



2018

## PROTEIN ENGINEERING IN THE STUDY OF PROTEIN LABELING AND DEGRADATION

Xinyi Zhang

University of Kentucky, xinyi.zhang@uky.edu

Digital Object Identifier: <https://doi.org/10.13023/ETD.2018.075>

[Right click to open a feedback form in a new tab to let us know how this document benefits you.](#)

---

### Recommended Citation

Zhang, Xinyi, "PROTEIN ENGINEERING IN THE STUDY OF PROTEIN LABELING AND DEGRADATION" (2018). *Theses and Dissertations--Chemistry*. 93.  
[https://uknowledge.uky.edu/chemistry\\_etds/93](https://uknowledge.uky.edu/chemistry_etds/93)

This Doctoral Dissertation is brought to you for free and open access by the Chemistry at UKnowledge. It has been accepted for inclusion in Theses and Dissertations--Chemistry by an authorized administrator of UKnowledge. For more information, please contact [UKnowledge@lsv.uky.edu](mailto:UKnowledge@lsv.uky.edu).

## **STUDENT AGREEMENT:**

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

## **REVIEW, APPROVAL AND ACCEPTANCE**

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's thesis including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Xinyi Zhang, Student

Dr. Yinan Wei, Major Professor

Dr. Mark A. Lovell, Director of Graduate Studies

PROTEIN ENGINEERING IN THE STUDY OF PROTEIN LABELING AND  
DEGRADATION

---

DISSERTATION

---

A dissertation submitted in partial fulfillment of the  
requirements for the degree of Doctor of Philosophy in the College  
of Arts and Sciences at the University of Kentucky

By

Xinyi Zhang

Lexington, Kentucky

Director: Dr. Yinan Wei, Professor of Chemistry

Lexington, Kentucky

2018

Copyright © Xinyi Zhang 2018

## ABSTRACT OF DISSERTATION

### PROTEIN ENGINEERING IN THE STUDY OF PROTEIN LABELING AND DEGRADATION

Proteins are large macromolecules that play important roles in nature. With the development of modern molecular biology techniques, protein engineering has emerged as a useful tool and found many applications in areas ranging from food industry, environmental protection, to medical and life science. Biomimetic membrane incorporates biological elements, such as proteins, to form membranes that mimic the high specificity and conductance of natural biological membranes. For any application involving the usage of proteins, the first barrier is always the production of proteins with sufficient stability, and the incorporation of proteins into the artificial matrix. This thesis contains two major parts, the first part is focused on the development and testing of method to immobilize active enzymes. The second part is devoted to study the degradation of membrane proteins in *E. coli* cells.

In the immobilization study, Pyrophosphatase (PpaC) was chose as a model enzyme. A dual functional tag consist of histidine and methionine has been developed, in which histidine is used for purification while methionine is metabolically replaced with azidohomoalanine (AHA) for immobilization. We found that the addition of the tag and the incorporation of AHA did not significantly impair the properties of proteins, and the histidine–AHA tag can facilitate protein purification, immobilization, and labeling. This tag is expected to be useful in general for many proteins.

Degradation of soluble protein has been well characterized, but the membrane protein degradation process remains elusive. SsrA tag is a well-known recognition sequence for soluble protein degradation, which marks prematurely terminated protein products translated from damaged mRNA. SsrA tagged membrane proteins was found to be substrate of a cytosolic protease complex ClpXP, which mediated complete degradation.

Keywords: protein engineering, protein tag, immobilization, degradation

Xinyi Zhang

---

Feb 14<sup>th</sup>, 2018

---

PROTEIN ENGINEERING IN THE STUDY OF PROTEIN LABELING AND  
DEGRADATION

By

Xinyi Zhang

Dr. Yinan Wei

---

Director of Dissertation

Dr. Mark A. Lovell

---

Director of Graduate Studies

Feb. 14<sup>th</sup>, 2018

---

## ACKNOWLEDGEMENTS

First of all, I would like to express my sincere and honest appreciation to my advisor, Dr. Yinan Wei. She is a knowledgeable, humble and amiable person, it is my fortune to be her student. Being an advisor, she is nice, patient and supportive, giving me guidance and encouraging me to solve the problems independently. Being a researcher, her knowledge, passion and attitude towards biochemistry always inspire me, make me want to be a better person “Stay hungry, Stay foolish”. Being a friend, she always cares about the others and willing to help them out. To me, she sets a great example for not only being an insightful scientist but also a decent person. I want to thank her for everything she has taught me during these years.

Also, I would like to thank my doctoral committee members. Dr. D. Allan Butterfield, for his suggestions and advises on graduate study. Dr. Robert B. Grossman, for inspiring me to think outside of the box. Dr. Harry LeVine, for his insightful advices on my research. They sacrificed their time and attended to my committee meetings, proof-reading my dissertation. Their advices always remind me to think thoroughly and deeply about the research. I appreciate Dr. Tianyan Gao, for taking her time and being my outside examiner.

Secondly, I would like to thank Dr. Dibakar Bhattacharyya for broaden my view to engineering and water technologies and Dr. Sebastián Hernández Sierra, Cassandra Porter for collaboration on layer-by layer assembled membranes. Dr. Isabel C. Escobar and Priyesh Wagh for working with biomimetic membranes. Dr. Brad Berron and Anuhya Gottipati for their help and training to use the fluorescence microscope.

I would like to thank my previous and current group members. Our postdoc Dr. Wei Lu, who taught me basic biochemistry techniques and gave me suggestions on the details of my experiments. It is always helpful to discuss with our previous group members, Dr. Linliang Yu, Dr. Meng Zhong, Dr. Cui Ye, Dr. Qian Chai, and Dr. Zhaoshuai Wang. And it is pleasant to work with our current lab mates, Thilini Abeywansha, Prasangi Rajapaksha, Isoiza Ojo, Lan Li, Ankit Pandeya, and our visiting scholar Ling Yang. I also want to thank Michael Bale and Kevin Kwan, who worked with me on the projects through their undergraduate study.

I want to convey my thankfulness to the ladies and gentlemen from main office and IBU office, who have offered me numerous help. Their kindness has created such a warm and friendly climate in Chemistry department.

I would like to thank Kentucky NSF EPSCoR RII Award 1355438, National Science Foundation (MCB 1158036 and CHE -1709381), and National Institute of Allergy and Infectious Diseases (1R21AI103717) for the funding support which makes my research possible.

Last but not least, I would like to thank my family and friends. Especially my parents, without their unconditional love, support, and trust, I could not be here and pursue my graduate study. I can hardly express how grateful I am to be their daughter, and my family is always my strength.



## TABLE OF CONTENTS

ACKNOWLEDGEMENTS .....	iii
LIST OF TABLES .....	ix
LIST OF FIGURES .....	x
LIST OF ABBREVIATIONS .....	xii
Chapter I Introduction.....	1
1.1 The brief history of protein engineering .....	1
1.2 Methods of Protein engineering .....	3
1.3 Protein functional tag .....	4
1.4 Special protein tag facilitated protein degradation.....	13
Chapter II A dual-functional tag facilitated protein labeling and immobilization.....	20
2.1 Introduction .....	20
2.2 Materials and methods .....	22
2.2.1 Plasmid construction.....	22
2.2.2 AHA incorporated into target protein .....	23
2.2.3 Protein purification .....	23
2.2.4 Bacteria growth curve .....	24
2.2.5 Protein biotinylation via click chemistry .....	24
2.2.6 Protein immobilization to alkyne agarose resin .....	25
2.2.7 Gel electrophoresis and Western Blot.....	25
2.2.8 Fluorescence spectroscopy.....	25
2.2.9 Pyrophosphatase activity assay .....	26
2.3 Results and discussion.....	27

2.3.1 Incorporation of AHA into the target proteins.....	27
2.3.2 Incorporation of a dual functional tag at the C-terminus of sfGFP .....	27
2.3.3 Addition of the GS3M4 tag to a pyrophosphatase PpaC .....	33
2.4 Conclusion.....	37
Chapter III Impairment of <i>S. aureus</i> PpaC activity by metal ion binding and dimer	
dissociation .....	39
3.1 Introduction .....	39
3.2 Materials and methods .....	40
3.2.1 Cloning and mutagenesis .....	40
3.2.2 Expression and purification .....	41
3.2.3 Pyrophosphatase activity assay.....	42
3.2.4 Circular Dichroism (CD) spectroscopy .....	43
3.2.5 Gel electrophoresis and size exclusion chromatography .....	43
3.3 Results and discussion.....	44
3.3.1 Effect of metal binding on activity .....	44
3.3.2 Effect of metal ion binding on protein stability .....	45
3.3.3 PpaC mutations that disrupt the dimer interface.....	47
3.3.4 Inhibitory effect of Mn <sup>2+</sup> and Mg <sup>2+</sup> .....	51
3.4 Conclusion.....	53
Chapter IV Structure effect on the degradation of MscS with ClpXP system.....	56
4.1 Introduction .....	56
4.2 Material and Method .....	58
4.2.1 Plasmid construction.....	58
4.2.2 Protein expression and purification .....	59
4.2.3 Degradation in cells .....	61

4.2.4 Degradation in reconstituted system .....	62
4.3 Result and discussion .....	64
4.3.1 The <i>ssrA</i> tag can facilitate degradation of full length <i>MscS</i> .....	64
4.3.2 <i>MscS</i> is membrane integrated .....	64
4.3.3 Detergent solubilized <i>MscS-ssrA</i> can be degraded using purified <i>ClpXP</i> .....	65
4.3.4 Development of a fluorescent polarization assay for real-time monitoring of degradation.....	66
4.3.5 <i>SspB</i> increases the rate of association of <i>ClpX</i> with substrate .....	68
4.3.6 Establishment of a calibration curve to measure degradation using FP .....	69
4.3.7 <i>S9C</i> degradation-Formation of disulfide bond did not affect degradation <i>in vivo</i> .....	71
4.3.8 Effect of substrate structure on the degradation with <i>ClpXP</i> .....	75
Chapter V Construction of <i>AcrB</i> nanodisc .....	76
5.1 Introduction .....	76
5.2 Material and methods .....	77
5.2.1 Cloning, expression and purification of <i>MSP1E3</i> .....	77
5.2.2 Assembly of nanodisc .....	78
5.2.3 Assembly of nanodisc containing <i>AcrB</i> trimer.....	79
5.2.4 Nanodisc detection and identification.....	79
5.3 Result and discussion .....	80
5.3.1 <i>MSP1E3</i> expression and purification.....	80
5.3.2 Formation of the Nanodiscs with fluorescent lipids .....	81
5.3.3 Formation of <i>AcrB</i> nanodisc.....	82
Chapter VI Conclusion .....	85
Appendix.....	86
A.1 Introduction .....	86

A.2 Material and methods .....	86
A.2.1 DGK expression and purification .....	86
A.2.2 DGK activity assay .....	87
A.3 Result .....	90
A.3.1 DGK expression .....	90
A.3.2 DGK activity .....	90
References .....	92
Vita .....	123

## LIST OF TABLES

Table 1.1, Sequence and ligand of protein functional tag.....	12
Table 2.1, Primer and tag sequence for GFP and PpaC.....	38
Table 4.1, Primer design for MscS mutants.....	59

## LIST OF FIGURES

Figure 1.1, Proposed secondary structure of <i>E. coli</i> tmRNA.....	14
Figure 1.2, Trans-translation model of tmRNA.....	16
Figure 1.3, ClpX hexamer structures.....	18
Figure 1.4, Substrate recognition and degradation by ClpXP.....	18
Figure 2.1, Structure of Met and AHA .....	28
Figure 2.2, Structure of tagged sfGFP and expression of sfGFP mutants.....	29
Figure 2.3, DL41(DE3) growth curve.....	30
Figure 2.4, Fluorescence spectra of sfGFP with Met or AHA.....	31
Figure 2.5, Labeling of three AHA-containing sfGFP with biotin alkyne.....	32
Figure 2.6, Fluorescence image of sfGFP-H6M4 binding on alkyne agarose resin.....	33
Figure 2.7, Structure, catalytic activity and labeling of PpaC.....	36
Figure 3.1, Catalytic activity of SA PpaC.....	44
Figure 3.2, The presence of bound Mn <sup>2+</sup> affects the thermal denaturation of PpaC.....	46
Figure 3.3, Structure and sequence of PpaC.....	48
Figure 3.4, SA PpaC mutants disrupted the dimer interface.....	50
Figure 3.5, Mn <sup>2+</sup> and Mg <sup>2+</sup> are competitive inhibitors of SA PpaC activities.....	52
Figure 4.1, Crystal structure and expression of MscS-ssrA. ....	63

Figure 4.2, Expression of MscS-ssrA.....	65
Figure 4.3, Degradation of detergent solubilized MscS-ssrA using ClpXP .....	66
Figure 4.4, Fluorescence polarization monitor the degradation of MscS-ssrA .....	68
Figure 4.5, SspB increases ClpX association with substrate. ....	69
Figure 4.6, Calibration curve to convert change of FP into percent of full length protein left .....	71
Figure 4.7, Degradation of MscS containing an inter-subunit disulfide bond.....	73
Figure 4.8, Degradation of MscS-S9C in reconstituted system.....	74
Figure 4.9, Substrate structure effect on the degradation with ClpXP.....	75
Figure 5.1, Schematic of MSP.....	77
Figure 5.2, Expression of MSP1E3.....	80
Figure 5.3, Formation of empty nanodisc with fluorescent lipid.....	82
Figure 5.4, Chromatogram of AcrB-nanodisc.....	83
Figure 5.5, Detection of AcrB-nanodisc.....	84

## LIST OF ABBREVIATIONS

AAA+ : ATPases associated with various cellular activities

ADP: Adenosine diphosphate

AHA: Azidohomoalanine

apo A1: apolipoprotein A-1

Arg: Arginine

ATP: Adenosine triphosphate

BSA: Bovine serum albumin

CBD: cellulose-binding domain

CB: Commassie blue stain

CBP: Calmodulin-binding peptide

CD: Circular Dichroism

CPB: carboxypeptidase B

CT: C-terminal domain

CuBr: Copper(I) bromide

DDM: n-dodecyl  $\beta$ -D-Maltopyranoside

DGK: Diacylglycerol kinase

DM: Decyl  $\beta$ -D-maltopyranoside



DNA: Deoxyribonucleic acid

DOG: 1,2-dioctanoyl-sn-glycerol

DTT: Dithiothreitol

EGTA: Ethylene glycol-bis( $\beta$ -aminoethyl ether)-N,N,N',N'-tetraacetic acid

EDTA: Ethylenediaminetetraacetic acid

ELISA: Enzyme-linked immunosorbent assay

FP: Fluorescence polarization

GFP: Green fluorescent protein

GST: Glutathione S-transferase

HEPES: 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

HPLC: High-performance liquid chromatography

IgG: Immunoglobulin G

IMAC: Immobilized metal-affinity chromatography

IPTG: isopropyl-D-thiogalactopyranoside

KCl: Potassium chloride

KOH: Potassium hydroxide

LB: Luria Broth

LDH: Lactate dehydrogenase

MBP: Maltose-binding protein

Met: Methionine

MnCl<sub>2</sub>: Manganese(II) chloride

MnPPi: Manganese pyrophosphate

mRNA: Messenger RNA

MSP: Membrane Scaffold Protein

MscS: Mechanosensitive Channel protein

NAD<sup>+</sup>: Nicotinamide adenine dinucleotide- oxidized form

NaCl: Sodium chloride

NADH: Nicotinamide adenine dinucleotide-reduced form

NBD-PE:1,2-dipalmitoyl-sn-glycero-3-phosphoethanolamine-N-(7-nitro-2-1,3-benzoxadiazol-4-yl)

NMR: Nuclear magnetic resonance spectroscopy

NTA: Nitrilotriacetic acid

NT: N-terminal domain

PAHs: Polycyclic aromatic hydrocarbons

PCR: Polymerase Chain Reaction

PDB: Protein Data Bank

PEG: Polyethylene glycol

PEP: Phosphoenolpyruvate

PIPES: Piperazine-N,N'-bis(2-ethanesulfonic acid)

PK: Pyruvate kinase

PMSF: Phenylmethanesulfonyl fluoride

PPase: Pyrophosphatase

POPC: 1-Palmitoyl-2-Oleoyl-sn-Glycero-3-Phosphocholine

PpaC: Pyrophosphatase

PPi: Pyrophosphate

PVDF: Polyvinylidene difluoride

RNA: Ribonucleic acid

SEC: Size exclusion chromatography

SDS-PAGE: sodium dodecyl sulfate-polyacrylamide gel electrophoresis

sfGFP: superfolder Green Fluorescent Protein

TBTA: Tris[(1-benzyl-1H-1,2,3-triazol-4-yl) methyl]amine

tmRNA: Transfer-messenger RNA

tRNA: Transfer RNA

WB: Western blot

$\Delta\text{acrB}\Delta\text{clp}$ : *acrB* and *clp* knocked out strain

## Chapter I Introduction

### 1.1 The brief history of protein engineering

Protein engineering, a process to develop useful novel proteins, was first described three decades ago.<sup>1, 2</sup> A lot of techniques have been developed since then. In order to construct proteins with function, detailed understanding about protein structure and organization is necessary. The first protein sequence was determined in 1951. The sequence of insulin was determined by analyzing hydrolyzed protein fractions using paper chromatography.<sup>3</sup> In 1960s, the crystal structure of myoglobin was determined by X-ray diffraction.<sup>4</sup> Protein Data Bank (PDB), established in 1971, provides researchers a crystallographic database for various molecules including protein, nucleic acid, and oligosaccharides.<sup>5</sup> In 1980s, nuclear magnetic resonance spectroscopy (NMR), an alternative method from X-Ray, was first used to determine protein structure.<sup>6</sup> Later on, the development in molecular biology techniques, especially the Polymerase Chain Reaction (PCR) technique, enabled the development of methods for protein engineering by changing specific residues.

