* **8**, doi:10.3390/antibiotics8020048 (2019).