Nowadays, protein engineering is used in various areas, ranging from food industry to environmental and medical applications, to bioassays.<sup>7</sup> Engineering of enzyme used in food industry was focus on improving its catalytic property, such as thermostability, specificity, and catalytic efficiency, which could be optimized using protein engineering.<sup>7</sup> For example, Altamirano et al. evolved a new phosphoribosylanthranilate isomerase that has catalytic properties similar to those of the natural enzyme but with higher specificity constant by combining the binding site of phosphoribosylanthranilate with a catalytic template required for isomerase activity. This study demonstrated a new direction for the evolution of new

biocatalysis.<sup>8</sup> The recombinant DNA technology of protein engineering has been used successfully in developing novel enzymes for food industry. Bacteria host strains have also been modified to increase enzyme yield.<sup>9</sup> Engineered transglutaminase, for example, an enzyme crosslinking protein molecules, has been widely used in food industry to improve protein properties like elasticity, viscosity, heat stability and water holding capacity.<sup>10</sup>

Protein engineering is also used in environmental applications, such as organic pollution elimination, petroleum biorefining, and microbial bioplastics. Peroxidases and laccases catalyze the oxidation of a broad range of organic substrate such as polycyclic aromatic hydrocarbons (PAHs), phenols, organophosphorus pesticides and azo dyes. However, factors including low enzyme stability in organic solvent, unfavorable substrate distribution, inhibition of enzyme-substrate complex, low reaction rate, high cost of the enzymes, limit the usage of these enzyme for commercial application. To overcome these problems, protein engineering or chemical modifications were used to obtain more robust enzymes.<sup>11</sup> Another example is the application of biotechnology in petroleum refining. For instance, desulfurization and denitrogenation of fuels are important refining objects in petroleum refining. Using genetic modification to recombine different organisms enabled multi-compounds transformation, higher activities refining, and stabilization of process conditions. Also, enhanced enzymes with higher solubility in organic media, higher activities and stability were created by protein engineering.<sup>12</sup>

Last but not least, protein engineering has found diverse applications in the production of medical therapeutics. Protein engineering for cancer treatment is a major area of interest. Modified proteins, polypeptides and other drug delivery vehicles at nano scale can be used in diagnosis and selectively targeting in cancer therapies.<sup>13</sup> Secreted protein such as

erythropoietin, insulin, interferon, plasminogen activators, growth hormone and colony-stimulating factors are new sourced that are engineered to produce therapeutics.<sup>14</sup> Meganucleases, sequence-specific endonucleases, could be used in gene therapy to correct mutated genes.<sup>15</sup> Modified antibody with altered antigen binding sites and size have been generated and evaluated as imaging probes to target tissues of interests, which was accomplished through protein engineering. It is obvious that antibody-based molecular imaging will be a convenient tool to diagnose cancer and other diseases,<sup>16</sup> and biopolymer production has great potentials in medical applications.<sup>17</sup>

## **1.2 Methods of Protein engineering**

More and more methods are becoming available for protein engineering with the development of biological science. Basically, there are two flavors of protein engineering. In “Rational design”, specific sites were modified on the protein. This method, in general, has to rely on detailed knowledge of the target protein. In “directed evolution”, mutations are randomly introduced into proteins and then hits with desired properties are selected.<sup>7</sup> The random method overcomes some problems in the rational method, which is limited by the level of understanding of the target protein. Replacing one amino acid with another may lead to structure and function change. Therefore, randomly mutation coupled with selection for the desired feature is a powerful method.<sup>18</sup> A simplest and most common technique used for random mutagenesis is called “saturation mutagenesis”, a method that alters a single amino acid to each of the remaining natural amino acids to provide every possible variant of the protein.<sup>18, 19</sup> Evolutionary methods is used for randomization of the entire gene of interest by error prone PCR to create variants with desired changes.<sup>18, 20</sup>

Another useful tool for protein engineering is “cell-free protein expression”, which solves certain limitation of recombinant protein expression in living cells, such as aggregation or degradation.<sup>21</sup> In this method, proteins are produced by combining DNA encoding the target protein with all component needed for protein translation. This method may substitute *in vivo* protein expression in the future due to its controllability and simplicity.<sup>22</sup>

A good example of protein engineering is Green fluorescent protein (GFP), which is widely used in life science research for its high stability and spectroscopic characteristics.<sup>23</sup> GFP has been used to signal gene expression,<sup>24</sup> protein localization,<sup>25</sup> biosensor,<sup>26</sup> and protein-protein interactions.<sup>27, 28</sup> Protein engineering studies have helped improve the folding of GFP and produce variants of different colors.<sup>29</sup>

### **1.3 Protein functional tag**

In the study of proteomics, protein engineering techniques provide the methods to modify proteins, so that the protein of interest is more amendable for study. Protein tags, a peptide sequences facilitating purification, labeling, detecting, or imaging of proteins, have been attached to the proteins of interests by genetically modifying the gene encoding the target protein and used broadly both in basic research and commercial applications.<sup>30</sup>

In order to efficiently purify target protein with low cost and labor consumption, affinity tags have been fused onto the protein to facilitate easy purification, including protein purification based on protein-protein interactions, enzyme-substrate interactions, protein-metal interactions etc.<sup>31</sup> Several features are shared by affinity tags, such as one-step



adsorption, minimum effect on protein structure and function, specifically removable, simple and accurate binding, and applicable to different proteins.<sup>30, 32</sup> When choosing a system for protein purification several factors need to be considered. Small tags such as flag-, poly-His-, poly-Arg-, S-, Strep- tags etc, have less influence on protein structure and activity on specific location of target protein, are often not removed after purification, while larger peptide or fusion proteins have to be removed so that can be used for different applications later on, such as crystallization or antibody production.<sup>33</sup> Several functional tags are listed in Table 1.1.

### **Polyarginine tag**

The Arg-tag is the peptide consists of five to six arginines and was first described in 1984.<sup>34</sup> The polyarginine was fused at the C-terminal of human urogastrone, and a two-step ion exchange chromatography led to >95% purity with 44% yield. The basic polyarginine tag was then removed by carboxypeptidase to avoid potential interference of protein structure and function.<sup>34</sup> Arg<sub>9</sub>-tag had been fused to bovine pancreatic ribonuclease (RNase A) at the C-terminus, and the positive charged tag strengthened the interaction of the protein with cation-exchange resins. Thus, most Arg<sub>9</sub>-tagged RNase A bound to the column, and could be eluted with concentrated NaCl solution (> 1 M). The exoprotease carboxypeptidase B (CPB), which catalyzes the hydrolysis of peptide bonds at the C-terminus of basic residues, was used to remove the Arg<sub>9</sub>-tag. Although the Arg<sub>9</sub>-tag did not affect the activity of RNase A, it destabilized the protein due to unfavorable interaction between cationic the Arg-tag and the overall cationic RNase A.<sup>35</sup> Polyarginine peptides bind to negatively charged mica surfaces, which enables oriented protein immobilization on flat surfaces.<sup>36</sup> Hexa-arginine-tagged esterase has been effectively immobilized on gold-

coated magnetic nanoparticles. The immobilization decreased the activity by 40% but can be retained after several uses by magnetic decantation.<sup>37</sup>

### **Polyhistidine tag**

Immobilized metal-affinity chromatography (IMAC), described by Porath et al in 1975, is a technique based on the interaction between transition metal ions ( $\text{Ni}^{2+}$ ,  $\text{Zn}^{2+}$ ,  $\text{Cu}^{2+}$ ,  $\text{Co}^{2+}$ ) chelated on solid support and the side chain of specific amino acid.<sup>38</sup> Histidine exhibits the strongest interaction with immobilized metal ion matrices.<sup>30</sup> Histidine tagged protein binds to the chelating matrices at neutral or slightly alkaline pH, then contaminant molecules can be washed away by buffer containing modest concentration of imidazole (~50 mM), elution of his-tagged protein can be easily achieved by either decreasing the pH (pH 4-5) or adding imidazole (> 250 mM).<sup>38</sup> In 1987, Hochuli developed a nitrilotriacetic acid (NTA) adsorbent for metal chelate chromatography, the purification method with histidine tag was further improved. The NTA ligand is quadridentate and is suitable for metal ions with coordination number of six, leaving the two remaining valencies for reversible binding of biomolecules.<sup>39</sup> A model protein mouse dihydrofolate reductase was fused with a poly-histidine peptide and expressed in *E. coli*, and the fusion protein was successfully purified through Ni-NTA.<sup>40</sup> Besides washing with low concentration of imidazole, EDTA at low concentration of about 0.5 mM was used to remove non-specifically bound proteins.<sup>41</sup> Imidazole can affect NMR experiments, competition studies, or crystallographic trials, as well as leading to protein aggregations, and thus is normally removed through dialysis. A method with two-step purification that removes hexahistidine tag with a thrombin cleavage site has been established.<sup>42</sup> Histidine tag can be placed on either C- or N-terminus based on individual protein, but it may have negative impact on protein structures and activities.

The biochemical properties of the protein might be altered due to the presence of histidine-tag<sup>43</sup>, and C-terminally tagged l-lactate dehydrogenase (LDH) displayed lower activity than wild type enzyme.<sup>44</sup> Nevertheless, purification through histidine tag with affinity column is still the most popular method.

### **Strep-tag**

Strep-tag is a nine amino acid peptide (AWRHPQFGG) which is developed for purification with streptavidin columns.<sup>45</sup> Strep-tag was originally selected from a library of peptides to specifically bind to the core part of streptavidin. Streptavidin or avidin can tightly and specifically bind to biotin, together with its high intrinsic stability and low nonspecific interaction, streptavidin or other engineered versions have been broadly used for purification as well as detection method.<sup>46</sup> Strep-tagged protein reversibly binds to the same pocket where natural ligand D-biotin is complexed on the streptavidin immobilized affinity column, and can be eluted competitively using D-biotin or a suitable derivative with less strong affinity, such as D-desthiobiotin.<sup>47</sup> Over the years, both strep-tag and its interaction partner streptavidin have been improved. The optimized strep-tag II is an octapeptide (WSHPQFEK), which has more flexible attachment site. And the streptavidin variant with higher binding capacity named Strep-Tactin has also been developed. Another advantage of Strep-tag II is that it does not affect protein folding nor interfere protein function.<sup>47</sup> The use of Strep-tag can be applied both under denaturing condition like western blot and native state as ELISA.<sup>48</sup> Recombinant proteins with strep-tag has been purified and crystalized.<sup>49</sup>

## **FLAG-tag**

Flag-tag, consist of eight amino acid DYKDDDDK, can be added to a protein for immunoaffinity chromatography. It is based on a calcium-dependent antibody binding, which recognizes the first three amino acids of this peptide segment. Bound flag-tagged protein can be eluted with EDTA or low pH ~3.0.<sup>50, 51</sup> Flag tag can be placed on either end of the protein. It is found that a shorter version of the flag tag with four peptide (DYKD) binds with the antibody M1 with the same affinity.<sup>52</sup> Another monoclonal antibody (anti-FLAG M2) was found to bind with flag peptide on either C- or N-terminal flag fused protein without calcium dependent. Therefore, target protein can only be eluted with lower pH or competition with free flag peptide.<sup>53</sup> The purity of proteins purified via flag tag about 90%.<sup>54</sup> One disadvantage of the system compared with Ni<sup>2+</sup>-NTA or Strep-Tactin is that the monoclonal antibody matrix is not as stable as other resins.<sup>30</sup> Flag tag can be used in purification and detection, crystalized flag-fused protein has similar structure as untagged protein.<sup>55</sup>

## **c-myc-tag**

The murin anti-c-myc antibody 9E10 was developed for binding with an antibody epitope of ten amino acids (EQKLISEEDL) which is called c-myc-tag.<sup>56</sup> It can be fused at N- or C-terminus of proteins. C-myc-tag is mainly used for detection such as western blot assay, immunoprecipitation, and flow cytometry,<sup>57</sup> although it can also be used for purification. A fragment of insulin-like growth factor-1 has been overexpressed in mammalian cells and purified via the c-myc-tag, then eluted at low pH ~3.0. The purified protein has been crystalized.<sup>58</sup>

## **S-tag**

When pancreatic ribonuclease A (RNase A) is digested into two fragments, the 15-amino acid S-tag and 103-amino-acid S-protein still bind together tightly.<sup>59</sup> S-tag contains KETAAAKFERGHMDS, which bind with S-protein with  $k_d$  of  $\sim 0.1 \mu\text{M}$ .<sup>60</sup> There are four cationic, three anionic, three charged polar amino acids in S-tag, and this composition makes the S-tag highly soluble. The tag can be fused to either N- or C-terminus of a protein. S-tagged protein bound to S-protein-sepharose column can be eluted by low pH  $\sim 2.0$ .<sup>61</sup> This system is useful in the development of assays with high sensitivity.<sup>62</sup>

## **Glutathione S-transferase-tag**

Glutathione S-transferase (GST) belongs to a family of enzymes that detoxicate electrophiles by glutathione conjugation.<sup>63</sup> The interaction between GST and its substrate glutathione has been applied in both detection and immunoassay.<sup>64,65</sup> Single step purification of peptides fusions with glutathione S-transferase has been described by Smith and Johnson in 1988.<sup>66</sup> The 26 kDa GST, when fused to protein of interest, has been shown to improve solubility and stability of the protein. The fusion protein binds tightly to glutathione-agarose, and can be competitively eluted with reduced glutathione. However, the use of reduced glutathione might affect target proteins contains disulfide bond.<sup>67</sup> Also, the purification need to be performed under non-denatured conditions since the binding site of GST would be destroyed by the denaturing reagents. The system has high purification efficiency but may lead to dimerization of the target protein and may interfere protein functions.<sup>68</sup> The GST domain can be cleaved from the fusion protein by engineering the protease factor X cutting site between the tag and the protein of interest.<sup>66</sup> Normally, GST tag is removed before further studies such as NMR, crystallography. Successful

studies of protein-protein interactions,<sup>69</sup> DNA-protein interactions and directional immobilization have been developed with GST assay.<sup>70</sup>

### **Cellulose-binding domain**

More than one hundred of cellulose-binding domain (CBD) sequences have been classified into more than 10 families. The CBDs are derived from cellulases, xylanases and other hydrolases, and they can be placed at N- or C- terminus, or the internal, of target proteins.<sup>71</sup> Family I CBD binds reversibly with crystalline cellulose and is used for purification in affinity chromatography. In this condition, hydrogen bond formation and van der Waals interaction are the main driving forces for binding. The fusion protein binds with cellulose column in the presence of ammonium sulfate and can be eluted with pure water.<sup>72</sup> In contrast, Family II and III CBDs can be eluted with ethylene glycol.<sup>73</sup> Some CBD fusion proteins have been irreversibly immobilized on the cellulose materials and used as immobilized enzymes.<sup>74</sup>

### **Calmodulin-binding peptide**

The purification of fusion protein containing Calmodulin-binding peptide(CBP) was first described in 1992. This peptide was derived from the C-terminus of skeletal muscle myosin light-chain kinase that binds calmodulin with nanomolar affinity.<sup>75</sup> The calmodulin binding protein can be purified by one-step affinity chromatography on calmodulin-agarose. The interaction between calmodulin and CBP is  $\text{Ca}^{2+}$  dependent, the tight binding allows more stringent washing condition to ensure high purity of the fusion protein. The fusion protein can be eluted with EDTA/EGTA.<sup>76</sup> The system can be used to purify fusion protein expressed in *E. coli* but not in eukaryotic cells, in which many endogenous proteins interact

with calmodulin in a calcium-dependent manner.<sup>77</sup> CBP can be located on N- or C-terminal. However, N-terminal tag may reduce the efficiency of translation, while tag at the C-terminus normally shows higher expression levels.<sup>78</sup>

### **Maltose-binding protein**

Maltose-binding protein (MBP) is encoded by the *malE* gene of *E. coli* K12.<sup>79</sup> The mature MBP has 370 amino acid residues and is around 40 kDa, and it specifically binds with maltose and maltodextrins with a  $K_d$  around 1  $\mu$ M.<sup>80</sup> Fusion protein with MBP tag has been purified by one-step affinity chromatography on cross-linked amylose. After binding, fusion protein can be eluted with maltose.<sup>81</sup> Buffer conditions for elution can be in the pH range from 7.0 to 8.5, and denaturing reagent cannot be used.<sup>30</sup> The MBP can be located on N- or C-terminal. It is found that MBP can function as a molecular chaperone, facilitating the solubilizing and promoting proper folding of the fusion protein.<sup>82</sup> MBP has been widely used for affinity purification. A ten asparagine residues linker between the target protein and MBP improves the binding affinity in this system. The MBP tag can be easily cleaved with site-specific protease.<sup>30</sup>

### **Other tags**

Beside all the tags described above, there are several other tags. A 13 amino acids biotag can be used, which lead to *in vivo* biotinylation in *E. coli*, for detection, immobilization and purification.<sup>83</sup> Affinity purification using the native interaction of staphylococcal protein A with immunoglobulin G (IgG) has been reported.<sup>84</sup> Intein from the *Saccharomyces cerevisiae* VMA1 gene has been used as chitin-binding domain, which

binds to chitin resin.<sup>32</sup> Avitag, PinPoint X<sub>a</sub> protein purification system used for purification<sup>85, 86</sup>, and T7 tag, HA (hemagglutinin A) tag can be used for detection.<sup>87</sup>

Table 1.1 Sequence and ligand of protein functional tag

<b>Tag</b>	<b>Sequence</b>	<b>Matrix / ligand</b>	<b>Elution condition</b>
Poly-arginine	RRRRRR	Cation exchange resin	Gradient NaCl at alkaline pH>8.0
Poly-histidine	HHHHHH	Ni-NTA	Imidazole > 250 mM or low pH
Strep-tag	WRHPQFGG	streptavidin/avidin	Biotin or desthiobiotin
Strep-tag II	WSHPQFEK	streptavidin/avidin	Biotin or desthiobiotin
Flag-tag	DYKDDDDK	monoclonalAbM1,M2	EDTA, low pH
c-myc-tag	EQKLISEEDL	monoclonal Ab 9E10	Low pH
S-tag	KETAAAKFERGH MDS	S-protein (S fragment of RNaseA)	low pH



Table 1.1 (continued)

Glutathione S-transferase-tag	26 kDa protein domain	glutathione-Sepharose	Reduced glutathione
Cellulose-binding domain	Varies	Cellulose	ethylene glycol
Calmodulin-binding peptide	KRRWKKNFIAVS AANRFKKISSSG AL	calmodulin	EGTA
Maltose-binding protein	40 kDa protein domain	cross-linked amylose	maltose

#### 1.4 Special protein tag facilitated protein degradation

In bacteria, translation involving damaged mRNA lacking the stop codon leads to stalled ribosomes, which are rescued by the transfer-messenger RNA (tmRNA) system.<sup>88-90</sup> The tmRNA, also known as 10Sa RNA or *ssrA* RNA, is a highly conserved bacteria molecule that has properties of both tRNA and mRNA.<sup>91</sup> The secondary structure of tmRNA is depicted in Figure 1.1.<sup>88</sup>

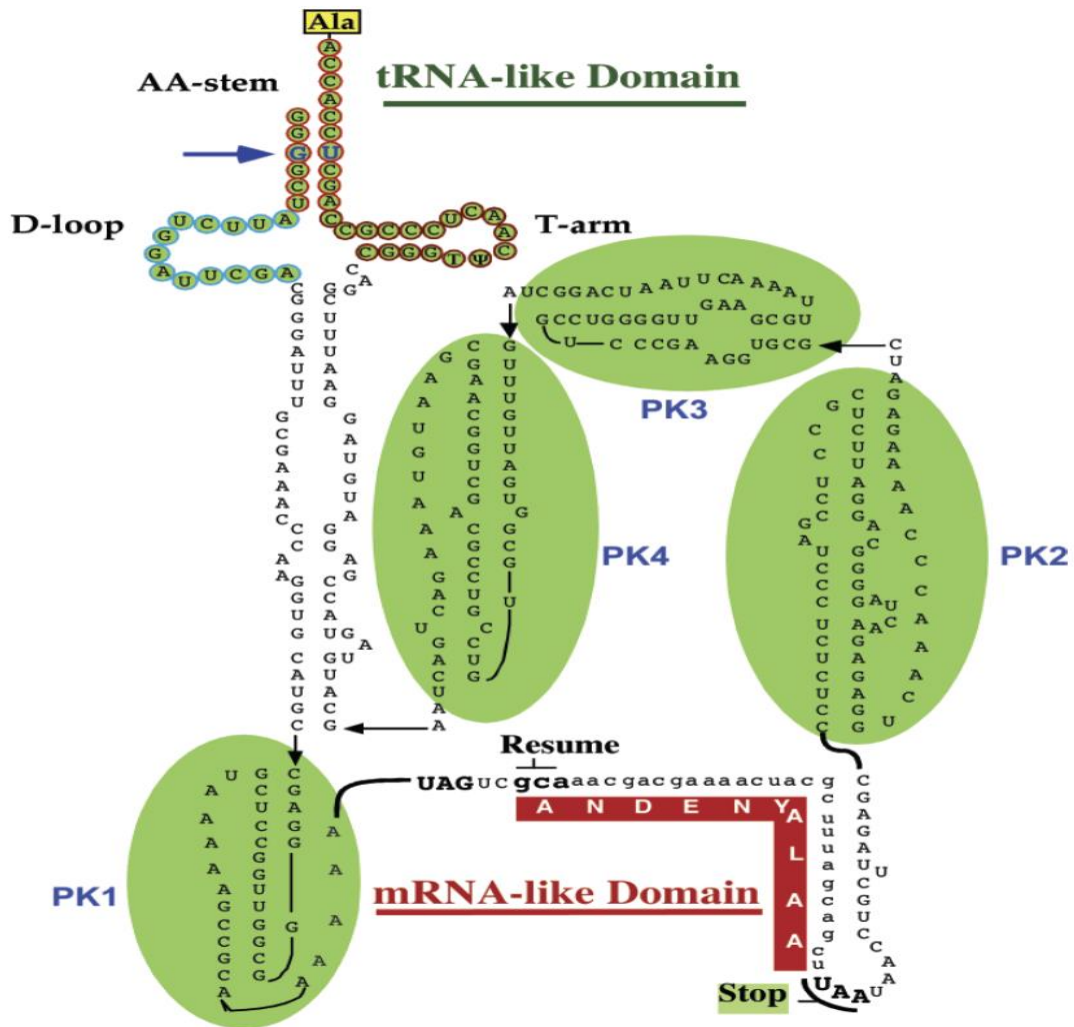


Figure 1.1 Proposed secondary structure of *E. coli* tmRNA. <sup>88</sup> Reprinted (adapted) with permission from ref <sup>88</sup>. Copyright (2007) American Chemical Society.

The 3' and 5' ends of tmRNA form a tRNA-like domain, including a D-loop, a T-arm and an amino acid acceptor stem. Unlike tRNA, instead of the anticodon stem, the tRNA-like domain is followed by a long stem and then connected with a mRNA-like domain. After that, three pseudoknots (PK2-4) link mRNA-like domain to the 3' end.<sup>92,93</sup> The tRNA-like

domain is recognized by alanyl-tRNA synthetase so as to transfer alanine to the nascent protein, and thus the stalled mRNA is released.<sup>91, 94</sup> Alter the tRNA<sup>ala</sup>-like domain to tRNA<sup>his</sup>-like domain would lead to the incorporate of histidine at the first amino acid of the 11-amino acid peptide coded by tmRNA.<sup>95</sup> The mRNA-like domain was found by Tu et al. when their effort of expressing interleukin-6 (IL-6) in *E. coli* resulted in a collection of IL-6 fragments.<sup>96</sup> All truncated fragment of IL-6 were found to be tagged with an 11-amino acid peptide, AANDENYALAA. The last 10 amino acids were encoded by the mRNA-like domain. The C-terminal tagged proteins become substrates for cellular proteases. The process of tmRNA mediated protein degradation and ribosome rescue is depicted in Figure 1.2. In *E. coli*, at least four proteases, Tsp, FtsH, ClpXP, and ClpAP, degrade the C-terminal tagged polypeptide.<sup>89, 97-99</sup>

Tsp, also called Prc, is a tail-specific periplasmic protease that recognize protein with non-polar C-terminal residues. It has been shown to degrade a variant of N-terminal domain of  $\lambda$  repressor with the non-polar C-terminal tag of WVAAA, but not polar C-terminal sequence RSEYE.<sup>100</sup> Also, Tsp is found to be more effective to degrade substrates that are not stable, and a peptide substrate is preferred than a protein substrate for cleavage.<sup>101</sup> The C-terminal residues of *ssrA* (YALAA) is similar to that of Tsp recognizing C-terminal tag (WVAAA). A periplasmic protein (cytochrome *b*<sub>562</sub>) was constructed with *ssrA* tag, and subsequently degraded by Tsp *in vivo*. This supports the hypothesis that Tsp degrade the *ssrA* tagged periplasmic protein.<sup>97</sup> Tsp degraded  $\lambda$  repressor (a cytoplasmic protein) with nonpolar C-terminal tag *in vitro*, but not *in vivo*, which indicates that at least one additional protease also works for C-terminal specific proteolysis.<sup>102</sup>

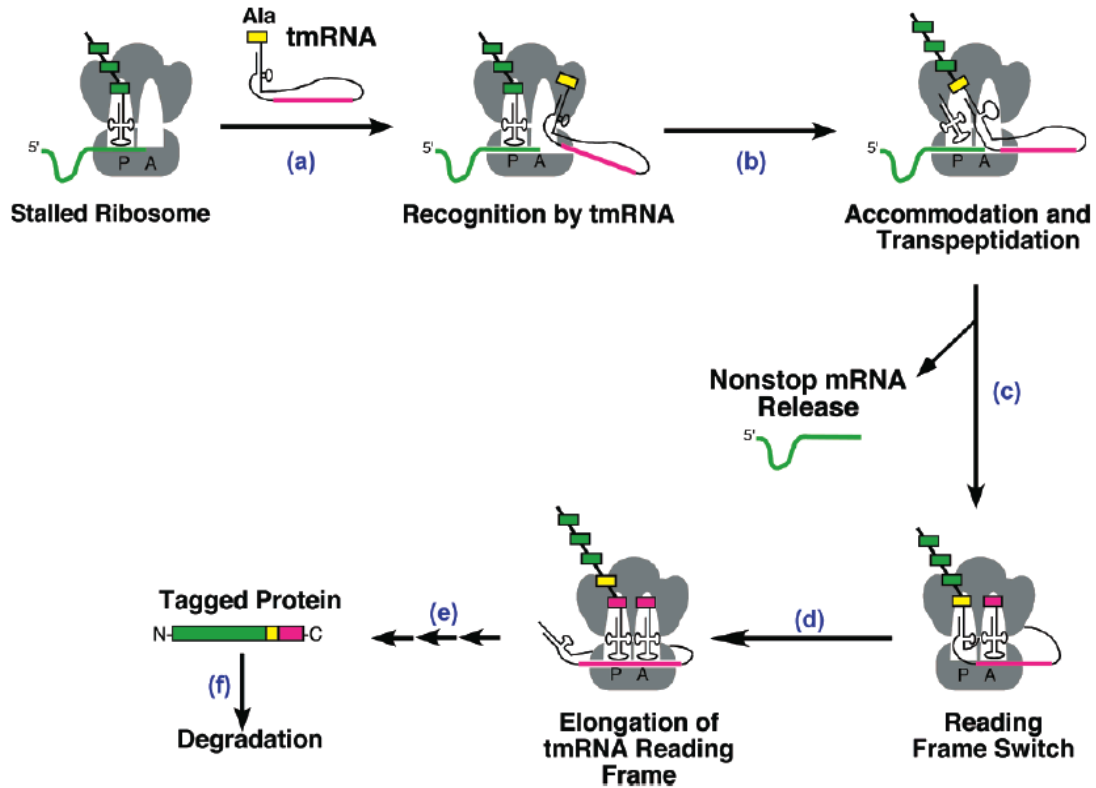


Figure 1.2 Trans-translation model of tmRNA.<sup>88</sup> (a) tmRNA charged with alanine enter the A site of stalled ribosome, (b) transfer tmRNA to the nascent chain (c) release of the faulty mRNA and replace with open reading frame on the tmRNA (d) translate the tmRNA until the stop codon is reached (e) the tagged protein is released and degraded by specific proteases. Reprinted (adapted) with permission from ref<sup>88</sup>. Copyright (2007) American Chemical Society.

FtsH, an essential protease, is an integral membrane protein with a ATP-binding domain at the cytoplasmic C-terminus.<sup>103, 104</sup> It degrades the heat-shock transcription factor  $\sigma^{32}$ , which is involved in the gene regulation in *E.coli*.<sup>105</sup> Moreover, FtsH degrades a set of meta-stable proteins, also some misassembled membrane proteins, as a means of the cellular protein quality control.<sup>106-109</sup> SsrA tagged phage  $\lambda$  cI repressor, a cytoplasmic

protein, is rapidly degraded with FtsH *in vitro*.<sup>99</sup> Later, an inner-membrane protein, ProW1–182, tagged with *ssrA* tag has been shown to be degraded by FtsH *in vivo*, suggest that FtsH plays a role in maintaining protein quality in the inner membrane.<sup>110</sup> ClpXP and ClpAP are two well studied protease complexes, and both using ClpP as the protease.<sup>111</sup>,<sup>112</sup> *E.coli* ClpP is expressed as a proenzyme,<sup>113</sup> and it alone cannot function efficiently as a protease. ClpP needs to associate with its partner ATP-driven unfoldase to cleave polypeptides.<sup>114</sup> ClpX and ClpA are AAA+ (ATPases associated with various cellular activities) family unfoldase.<sup>111</sup> ClpX and ClpA use the energy of ATP hydrolysis to help unfold the substrates and feed them into the active chamber of ClpP.<sup>115, 116</sup> Several short sequence motifs are recognized by both ClpX and ClpA, and the 11-amino acid (*ssrA*) tag is the best characterized recognition motif.<sup>117</sup> The hexamer structure of ClpX is depicted in Figure 1.3. The N domain of each subunit is coordinated with a zinc atom,<sup>118</sup> while the rest forms the AAA+ module. It is found that ClpX and ClpX<sup>ΔN</sup>, which lacks the N domain, both form hexamer rings, and ClpX<sup>ΔN</sup> binds with ClpP and supports the degradation of some protein substrate at a similar rate to ClpX, indicates that AAA+ domain plays the major function for ClpX.<sup>119,120</sup> The N domain contribute to hexamer stability and recognition of adaptors, such as SspB.<sup>121</sup> SspB improves the ClpXP degradation of *ssrA* tagged proteins by lowering  $K_M$ .<sup>122</sup> Two features of SspB help tether substrates to ClpX and enhance degradation. First, SspB dimer binds to the N-terminal of *ssrA* tag (AANDENY), leaving the last three amino acid (LAA) for ClpX binding.<sup>123</sup> Second, the C-terminal of each SspB unit binds with ClpX and stimulate ClpXP degradation.<sup>118, 124, 125</sup> In the presence of SspB adaptor, degradation rate increases dramatically. For example, mutated C-terminal *ssrA* tagged substrates for weak ClpX binding were degraded around

100-fold faster in the presence of SspB than in the absence.<sup>126</sup> The ClpXP degradation process is depicted in Figure 1.4.<sup>121</sup>

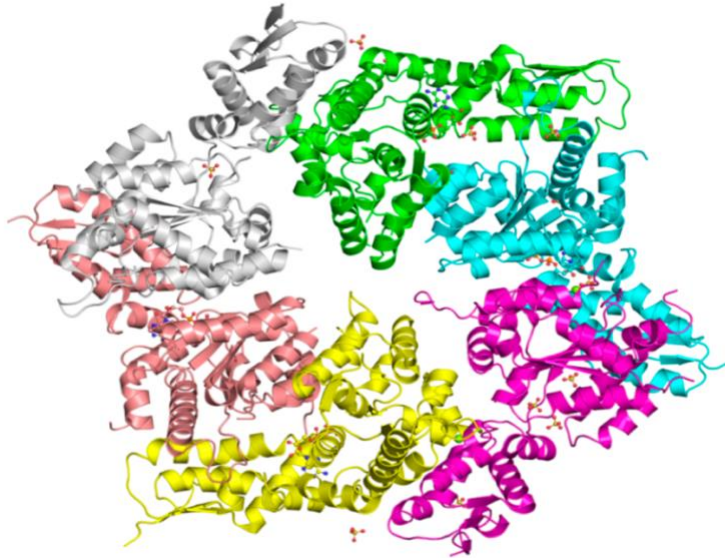


Figure 1.3 ClpX hexamer structures. Face view colored by subunit. PDB ID: 3HWS.

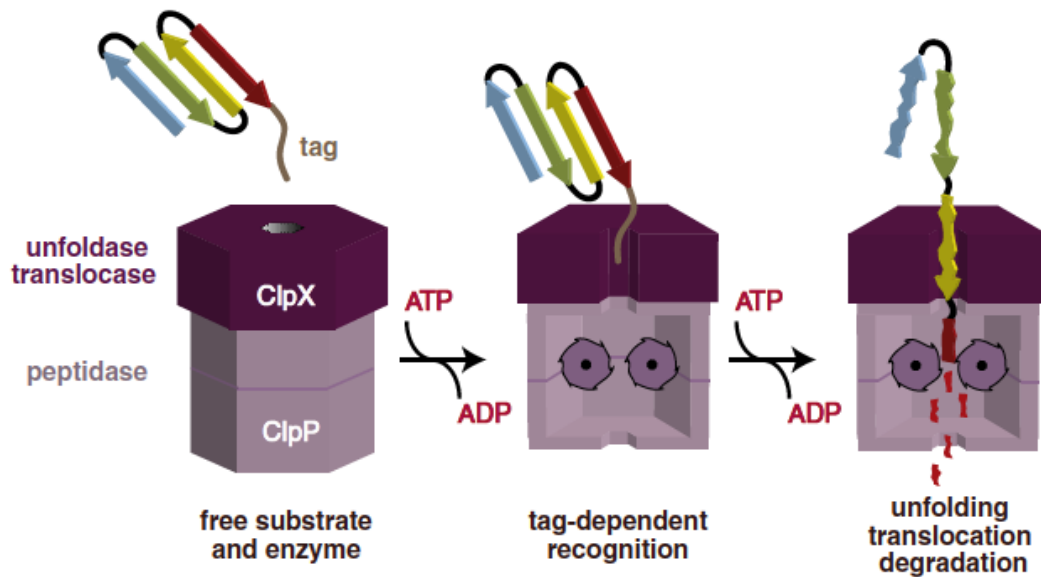


Figure 1.4 Substrate recognition and degradation by ClpXP. First, ClpX recognizes the tagged protein and binds the substrate to the axial pore of ClpX. Second, ClpX unfolds the

substrate with the energy from ATP hydrolysis and translocates the unfolded substrate into ClpP chamber for proteolysis. Third, ClpP cleave the substrate into small peptide fragments.<sup>121</sup> Adapted with permission from ref.<sup>127</sup>

Accumulation of stalled ribosome leads to many undesired consequences. There would be a significant decrease of translational efficiency due to the lack of ribosome for new mRNAs. Also, the protein products translated from defective mRNAs might be harmful to the cell.<sup>88</sup> Having multiple degradation systems dealing with ssrA-tagged proteins will guarantee normal translation, and thus, maintain the well-being of cells. SsrA tag serves as a proteolysis signal, recognized by different proteases, and enable protein quality-control in the cell and regulate the cell growth.<sup>91</sup>

## Chapter II A dual-functional tag facilitated protein labeling and immobilization

### 2.1 Introduction

Proteins are functional materials created by nature to execute diverse activities in living organisms. The superb selectivity, specificity, and performance of proteins have made them highly desirable components in the creation of bio-mimetic materials. A major obstacle in the utilization of proteins in biotechnology is the difficulty with site-specific immobilization of proteins without compromising their functions. The introduction of a peptide tag has been a popular method to facilitate protein immobilization or detection. Such peptide tags include the GST tag<sup>66, 128</sup>, strep tag<sup>47, 129, 130</sup>, HA tag<sup>131, 132</sup>, flag tag<sup>50, 52, 133</sup>, arg-tag<sup>34, 36</sup>, c-myc-tag<sup>56, 134</sup>, and histag<sup>39, 135</sup>. Proteins bearing these tags can be captured or detected via interactions with their corresponding binding partner modules or antibodies. Peptide tags are usually directly encoded into the gene of the target protein, and thus have the advantage of being convenient and highly specific. While they are very useful in the purification and detection of the tagged proteins, the application of most peptide tags in protein immobilization and modification is limited by the noncovalent nature of the interaction, which suffers from drawbacks including a high off-rate and low mechanical resilience. Several systems have been developed to promote the formation of covalent bond between a residue in the peptide tag and its catcher module, including split inteins<sup>136</sup>, the SpyTag/SpyCatcher<sup>137-139</sup>, cysteine and  $\alpha$ -chloroacetyl interaction facilitated by the coil-coil interaction<sup>140</sup>, the Histag and nitrilotriacetate-based arylazide photoreactive label via click chemistry<sup>141, 142</sup>, and the employment of the substrate peptide of a ligase or transferase, such as the LAP tag<sup>143</sup>, the Q-tag<sup>144</sup>, the sortagging motif<sup>145</sup>, the formyl glycine tag<sup>146</sup>, and the peptides A1 and S6<sup>147</sup>.



The incorporation of non-natural amino acid is another useful technique used in the site-specific modification of proteins. Examples include the azido- or alkyne- containing residues via click chemistry <sup>148</sup>, norbornene-containing residues that react with tetrazine-based probes <sup>149</sup>, and p-azido-L-phenylalanine via photocatalytic reaction <sup>150</sup>. Non-natural amino acids are usually incorporated into the protein of interest through the introduction of a dedicated orthogonal pair of tRNA and tRNA synthetase that translate a stop codon (typically TAG) into the specific non-natural amino acid. Recently, this method has been coupled with in situ biosynthesis to incorporate a sulfur-containing noncanonical amino acid S-allyl-L-cysteine (Sac) into *E. coli* proteins<sup>151</sup>. Subsequently, the non-natural amino acid can introduce unique chemistry to facilitate site-specific modification. Alternatively, if the structure of a non-natural amino acid is very similar to a natural amino acid, it could be incorporated metabolically by using the corresponding auxotrophic strain. In this case, no additional cellular machinery need to be introduced. For example, azidohomoalanine (AHA) can be incorporated into the sequence of proteins by Met tRNA and tRNA synthetase (Figure 2.1). This method has been used by several groups to label and modify proteins <sup>152-162</sup>. However, labeling of the incorporated AHAs often suffers from poor efficiency, since Met residues usually form part of the hydrophobic core of a protein and have limited accessibility to reaction. Although all proteins contain Met in their amino acid sequence, only a small percentage of these intrinsic Met residues seemed to be available for labeling. One potential solution is to denature and unfold the target protein before labeling to improve the accessibility of the AHA residues<sup>148, 163</sup>. This approach is very useful in the detection of a target protein, but not suitable for applications that demand

active protein. Otherwise Met can be introduced at the surface of the protein via site-directed mutagenesis<sup>157</sup>. Here we take advantage of the convenience of metabolic incorporation and address the issue of limited accessibility through the genetic introduction of a methionine-containing tag. To avoid the formation of a hydrophobic patch which may potentially affect protein folding, we inserted polar residues (histidine or serine) between neighboring methionine residues. The tag has two functions: histidine in the tag can facilitate protein purification via the conventional metal affinity chromatography, while methionine (or rather, their replacement by AHA) can be used in covalent labeling. The performance of the tags was evaluated using two model proteins, superfolder Green Fluorescent Protein (sfGFP) and an inorganic pyrophosphatase from *Staphylococcus aureus*, PpaC<sup>164</sup>. sfGFP was chosen due to its intrinsic fluorescence, while PpaC was chosen since its enzymatic activity can be conveniently measured using a colorimetric assay. Using these two model proteins, we demonstrated that the incorporation of the tag did not compromise protein function, and the tag could facilitate both protein purification and modification. We expect the dual-functional His-AHA tag to be useful in general for various biotechnological applications.

## **2.2 Materials and methods**

### **2.2.1 Plasmid construction**

Plasmids pET22-sfGFP and pET22-PpaC were created in earlier studies<sup>164, 165</sup>. C-terminal tags were introduced via the fast cloning method using pET22-sfGFP or pET22-PpaC as the template and primers as listed in Table 2.1<sup>166</sup>. All coding sequences were confirmed through DNA sequencing.

### **2.2.2 AHA incorporated into target protein**

For protein expression, the corresponding plasmid was transformed into *E. coli* strain *DL41(DE3)*. Single colonies were first cultured at 37 °C in M9 medium containing ampicillin (100 mg/L) overnight, and then the Met-starved overnight culture was used to inoculate 30 mL of fresh M9 medium containing 19 essential amino acids (each at 40 µg/mL without Met) and 50 ug/mL AHA with 10-fold dilution. The cells were grown at 37°C with shaking at 250 rpm until its absorbance at 600 nm (OD 600 nm) reached 0.8, and then was induced with 1 mM isopropyl-D-thiogalactopyranoside (IPTG). After overnight expression, the cells were harvested by centrifugation at 9,800 ×g for 10 min. Cell pellet was stored at -80°C. Control proteins with normal Met in the tag were expressed the same as described except replacing 50 ug/ml AHA with 40 ug/ml methionine in the culture medium.

### **2.2.3 Protein purification**

For purification, the cell pellet was resuspended in 30 mL lysis buffer (50 mM HEPES, 200 mM NaCl, pH 7.5 for PpaC, or 20 mM phosphate, 200 mM NaCl, pH 7.5 for sfGFP). The protease inhibitor PMSF stocked in 95% ethanol was added freshly to a final concentration of 0.5 mM. Cells were lysed through sonication and then centrifuged at 15,300 ×g, 4 °C for 20 minutes. The supernatant was collected and incubated with Ni-NTA agarose beads (Qiagen) for 40 min at 4 °C with shaking. The resins were then loaded into an empty column, drained, and washed with the corresponding lysis buffer supplemented with 40 mM imidazole. Finally, proteins were eluted using the corresponding lysis buffer

supplemented with 500 mM imidazole. After purification, imidazole in the samples was removed by dialysis against the lysis buffer.

#### **2.2.4 Bacteria growth curve**

*DL41(DE3)* strain was cultured in M9 medium supplemented with 20 amino acids overnight. The next morning, the overnight culture was used to inoculate three cultures of M9 medium supplemented either with 19 amino acids (each at 40 µg/mL, no Met) plus AHA (50 µg/mL), with 19 amino acids (each at 40 µg/mL, no Met), or with all 20 amino acids (each at 40 µg/mL). Cell growth was measured by monitoring the absorbance of the cell cultures at 600 nm (OD 600 nm) at the indicated time.

#### **2.2.5 Protein biotinylation via click chemistry**

The reactivity of AHA residues in the structure of sfGFP and PpaC was examined through their reaction with PEG4 carboxamide-propargyl biotin (biotin alkyne). To initiate the reaction, biotin-alkyne (50 µM), Tris[(1-benzyl-1H-1,2,3-triazol-4-yl) methyl] amine (TBTA, 600 µM), and CuBr (600 µM) were added into purified protein sample in HEPES buffer. The reaction mixture was incubated at room temperature with shaking for 5 min for sfGFP constructs and 1 hour for PpaC constructs, and then analyzed using anti-biotin Western blot.

### **2.2.6 Protein immobilization to alkyne agarose resin**

Alkyne agarose resins were purchased from Jena Bioscience. The alkyne agarose beads were first washed with 10 bed volumes of HEPES buffer (20 mM HEPES, 200 mM NaCl, pH 7.5) three times and then resuspended in two bed volume of HEPES buffer containing the indicated protein. TBTA was added to a final concentration of 600  $\mu$ M followed by mixing using pipette tip. CuBr was then added to a final concentration of 200  $\mu$ M. The reaction mixture was mixed thoroughly using pipette tip and incubated at room temperature for 2 hours. Finally, the beads were washed using 10 bed volume of HEPES buffer via centrifugation for three times.

### **2.2.7 Gel electrophoresis and Western Blot**

SDS-PAGE was conducted using 20% Tris-glycine gel. For Western blot, proteins were transferred to PVDF membrane after SDS-PAGE using semi-dry blotter (Denville, Holliston, MA) and detected using a monoclonal anti-biotin antibody-alkaline phosphatase conjugate.

### **2.2.8 Fluorescence spectroscopy**

Fluorescence emission spectra of fluorescein were collected using a Perkin-Elmer LS-55 fluorescence spectrometer (PerkinElmer, Waltham, MA) at 20 °C and excitation wavelength at 485 nm.

Fluorescent image of the sfGFP immobilized alkyne beads was taken using Eclipse Ti-S fluorescence microscope (Nikon, Melville, NY).

### **2.2.9 Pyrophosphatase activity assay**

The PPase activity was measured in Tris buffer using Mn pyrophosphate as the substrate as described <sup>167</sup>. Stock solutions of MnCl<sub>2</sub> and sodium pyrophosphate were mixed at a 1:1 molar ratio at 0.5 mM right before the analysis (mixing equal volume of 1.0 mM MnCl<sub>2</sub> and 1.0 mM sodium pyrophosphate) and diluted to the indicated substrate concentration for the activity measurement. At neutral pH, pyrophosphate is not fully deprotonated. The major species is MnH<sub>2</sub>PPi (MnPPi), which was considered the substrate for pyrophosphatase <sup>168</sup>. A stock solution of malachite green (0.12%, w/v) was made by dissolving the dye in 3 M sulfuric acid. A working solution was always made fresh by adding one volume of 7.5% (w/v) sodium molybdate into four volumes of the malachite green stock solution followed by the addition of Tween 20 to a final concentration of 0.2% (v/v). This solution is used to both terminate the enzyme reaction and initiate the colorimetric reaction to determine the concentration of phosphate. For activity measurement, PpaC was added into a freshly prepared reaction mixture containing the indicated concentration of MnPPi, in a reaction buffer (25 mM Tris-Cl, 50 mM NaCl, pH 7.0) at room temperature for 5 min. To stop the enzymatic reaction and determine the phosphate concentration of a sample, one volume of the working solution was mixed with four volumes of the enzymatic reaction mixture to be analyzed. For immobilized PpaC, the reaction mixture was subjected to a quick centrifugation and the supernatant was collected for analysis. The mixture was incubated for 5 min for the color to develop, and the absorbance at 630 nm was measured. We have experimented with the conditions and confirmed that under these conditions, the product phosphate accumulation over time was

linear. Therefore, the absorbance at 630 nm, after correction for background, directly correlates with the rate of hydrolysis. The Michaelis-Menten constant  $K_M$  and specific enzymatic activity were determined through fitting the measured values with SigmaPlot.

## **2.3 Results and discussion**

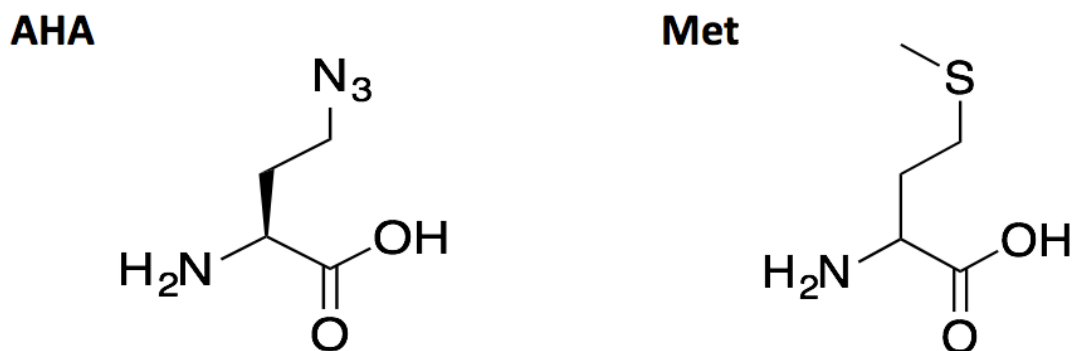
### **2.3.1 Incorporation of AHA into the target proteins**

Expression of AHA-containing proteins was performed using a Met auxotrophic strain, *DL41(DE3)*. *DL41(DE3)* could not grow in the absence of Met (Figure 2.3). The replacement of Met by AHA slowed down the growth of the strain and increased the doubling time at the exponential growth phase from 1 hour to 2 hours, but the two cultures grew to similar densities at saturation. This result indicates that AHA can be used effectively as a replacement of Met in the synthesis of proteins and support bacteria cell growth. To further confirm that AHA was incorporated to replace Met, we submitted AHA-containing PpaC-H6G3M4 for mass spectroscopy peptide fingerprinting analysis. There are 13 Met in the protein, including four in the tag. Among them peptides containing AHA-replaced Met1, Met125, or M142 were not detected. Peptides containing AHA replacement for all other 10 Met residues were identified, indicating that these residues were at least partially replaced by AHA.

### **2.3.2 Incorporation of a dual functional tag at the C-terminus of sfGFP**

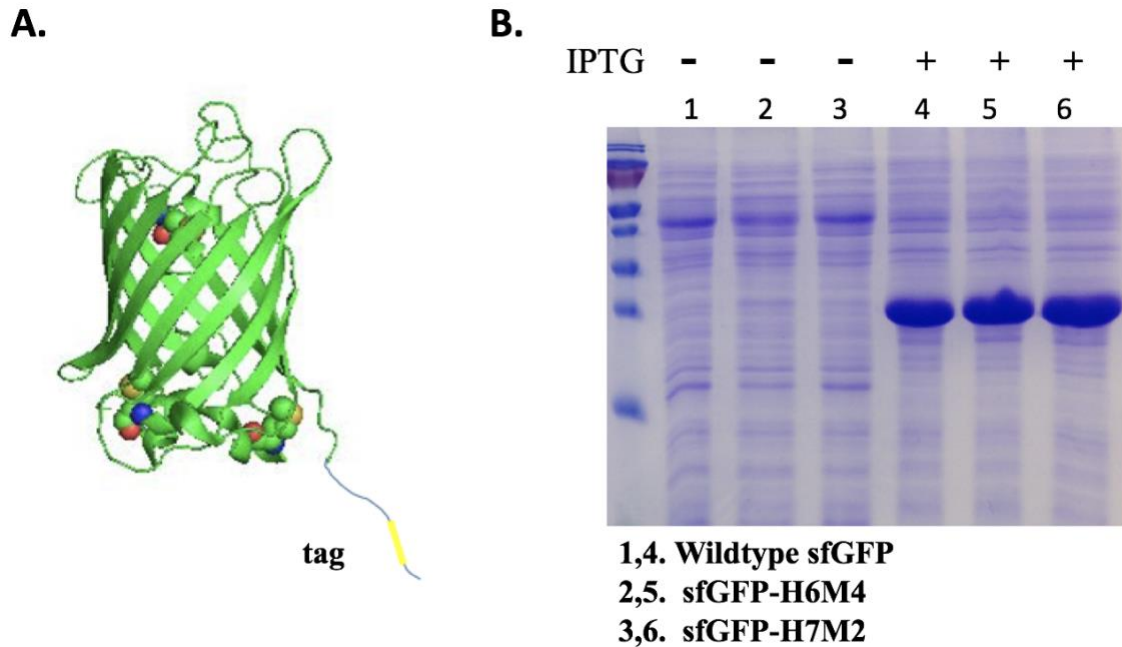
The intrinsic fluorescence of sfGFP makes it a popular model protein in studies involving protein modification, including ones using the azido-alkyne based click chemistry<sup>169</sup>. The

polyhistidine tag has been used extensively in the purification of sfGFP<sup>170</sup>. Here we have tested two different tag designs, with the addition of two or four Met residues (Figure 2.2, Table 2.1). To avoid the potential creation of a hydrophobic patch, we juxtaposed Met with His. For protein expression, *E. coli* strain *DL41(DE3)* containing plasmid sfGFP-H6, sfGFP-H7M2, or sfGFP-H6M4 were grown in the M9 medium supplemented with 20 essential amino acid as described in the Materials and Methods. As shown in Figure 2.2 B, the different tags did not affect the expression level of the protein. The two proteins containing Met residues in the tag can be purified similarly as that of the histagged sfGFP, with similar yield and purity.



**Figure 2.1** Structure of Met and AHA.

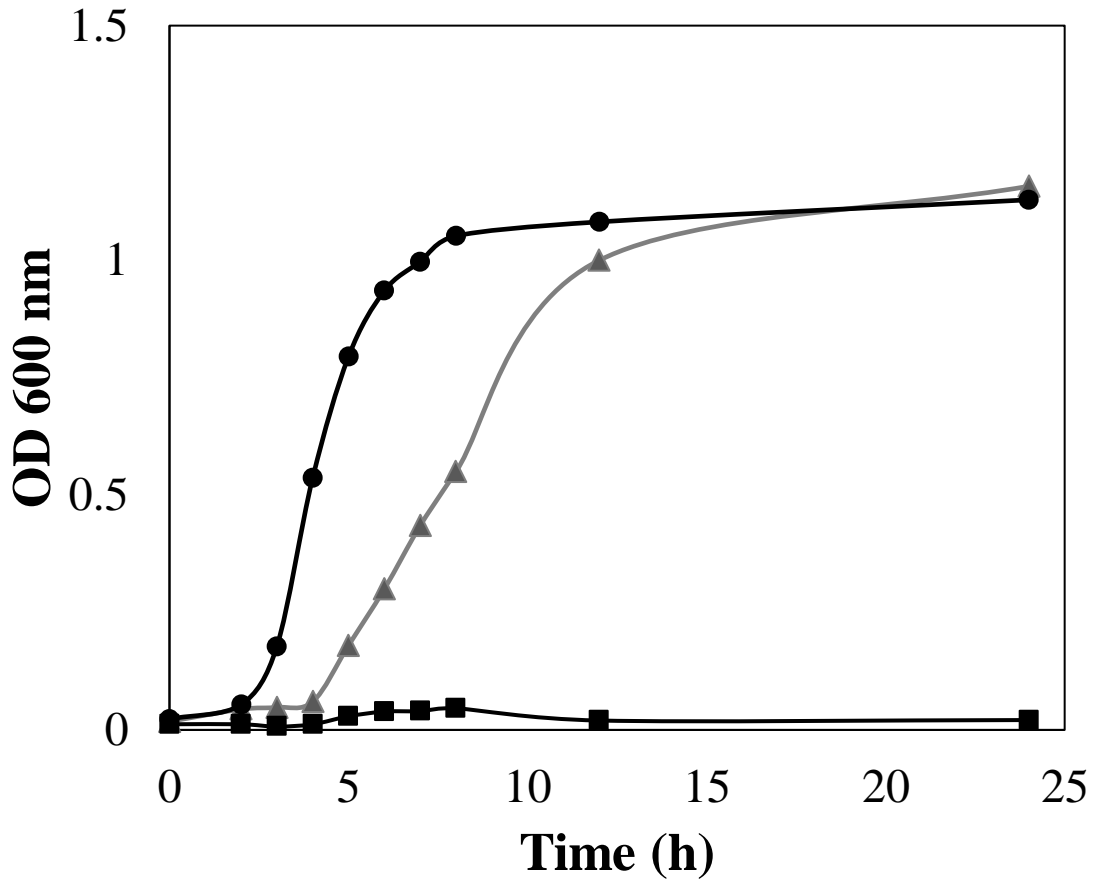




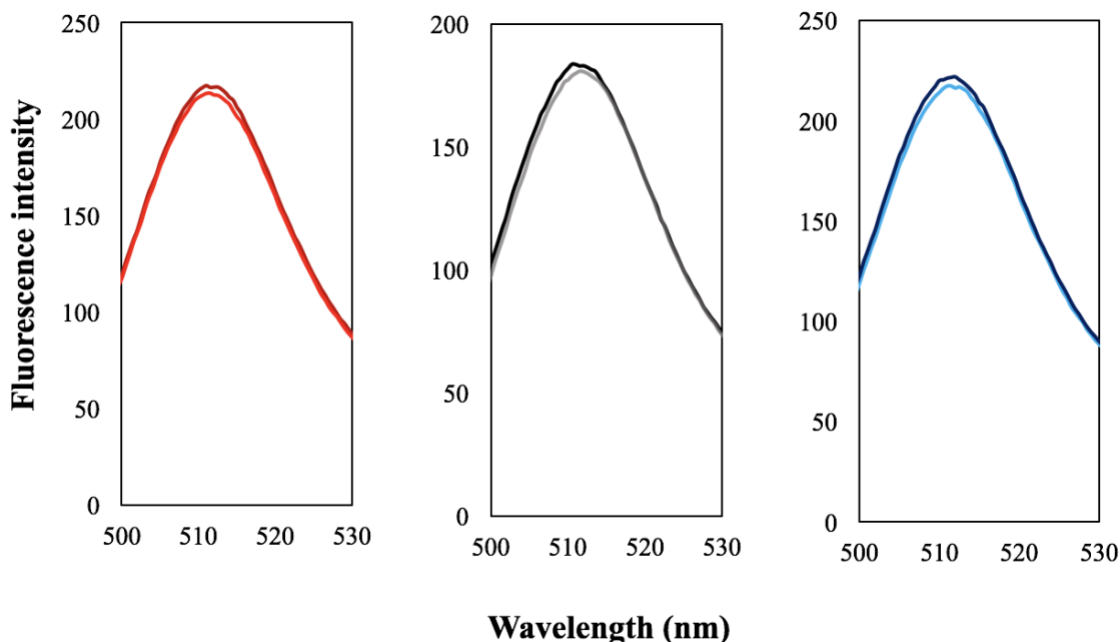
**Figure 2.2** Structure of tagged sfGFP and expression of sfGFP mutants. **A.** Structure of sfGFP with the sidechain of the three intrinsic Met residues highlighted with ball-and-stick model. The location of the C-terminal tag is shown. **B.** Cell lysate of *DL41(DE3)* transformed with the corresponding plasmids (for sfGFP) before or after IPTG induction. Molecular weight of bands in the marker, from top to bottom, are 130, 100, 70, 55, 40, 35, 25, and 15kD, respectively.

To examine the potential effect of the tag and the replacement of Met by AHA on protein structure, we measured the fluorescence spectra of the three proteins, sfGFP-H6, sfGFP-H7M2, and sfGFP-H6M4, containing Met or AHA. After expression and purification, the fluorescence spectra of the protein samples were collected (Figure 2.4). Six protein samples were examined, including the three different constructs of sfGFP expressed in the presence of either Met or AHA. The wavelength of the emission peak was not affected by the

addition of the tags or the incorporation of AHA. For each construct, the replacement of Met by AHA did not affect the fluorescence intensity. The fluorescence intensity of sfGFP-H6M4 was approximately 10% lower than that of sfGFP-H6 and sfGFP-H7M2 (data collected from 10 repeats).



**Figure 2.3** DL41(DE3) growth curve. DL41(DE3) cultured in the M9 medium supplemented with different amino acids: 19 amino acids (no Met) (squares), all 20 amino acids (circles), and 19 amino acids (no Met) plus AHA (triangles).

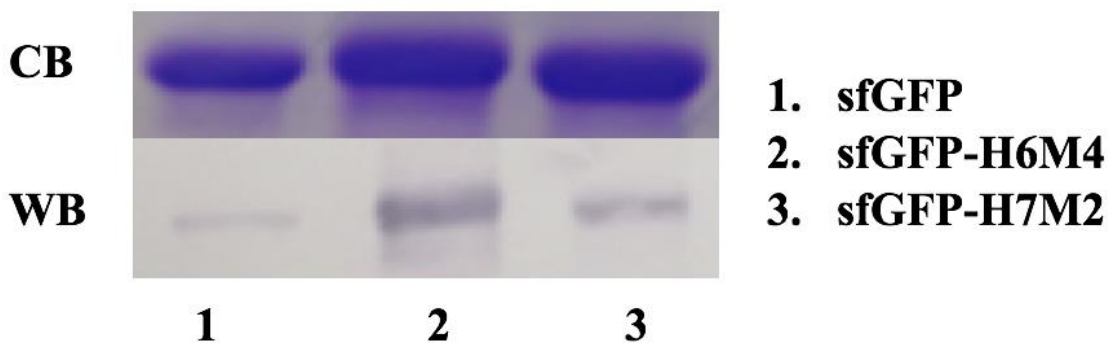


**Figure 2.4** Fluorescence spectra of sfGFP-H6 containing Met (red) or AHA (magenta); sfGFP-H7M2 containing Met (grey) or AHA (black); sfGFP-H6M4 containing Met (blue) or AHA (dark blue).

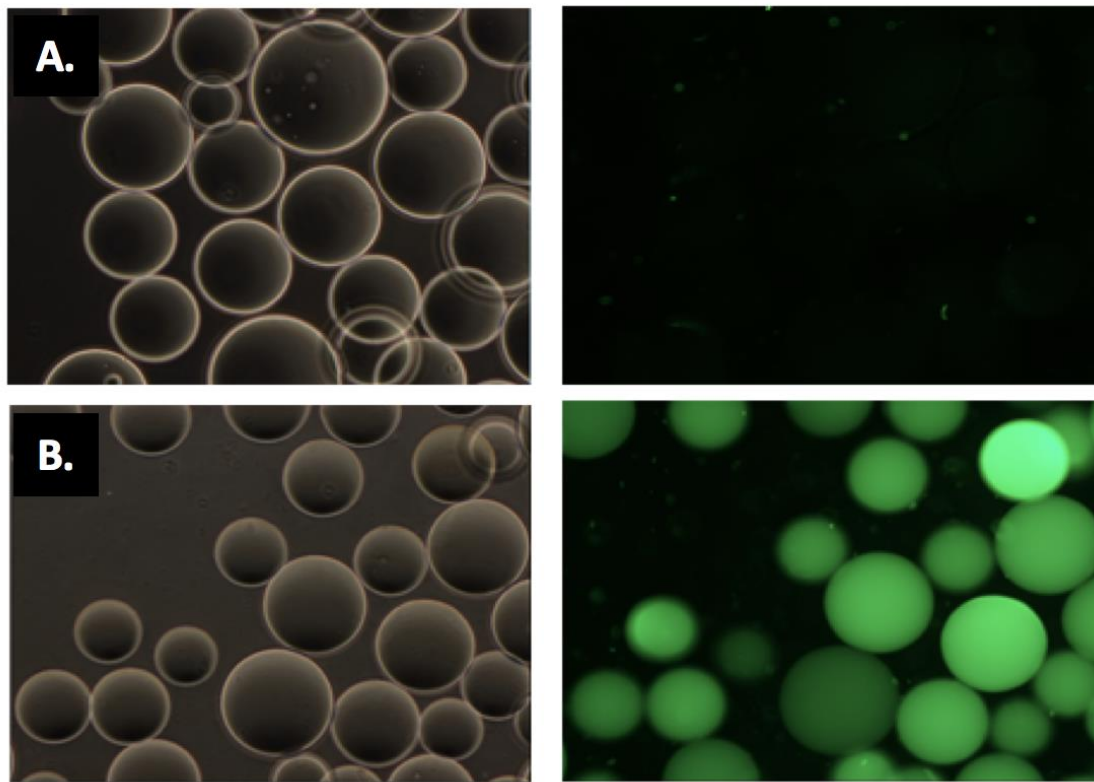
Next, we examined the effect of the Met-containing tag in promoting the efficiency of click chemistry reaction. Purified sfGFP bearing H6, H7M2 or H6M4 tag was subjected to labeling using biotin-alkyne followed by detection using anti-biotin Western blot. As shown in Figure 2.5, the level of labeling of sfGFP-H6M4 was significantly better than that of the sfGFP- H7M2, which was also better than that of the sfGFP-H6. The relative levels of labeling of H6M4 and H7M2 were approximately ten- and three-fold, respectively, to the level of sfGFP-H6. There are three intrinsic Met in sfGFP, as highlighted in Figure 2.2. The side chain of these Met residues are likely involved in hydrophobic interactions in the

native structure of sfGFP. Therefore, when AHAs replace these Mets, their side chains are likely having limited accessibility for reaction.

To demonstrate the usefulness of the tag in facilitating immobilization, we expressed sfGFP-H6M4 in the presence of AHA or Met, and incubated alkyne agarose resin with purified proteins for immobilization using click chemistry. After immobilization and washing, images were taken of the modified resin under normal white light or blue light (Figure 2.6 A & 2.6 B). It is clear that the reaction with the alkyne agarose resin depends on the presence of AHA in the protein. The immobilization of AHA-containing sfGFP-H6M4 is highly specific. AHA-containing sfGFP-H6M4 readily attached to the alkyne beads, while no binding could be detected for sfGFP-H6M4 expressed in the absence of AHA.



**Figure 2.5** Labeling of three AHA-containing sfGFP with biotin alkyne. Commassie blue stain (CB) and anti-biotin Western blot (WB) analysis of three AHA-containing sfGFP constructs reacted with biotin alkyne.



**Figure 2.6** Fluorescence image of sfGFP-H6M4 binding on alkyne agarose resin. Agarose resin were monitored under white light, while bounded sfGFP were identified under blue light. **A.** Alkyne agarose resin incubated with sfGFP-H6M4 expressed with Met, and then imaged using normal white light (left) or blue light (right). **B.** Alkyne agarose resin clicked with sfGFP-H6M4 expressed with AHA, and then imaged using normal white light (left) or blue light (right).

### 2.3.3 Addition of the GS3M4 tag to a pyrophosphatase PpaC

We used a pyrophosphatase PpaC as a model enzyme to further examine the performance of the methionine-containing tag. Here we added an octapeptide ‘GMSMSMSM’ after the histag at the C-terminal of PpaC (Table 2.1). PpaC is a good model enzyme because the

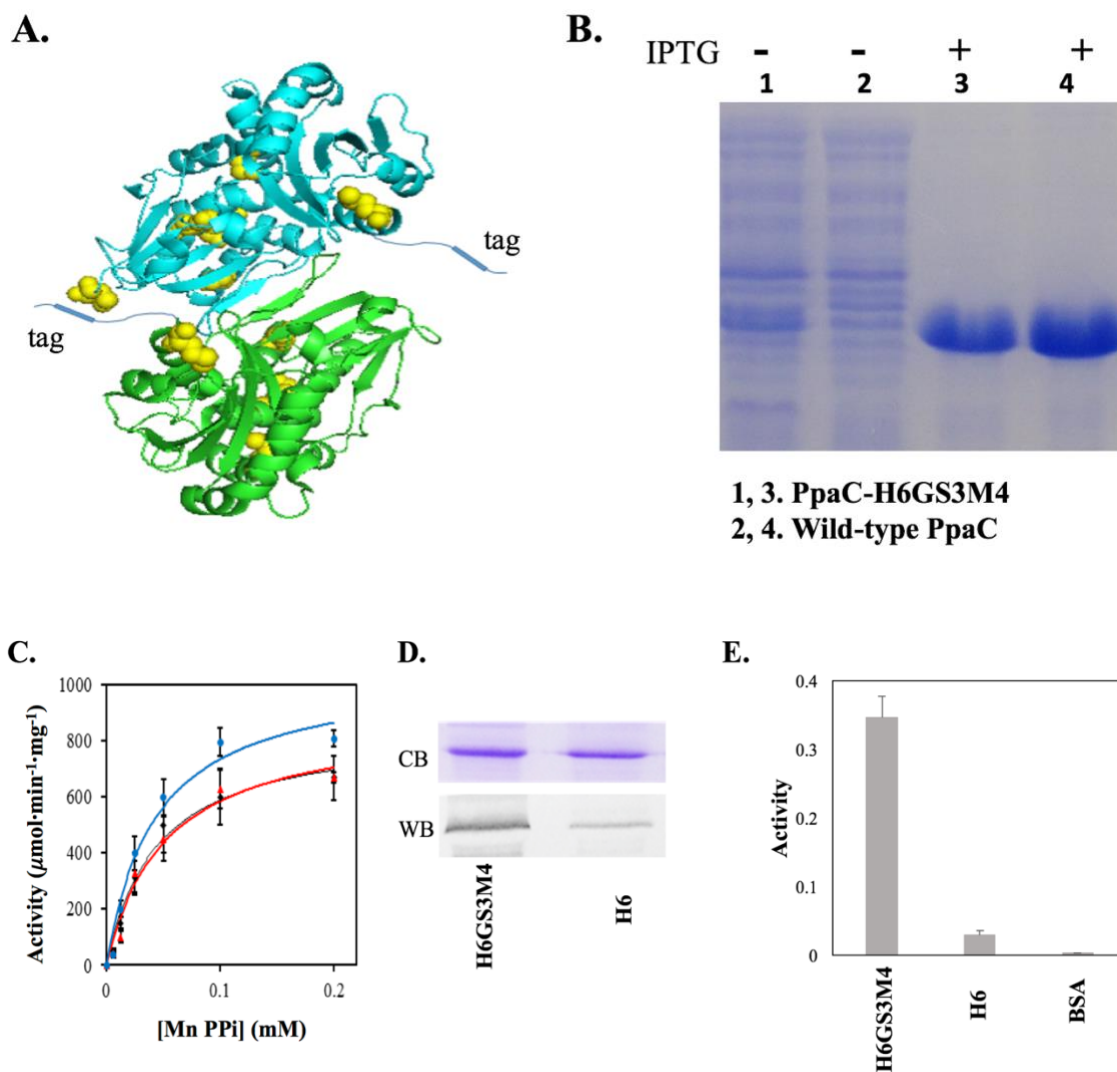
protein can be expressed and purified at high yield, and its catalytic activity can be measured using a convenient colorimetric assay. We have previously determined the crystal structure of PpaC-H6<sup>164</sup>. It exists as a dimer and each subunit contains 9 intrinsic Met residues (Figure 2.7 A). Similar as sf-GFP, this Met-containing tag would not affect the expression of PpaC either (Figure 2.7 B).

To examine the effect of the tag and the incorporation of AHA on the catalytic activity of PpaC, we first compared the catalytic activity of PpaC-H6 and PpaC-H6GS3M4 expressed in *DL41(DE3)* grown in the presence of Met (Figure 2.7 C, black and red). The fit to the Michaelis-Menten equation yielded the  $K_M$  and  $k_{cat}$  of  $44 \pm 9 \mu\text{M}$  and  $887 \pm 97 \mu\text{mol}\cdot\text{min}^{-1}\cdot\text{mg}^{-1}$ , respectively, for PpaC-H6GS3M4, which are not significantly different from the  $K_M$  and  $k_{cat}$  of PpaC-H6 ( $52 \pm 15 \mu\text{M}$  and  $897 \pm 98 \mu\text{mol}\cdot\text{min}^{-1}\cdot\text{mg}^{-1}$ , respectively). Next, we compared the  $K_M$  and  $k_{cat}$  of PpaC-H6GS3M4 expressed with AHA (Figure 2.7 C, red and blue). The  $K_M$  and  $k_{cat}$  of the protein expressed from cells grown in the presence of AHA are  $43 \pm 11 \mu\text{M}$  and  $1050 \pm 101 \mu\text{mol}\cdot\text{min}^{-1}\cdot\text{mg}^{-1}$ . The  $K_M$  and  $k_{cat}$  differences between the AHA-containing PpaC and non-AHA-containing PpaC are not statistically significant.

To compare the level of labeling of PpaC-H6 and PpaC-H6GS3M4, we expressed the two proteins in the presence of AHA as described. Although each PpaC subunit contains 9 intrinsic Met, the level of labeling of PpaC-H6 was much lower than the level of labeling of PpaC-H6GS3M4 (Figure 2.7 D). While similar amounts of proteins were used in the

experiment, the band intensity of PpaC-H6GS3M4 in the anti-biotin Western-blot was approximately four times the band intensity of PpaC-H6. In spite of the abundance of the intrinsic Met in the protein, the addition of a AHA(methionine) tag greatly improved the efficiency of modification.

Finally, we examined the usefulness of the methionine-containing tag in the immobilization of PpaC. PpaC-H6GS3M4 expressed in the presence of Met or AHA were incubated with alkyne agarose resin for immobilization as described in Materials and Methods. After immobilization, the resin was washed to remove free protein and the pyrophosphatase activity of protein attached to the resin was measured (Figure 2.7 E). For control, the alkyne agarose resin was incubated in the presence of Bovine serum albumin (BSA). The activity of immobilized PpaC-H6GS3M4 were measured as described in Materials and Methods. Without AHA, there was a low level of adsorption of PpaC-H6GS3M4 to the agarose resin, leading to a low level of activity. When AHA was incorporated into the structure of PpaC, the level of activity was increased approximately eight-fold, indicating more efficient immobilization. The control sample with BSA displayed no detectable activity.



**Figure 2.7** Structure, catalytic activity, and labeling of PpaC. **A.** Ribbon diagram of the structure of PpaC-H6 (created from 4RPA.pdb). The two subunits in the dimer are color-coded. Methionines in the structure are highlighted in yellow ball-and-stick models. **B.** Cell lysate of *DL41(DE3)* transformed with the corresponding plasmids (for PpaC) before or after IPTG induction. **C.** Michaelis-Menten plot of PpaC-H6 (black), PpaC-H6GS3M4 (red), and PpaC-H6GS3M4 containing AHA (blue). Each experiment was repeated three times. The average value and standard deviations are shown. Data was fitted using the Michaelis-Menten equation to determine  $K_M$  and  $k_{cat}$ . **D.** Commassie blue stain (CB) and



anti-biotin Western blot (WB) analysis of AHA-containing PpaC-H6 and PpaC-H6GS3M4 with biotin alkyne. **E.** Activity of PpaC immobilized on alkyne agarose resin. Each experiment was repeated three times. The average value and standard deviations are shown.

## **2.4 Conclusion**

In summary, we have shown that AHA-containing tags introduced at the C-terminus of two model proteins significantly increased the efficiency of the click chemistry reaction. The design of the tag is flexible. We tested three sequences, with two, three, or four extra Met in the tag respectively. All of them increased the level of labeling when the AHA-containing version of the respective protein was reacted with biotin alkyne. When combined with a polyhistidine tag, the histidine-AHA tag can facilitate both noncovalent interaction such as metal affinity purification and covalent binding via reaction with the alkyne functional group. We avoided the incorporation of multiple Met/AHA residues in a continuous stretch to avoid the potential problem with creating a hydrophobic patch, which may interfere with protein folding or non-specific interactions. Other than that, the design of the tag is really very flexible. The versatility of the tag design makes the concept broadly useful in the development of strategies to label and modify various proteins, and immobilize proteins onto diverse alkyne functionalized supporting matrices. A limitation of this method is the compatibility of the tag with specific protein structures. While in most cases a short peptide tag as described in this study is not expected to drastically affect the structure and function of the protein of interest, there is always a possibility that certain proteins cannot tolerate such modifications. Also, some proteins may not like being zapped with the click reaction components.

Table 2.1 Primer and tag sequences

Construct	Primers	Tag sequence
sfGFP-H7M2	5'-atg cac atg tgagatccggctgctaaca-3'	HHHHHHMHM
	5'-ggatctcacatgtgcatgtgggtgggtgggtg-3'	
sfGFP-H6M4	5'-catatgcacatgcacatgcatcactaacggctgctaacaagc-3'	HMHMHMHMHH
	5'-gcatgtgcatgtgcatatgcatgtgggatcctttagag-3'	
PpaC- H6GS3M4	5'-gcatgagcatgagcatgtgataatgagatccggctgc-3'	HHHHHHGMSMSMSM
	5'-gctcatgctcatgctcatgccatgatgggtgggtgggtgctcgag-3'	

## Chapter III Impairment of *S. aureus* PpaC activity by metal ion binding and dimer dissociation

### 3.1 Introduction

Inorganic pyrophosphatase (PPase) catalyzes the hydrolysis of inorganic pyrophosphate. It provides a thermodynamic pull for many biosynthetic reactions and is an essential enzyme in all domains of life <sup>171-173</sup>. Soluble inorganic PPases comprise two families, which are distinct in both amino acid sequence and structure <sup>174-178</sup>. Family I PPases occur in both prokaryotic and eukaryotic cells, whereas family II PPases occur almost exclusively in bacteria and have no homologue in multicellular eukaryotes <sup>179</sup>. Monomers of family I PPases have a simple cup-like single-domain structure, while monomers of family II PPases consists of two domains, with the active site at the domain interface. The catalytic site in family I PPases is a preformed structure and catalysis does not involve significant conformational change in the enzyme. In family II PPases, the C-terminal domain contains the high affinity substrate-binding site while the N-terminal domain contains the nucleophile-coordinating metal cations. Substrate binding to the C-terminal domain switches the enzyme from an open state to a closed state, which is required for catalysis <sup>180-183</sup>. Metal ion cofactor specificities of family I and II PPases for activity are also different. Family I PPases use  $Mg^{2+}$ , while family II PPases prefer  $Mn^{2+}$ .

*Staphylococcus aureus* (SA) PpaC was first identified as a family II PPase based on sequence homology <sup>184</sup>. We recently determined its  $Mn^{2+}$ -complexed structure by X-ray crystallography at 2.1 Å resolution <sup>164</sup>. The crystal structure and size-exclusion chromatography showed that SA PpaC, similarly to other characterized family II PPases,

is a dimer<sup>179</sup>. The active site, which is located at the interface between the N-terminal and C-terminal domains, contains two Mn<sup>2+</sup> binding sites, each half-occupied, which indicates that in the absence of the substrate one or the other, but not both of active sites can be occupied, with equal probabilities. In our structure, the two domains of SA PpaC form a closed active site, despite the absence of the substrate or products in the active site.

An important and yet unanswered question about the mechanism of family II PPases is the functional relevance of dimerization, as the active site is not at the dimer interface. To investigate if monomeric PpaC is functional, we created point mutations at the dimer interface. These mutations are not directly involved in the structure or formation of the active site, and thus should not directly affect the interaction and binding with substrate. We characterized the structure and activity of the mutants, as well as investigated the inhibitory effect of metal ions on SA PpaC.

## **3.2 Materials and methods**

### **3.2.1 Cloning and mutagenesis**

Plasmid pET22-PpaC for SA PpaC expression was constructed in a previous study<sup>164</sup>. A vector-derived his<sub>6</sub>-tag was introduced at the C-terminus of PpaC to facilitate purification via immobilized metal ion affinity chromatography (IMAC). Site directed mutagenesis was conducted using the QuikChange mutagenesis kit following the manufacturer's instructions (Agilent Technologies). DNA sequences were confirmed by sequencing.

### 3.2.2 Expression and purification

Plasmids encoding wild-type or mutant PpaC were transformed into *E. coli* strain ER2566 (ATCC) for expression. Cells were first cultured overnight at 37 °C in LB broth containing 100 mg/L ampicillin, which was then used to inoculate fresh LB-ampicillin media at a 50-fold dilution. The culture was incubated with shaking at 250 rpm at 37 °C until the absorbance at 600 nm reached 0.6, which usually took ~2.5 h. IPTG was added to a final concentration of 1 mM, and the culture was incubated for an additional 4 h. Cells were harvested by centrifugation at 6000×g for 10 min at 4 °C and re-suspended in 30 mL of lysis buffer (30 mM Tris-Cl, 0.3 M NaCl, 2 mM β-mercaptoethanol, 0.5 mM PMSF, pH 8.0). The cell suspension was sonicated to rupture the cells. The supernatant was collected after centrifugation at 10,000×g for 10 min at 4°C and imidazole was added to a final concentration of 7.5 mM. The solution was loaded onto a Ni-NTA (Qiagen) column equilibrated with the same buffer. The column was washed with 10 bed volumes of the lysis buffer supplemented with 40 mM imidazole. SA PpaC was eluted with lysis buffer containing 0.5 M imidazole. Purity of SA PpaC was examined by SDS-PAGE on a 12% Tris-Cl polyacrylamide gel after Coomassie blue stain. Imidazole was removed via dialysis using a membrane with 3.5 kD molecular weight cutoff. Protein samples were dialyzed for 2 days at 4°C against the 100-fold larger volume of buffer A (30 mM Tris and 0.1 M NaCl, pH 7.5), with two buffer changes. To remove metal ions, purified SA PpaC was incubated with 10 mM EDTA for 2 h at room temperature, and then dialyzed against buffer A as described above.

### 3.2.3 Pyrophosphatase activity assay

The PPase activity was measured using Mn pyrophosphate as the substrate as described <sup>167</sup>. Stock solutions of MnCl<sub>2</sub> and sodium pyrophosphate were mixed at a 1:1 molar ratio at 0.5 mM right before the analysis (mixing equal volume of 1.0 mM MnCl<sub>2</sub> and 1.0 mM sodium pyrophosphate) and diluted to the indicated substrate concentration for the activity measurement. At neutral pH, pyrophosphate is not fully deprotonated. The major species is MnH<sub>2</sub>PPi (MnPPi), which was considered the substrate for pyrophosphatase <sup>168</sup>. A stock solution of malachite green (0.12%, w/v) was made by dissolving the dye in 3 M sulfuric acid. A working solution was always made fresh by adding one volume of 7.5% (w/v) ammonium molybdate into four volumes of the malachite green stock solution followed by the addition of Tween 20 to a final concentration of 0.2% (v/v). For activity measurement, purified SA PpaC was added into a freshly prepared reaction mixture containing the indicated concentration of MnPPi, in a reaction buffer (25 mM Tris-Cl, 50 mM NaCl, pH 7.0) at room temperature for 5 min. For the metal cofactor effect experiments, sodium PPi was used as the substrate. To stop the enzymatic reaction and determine the phosphate concentration of a sample, one volume of the working solution was mixed with four volumes of the enzymatic reaction mixture to be analyzed. The mixture was incubated for 5 min for the color to develop, and the absorbance at 630 nm was measured in a 1 cm pathlength cuvette on a BIOMATE 3" UV/vis spectrophotometer (Thermo Scientific). We have experimented with the conditions and confirmed that under these conditions, the product accumulation over time was linear. Therefore, the absorbance at 630 nm, after correction for background, directly correlates with the rate of hydrolysis. To obtain relative activities, the absorbance of the sample containing Mn<sup>2+</sup> was designated as 1 and used to

normalize the absorbance of other samples. The Michaelis-Menten constant  $K_M$  and specific enzymatic activity were determined using a nonlinear fit of the measured values with SigmaPlot. The inhibition constant  $K_I$  was obtained from fitting the measured values using:  $v = \frac{v_{max}[s]}{K_M\left(1+\frac{[I]}{K_I}\right)+[s]}$ , [s] and [I] are the concentrations of the substrate and inhibitor, respectively.

### **3.2.4 Circular Dichroism (CD) spectroscopy**

CD experiments were performed using a JASCO J-815 CD spectrometer with a Peltier temperature controller. Blank scans were collected from dialysis buffer and subtracted from the spectra containing enzyme. Cuvettes of 1 mm pathlength were used for far UV scans in the wavelength range of 190-250 nm. Thermal denaturation was measured by monitoring the ellipticity at 222 nm with the increase of temperature at 2.0 deg/min and recorded every 4°C with 8 s equilibration time.

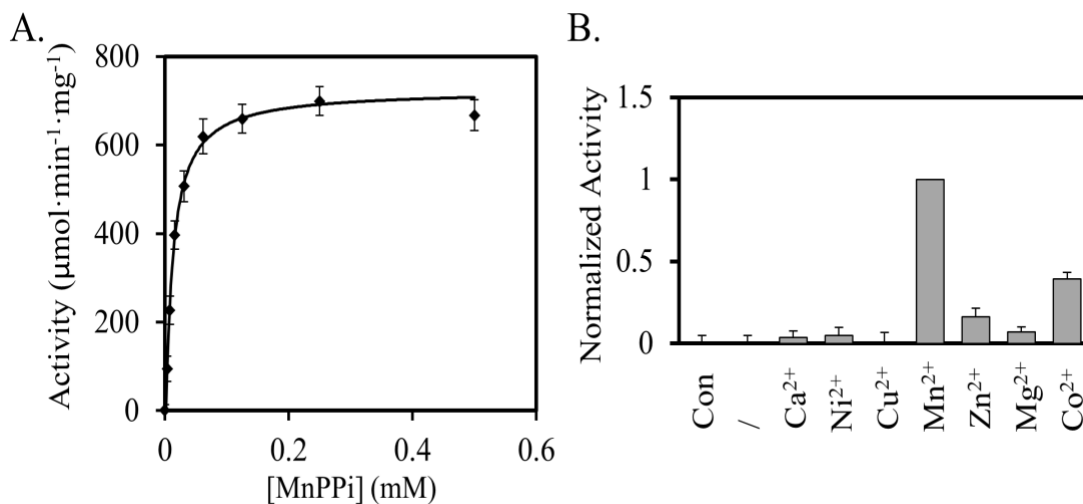
### **3.2.5 Gel electrophoresis and size exclusion chromatography**

SDS-PAGE analysis was performed using a 12% Tris-glycine gel with a Mini-Protean Tetra Cell system (Bio-Rad Laboratories, Inc.), at 200 V and room temperature until the dye front migrated out of the gel. Native gel electrophoresis was conducted similarly except that SDS was not added into the gel or running buffer. Size exclusion chromatography was performed using a Bio-Sil SEC 250 gel filtration column on a Bio-Rad Duo Flow HPLC system, using buffer A as the mobile phase.

### 3.3 Results and discussion

#### 3.3.1 Effect of metal binding on activity

Family II PPases are reported to prefer  $\text{Mn}^{2+}$  as a co-factor at their active site <sup>179</sup>. Therefore, we first measured the steady-state kinetics of  $\text{PP}_i$  hydrolysis by SA PpaC in the presence of  $\text{Mn}^{2+}$  by monitoring the phosphate production with the malachite green assay (Figure 3.1 A). Metal ions co-purified with SA PpaC were first removed by a treatment with EDTA as described above, and then  $\text{Mn}^{2+}$  was added to the concentrated SA PpaC at a final concentration of 1 mM. The pre-incubation with  $\text{Mn}^{2+}$  was critical for the catalytic activity of SA PpaC. When EDTA-treated SA PpaC was used directly in the activity assay using MnPPi as the substrate, the observed activity was less than 0.1% of the activity of the  $\text{Mn}^{2+}$  pre-incubated sample. For the assay, the protein sample was diluted  $\sim 10^4$ -fold into the reaction mixture; therefore, the residual amount of free  $\text{Mn}^{2+}$  originated from the protein solution was negligible compared to the final concentration of  $\text{Mn}^{2+}$  in the reaction mixture. The fit to the Michaelis-Menten equation yielded the  $K_M$  and  $k_{\text{cat}}$  of  $29 \pm 6 \mu\text{M}$  and  $740 \pm 40 \mu\text{mol} \cdot \text{min}^{-1} \cdot \text{mg}^{-1}$ , respectively.





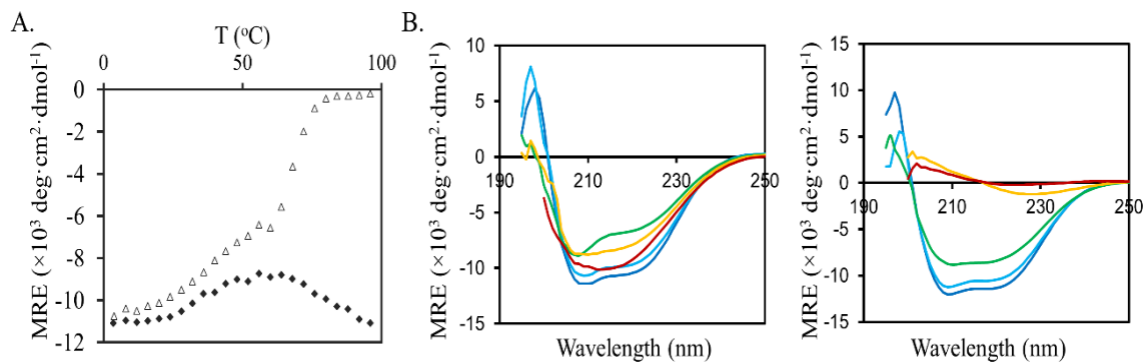
**Figure 3.1** Catalytic activity of SA PpaC. **A.** Michaelis-Menten plot of SA PpaC containing  $\text{Mn}^{2+}$ . The experiment was repeated six times. The average value and standard deviations are shown. Data was fitted using the Michaelis-Menten equation to determine  $K_M$  and  $k_{\text{cat}}$ . **B.** Relative activity of SA PpaC in the presence of different divalent cations.  $\text{Mn}^{2+}$  is the preferred metal ion for catalytic activity.

To determine the preference of metal cofactor, EDTA-treated PpaC sample was first diluted into buffer A supplemented with the indicated metal ions and incubated for 20 min. The corresponding metal ions were also added to the reaction mixture at a final concentration of 1.0 mM each before the addition of the appropriate amount of PpaC. The activity was then measured as described in Materials and Methods. As shown in Figure 3.1 B, while several metal ions could support activity,  $\text{Mn}^{2+}$  was the most effective, followed by  $\text{Co}^{2+}$ . This finding is consistent with the previously reported metal preference of family II PPases from other bacteria <sup>182</sup>.

### **3.3.2 Effect of metal ion binding on protein stability**

The thermal denaturation of SA PpaC was monitored in the absence/presence of  $\text{Mn}^{2+}$  to assess the effect of metal binding on protein stability (Figure 3.2 A). Ellipticity at 222 nm was recorded using CD during heating. Interestingly, the protein unfolding profiles differed significantly. Both the transition temperatures and the final unfolded states were different. In the presence of  $\text{Mn}^{2+}$ , PpaC displayed an unfolding curve with two transitions, a low temperature change initiated at 20 °C, and a high temperature change initiated at 50 °C. By

contrast, in the absence of metal ions, the ellipticity at 222 nm initially increased with the rise in temperature and then decreased, producing a bell-shaped profile. The unfolding of SA PpaC is clearly not cooperative, which is drastically different from the unfolding process of a close homologue *Bacillus subtilis* (BS) PpaC<sup>185</sup>. Metal-free BS PpaC unfolds cooperatively with a transition temperature at ~50 °C, while the presence of Mn<sup>2+</sup> leads to an increase of transition temperature to above 70 °C. Considering the high levels of similarity between the overall structures of the two proteins, we speculate that the difference in the unfolding behavior originates from the difference in tertiary structure packing and mode of interaction with Mn<sup>2+</sup>.



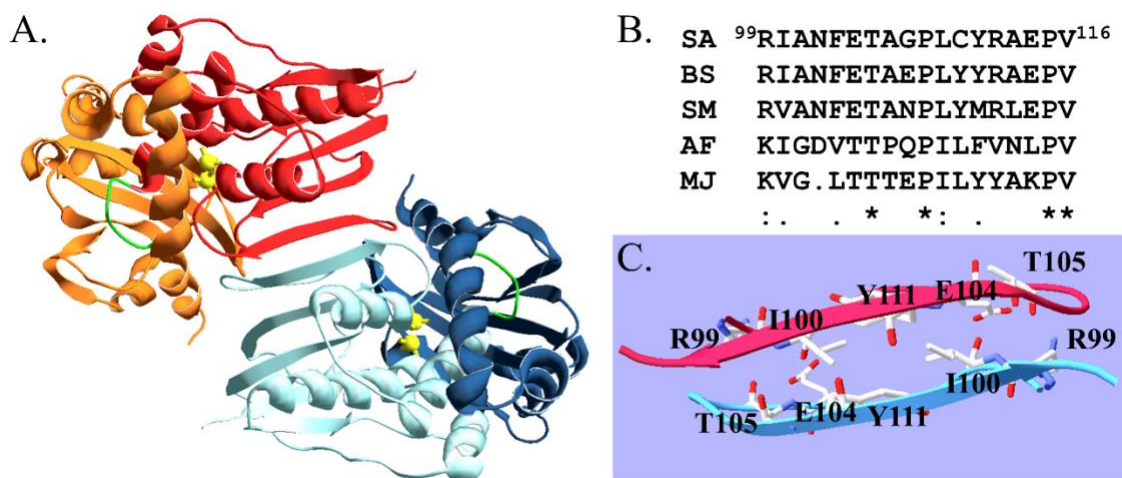
**Figure 3.2** The presence of bound Mn<sup>2+</sup> affects the thermal denaturation of PpaC. **A.** Mean residue ellipticity monitored at 222 nm upon heating in the absence of metal ions (black diamonds) or in the presence of 1.0 mM Mn<sup>2+</sup> (open triangles). **B.** Wavelength scans performed in the absence (left) or presence (right) of 1.0 mM Mn<sup>2+</sup>, at 8°C (blue), 28°C (cyan), 48°C (green), 68°C (orange), and 88°C (red).

To assign the signal to changes in secondary structure, we collected wavelength scans of PpaC in the absence or presence of 1.0 mM  $Mn^{2+}$  at five temperatures, 8, 28, 48, 68, and 88 °C (Figure 3.2 B). With or without  $Mn^{2+}$ , the spectra were similar at 8 and 28 °C. The two negative peaks at 208 and 222 nm suggest a high content of  $\alpha$ -helical structure. The enzyme structures in the absence and presence of the metal ion began to deviate from each other starting from 28 °C. In the presence of  $Mn^{2+}$ , PpaC steadily lost its secondary structure with increasing temperature, as signaled by the decrease of the negative peak. In the absence of the metal ion, the secondary structure content initially decreased as well, reaching a minimum level between 48 to 68 °C, upon which it underwent a structural transition into a mainly  $\beta$ -sheet structure. Thus, rather than increasing protein stability by increasing the melting temperature, metal ion binding seemed to help prevent the misfolding of SA PpaC into an alternative, potentially aggregation-prone conformation rich in  $\beta$ -sheet secondary structure.

### **3.3.3 PpaC mutations that disrupt the dimer interface**

Family II PPases function as homodimers. Binding to metal ion cofactors has been shown to strongly favor dimerization, accompanied by the acquisition of catalytic activity<sup>185, 186</sup>. However, it remains unclear if monomers are active. Metal binding is known to be required for catalytic function. Because metal binding strongly promotes dimerization of the wild-type family II PPases, it has not been possible to separate the two processes. To assess if the dimerization can be decoupled from metal binding and/or enzymatic activity, we designed and constructed three mutants. A PpaC monomer contains two globular domains, a N-terminal domain (NT) and a C-terminal domain (CT) connected by a linker, with the

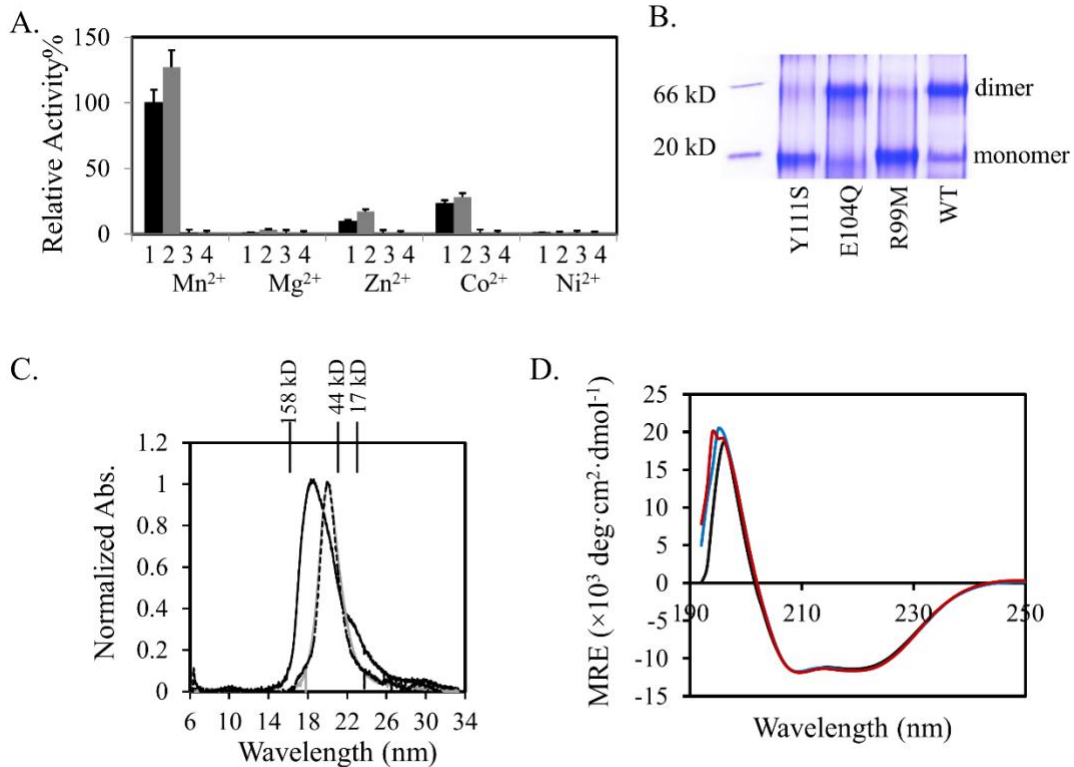
active site at the domain interface (Figure 3.3 A). Two monomers in a PpaC dimer interact via a hairpin loop involving a  $\beta$ -strand from the N-terminal domain of each monomer. The sequence of this hairpin, especially residues R99 to P115, is highly conserved (Figure 3.3 B). Since this region is not involved in the active site, we speculate that replacement of some of these residues may decrease dimer stability without affecting the overall protein structure including the active site. We identified three residues to mutate, R99, E104, and Y111 (Figure 3.3 C). R99 and E104 may potentially form an inter-subunit salt bridge. Y111 is engaged in hydrophobic interactions with I100, Y111, and A113 from the neighboring subunit. We created three single mutations, Y111S, R99M, and E104Q, aiming at disrupting dimerization.



**Figure 3.3** Structure and sequence of PpaC. **A.** Crystal structural of SA PpaC dimer complexed with  $Mn^{2+}$  (PDB ID: 4RPA<sup>164</sup>). The two monomers are color coded, with  $Mn^{2+}$  ions illustrated as yellow spheres. The NT domains in the two subunits are red and light blue, while the CT domains are orange and dark blue. The short linker between the domains is green. **B.** Sequence alignment of family II PPases from *S. aureus* (SA), *B. subtilis* (BS), *Streptococcus mutans* (SM), *Archaeoglobus fulgidus* (AF) and *Methanocaldococcus*

*jannaschii* (MJ). Identical, highly conserved, and semi-conserved residues are marked by asterisks, colons, and dots, respectively. **C.** A zoomed-in view of the hairpin loops that form the dimer interface. The side chains of residues R99, I100, E104, T105, and Y111 are highlighted.

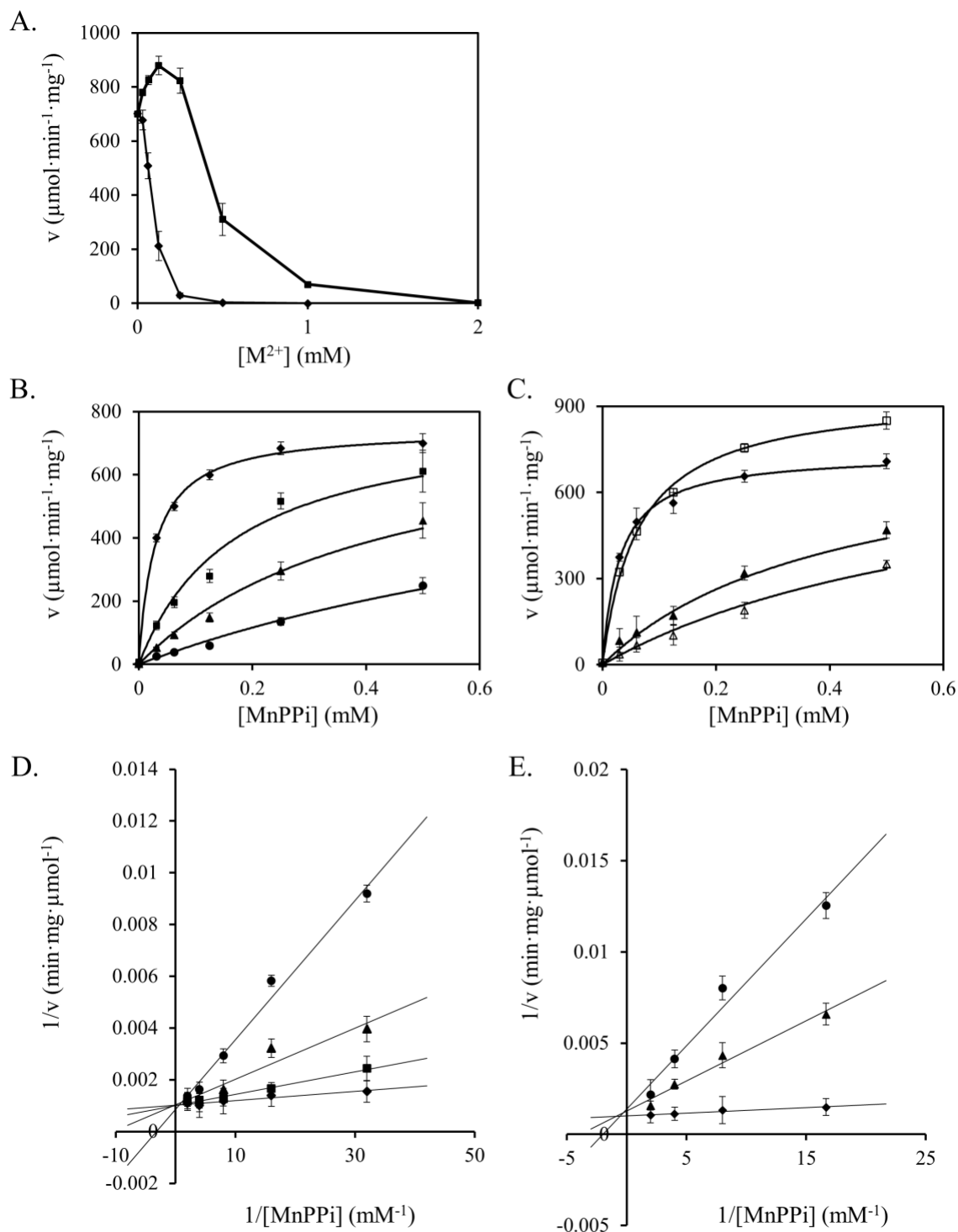
The catalytic activities of the three mutants were compared in the presence of five different metal ions. Y111S and R99M were completely inactive, while E104Q was largely active (Figure 3.4 A). To further verify the effect of mutations on the structure of PpaC, we expressed and purified the mutants for further characterization. Native gel electrophoresis was used to determine the oligomeric state of the proteins. For wild type and E104Q PpaC, a strong dimer band could be observed together with a weak monomer band (Figure 3.4 B). For the functionless mutants, the dimer band was much weaker and the monomer band became dominant. These results indicate that Y111S and R99M mutations weakened the dimerization. This conclusion was further confirmed by the size exclusion chromatography experiment, in which the mutants eluted mainly as monomers, while the wild type PpaC eluted as a dimer in equilibrium with a smaller population of monomer (Figure 3.4 C). Finally, the CD spectra of Y111S and R99M were compared with that of the wild type PpaC. The overall structures of the mutants were very similar to that of the wild type protein, confirming that the loss of activity was not due to global unfolding of the structure of the enzyme (Figure 3.4 D).



**Figure 3.4** SA PpaC mutants disrupted the dimer interface. **A.** Metal ion effect on the relative activity of wild type (1), E104Q (2), Y111S (3), and R99M (4) in the presence of 1.0 mM of the indicated metal ions. **B.** Native gel electrophoresis analysis of purified PpaC constructs. Wild type PpaC and E104Q migrated mainly as dimers, while Y111S and R99M migrated mainly as monomers. **C.** SEC traces of wild type SA PpaC (black continuous) and two mutants Y111S (grey continuous) and R99M (black dashed). Retention times of molecular weight standards are marked at the top of the plot. **D.** Far UV CD spectra show that the secondary structures of Y111S (blue) and R99M (red) are highly similar to that of the wild type PpaC (black).

### 3.3.4 Inhibitory effect of $\text{Mn}^{2+}$ and $\text{Mg}^{2+}$

When optimizing the condition for enzyme activity assay, we noticed that while  $\text{Mn}^{2+}$  was critical for PpaC activity, extra free  $\text{Mn}^{2+}$  (supplied in the form of  $\text{MnCl}_2$ ) was inhibitory (Figure 3.5 A). To investigate the mechanism of inhibition, we measured the catalytic activity of the protein as a function of  $\text{Mn}^{2+}$  concentration (Figure 3.5 B). As described in the Materials and Methods,  $\text{MnCl}_2$  and sodium pyrophosphate were mixed at a 1:1 molar ratio to make a 0.5 mM solution, which was then subjected to a serial two-fold dilution and used as the substrate. Enzyme activity was measured with the addition of extra  $\text{MnCl}_2$  at the indicated concentrations (80, 120, and 160  $\mu\text{M}$ ). The Lineweaver-Burk plots measured in the presence of different  $\text{MnCl}_2$  concentrations had very similar y-intercepts, indicating that the mechanism of inhibition is competitive (Figure 3.5 D). The inhibition constant  $K_I$  was determined to be  $15 \pm 5 \mu\text{M}$ . Similarly, we found that  $\text{Mg}^{2+}$  was also a competitive inhibitor at higher concentration (Figure 3.5 A, 3.5 C, and 3.5 E). The  $K_I$  was determined to be  $48 \pm 10 \mu\text{M}$ . However, at lower concentration,  $\text{Mg}^{2+}$  modestly stimulated the catalysis.



**Figure 3.5**  $\text{Mn}^{2+}$  and  $\text{Mg}^{2+}$  are competitive inhibitors of SA PpaC activities. **A.** Pyrophosphatase activity measured with 0.25 mM MnPPi as the substrate, in the presence



of MgCl<sub>2</sub> (squares) or extra MnCl<sub>2</sub> (diamonds). **B.** Pyrophosphatase activity measured in the presence of increasing concentrations of MnCl<sub>2</sub> (circles, 0.16 mM; triangles, 0.12 mM; squares, 0.08 mM; diamonds, no MnCl<sub>2</sub>). Curves are fitting of the data using the Michaelis-Menten equation as described in Materials and Methods. **C.** PpaC activity measured in the presence of increasing concentrations of MgCl<sub>2</sub> (circles, 1 mM; triangles, 0.5 mM; squares, 0.125 mM; diamonds, no Mg<sup>2+</sup>). Curves are fitting of the data using the Michaelis-Menten equation as described in materials and methods. **D.** Lineweaver-Burk plots revealing that Mn<sup>2+</sup> functioned as a competitive inhibitor. The same markers are used for the corresponding concentrations as in **B.** **E.** Lineweaver-Burk plots revealing that Mg<sup>2+</sup> also functioned as a competitive inhibitor. The same markers are used for the corresponding concentration as in **C.** All experiments were repeated six times. The average value and standard deviations are shown.

### **3.4 Conclusion**

In summary, we have characterized a family II inorganic pyrophosphatase from *S. aureus*. One unanswered question about the function of family II inorganic PPase is the functional relevance of dimerization. Since Mn<sup>2+</sup> binding, which is indispensable for catalytic activity, drives dimerization, it has not been possible to investigate the activity of monomeric PpaC. Using site-directed mutagenesis, we found that monomeric mutants containing single residue replacements at the dimer interface were completely inactive, indicated that dimerization is essential for the catalytic activity in the SA PpaC and likely for other family II PPases. In a previous study Halonen et al. found that mutations of H97 and H98 residues in BS PpaC, which are located at the subunit interface, greatly destabilized the BS PpaC dimer<sup>185</sup>. H97 and H98 are conserved in SA PpaC and are involved in Mn<sup>2+</sup> binding.

Mutation of R99 may have disrupted the local structure of the  $\text{Mn}^{2+}$  binding site as well and impaired the function of SA PpaC from both aspects of dimerization and metal ion binding.

Metal ion binding to family II inorganic PPases has been the subject of several studies<sup>180-183, 185-187</sup>. Family II inorganic PPases from different organisms have been reported in different studies to bind 2 to 4 metal ions, depending on the presence or absence of a substrate analog. One interesting observation in our previous study is the low  $\text{Mn}^{2+}$  to SA PpaC stoichiometry. In the absence of substrate, the wild type SA PpaC bound to one  $\text{Mn}^{2+}$  per subunit, which occupies one of two closely positioned sites at the substrate binding pocket<sup>164</sup>. This result is consistent with a previous study using solution equilibrium dialysis, which indicates that in the absence of substrate, family II PPases have one high affinity site for  $\text{Mn}^{2+}$  per monomer with a  $K_d$  in the nM range<sup>186</sup>. In the current study we found that  $\text{Mn}^{2+}$  functions as both a critical active-site co-factor and a competitive inhibitor, with similar apparent affinities ( $K_M \sim 30 \mu\text{M}$  and  $K_I \sim 15 \mu\text{M}$ ). In addition, we discovered that EDTA-treated SA PpaC needs to be incubated with  $\text{Mn}^{2+}$  to function properly. Therefore, there needs to be a  $\text{Mn}^{2+}$  at the active site before the binding of the substrate. These observations prompt us to speculate that the two half-occupied sites in the apo-PpaC crystal structure might both be occupied in the substrate-bound enzyme, one by a  $\text{Mn}^{2+}$  that enters the site during the pre-incubation, and the other by a  $\text{Mn}^{2+}$  that enters as the MnPPi complex. In the presence of excess free  $\text{Mn}^{2+}$ ,  $\text{Mn}^{2+}$  may compete with MnPPi to bind at the second site. In the case of  $\text{Mg}^{2+}$ , the inhibitory effect at high  $\text{Mg}^{2+}$  concentration could indicate the displacement of  $\text{Mn}^{2+}$  by  $\text{Mg}^{2+}$  from the active site, consistent with the observed

competitive inhibition mechanism. The modest and yet significant enhancement of activity at lower  $Mg^{2+}$  concentration is more difficult to speculate. It is possible that an additional metal binding site exists in the structure of the substrate-bound SA PpaC, as documented before in the case of BS PpaC<sup>182</sup>. Occupation of this site by  $Mg^{2+}$  accelerates the hydrolysis of PPi at the active site.

While manganese is known as an essential element, the physiological functions of  $Mn^{2+}$  remains poorly understood. With the discovery of an increasing number of  $Mn^{2+}$ -dependent enzymes and the abundance of  $Mn^{2+}$ -specific membrane transporters, the cellular function of manganese is attracting increasing research attention<sup>188</sup>. Many bacteria can vary cytoplasmic  $Mn^{2+}$  level over several logs of concentration, which make controlling  $Mn^{2+}$  concentration an appealing mechanism for activity regulation<sup>188</sup>. PPases play important roles in cellular metabolism and biosynthesis of macromolecules. The sensitivity of PpaC to the level of free  $Mn^{2+}$  might represent such an example of regulation. Maximum catalytic activity is only achieved in a narrow window of  $Mn^{2+}$  concentration.

## Chapter IV Structure effect on the degradation of MscS with ClpXP system

### 4.1 Introduction

Protein degradation is an important housekeeping mechanism for all cells. It removes damaged or excessive proteins in response to cellular condition and maintains protein homeostasis in the living cell. Complete degradation of protein substrates requires energy, which is necessary to unfold protein structures so they can fit into the active site of the proteases. The major class of proteases belong to the ATPases Associated with diverse cellular Activities (AAA+) family, which derive energy from ATP hydrolysis to unfold protein substrates.<sup>189-191</sup> Several types of AAA+ proteases have been discovered in bacteria, including ClpXP, ClpAP, ClpCP, HslUV, Lon and FtsH.<sup>192</sup> Lon and FtsH contain both the ATP-driven unfolding module and the proteolytic site in the same chamber.<sup>192, 193</sup> The rest are complexes of two proteins: an ATPase that unfolds the substrate (unfoldase) ClpX, ClpA, and HslU, and a peptidase that cleaves unfolded substrate ClpP or HslV.<sup>189, 194</sup> In this study, we focused on the degradation of a membrane protein substrate by ClpXP and SspB. Protein adaptor SspB enhances the binding and delivery of *ssrA* tagged protein for ClpXP degradation.<sup>122, 195</sup>

The C-terminal *ssrA* tag is an 11 amino acids sequence (AANDENYALAA in *E. coli*) that is recognized by the ClpXP complex as a signal for degradation. During protein synthesis, a truncated mRNA lacking the stop codon leads to the formation of a nonstop complex containing the mRNA, ribosome, and peptidyl-tRNA. Cells resolve this issue via a trans-translation process, in which the nonstop translation complex is released and the protein product is labeled with the *ssrA* tag at the C-terminus as a marker for degradation<sup>98, 122, 196,</sup>

<sup>197</sup>. ClpXP is one of the major players in the degradation of *ssrA* tagged soluble protein substrates. However, the degradation of *ssrA* tagged membrane protein has not been studied until recently. We found that *ssrA* tagged membrane protein AcrB could be degraded by ClpXP.<sup>198</sup> Shortly after, Sauer and co-workers reported that when they add the *ssrA* tag to a membrane protein fragment that contains the first three transmembrane helices of ProW, it could be completely degraded as well. However, FtsH, but not ClpXP, is responsible for the degradation<sup>110</sup>. Up to now these are the only *ssrA* tagged membrane protein substrates studied. Clearly, studies of the degradation of more membrane protein substrates are necessary to elucidate the mechanism of degradation and specificity. Here, we investigated the degradation of *ssrA* tagged MscS, aiming to investigate structure effects on the degradation of *ssrA* tagged membrane proteins by ClpXP proteases.

MscS is the mechanosensitive channel of small conductance. MscS directly responds to changes in membrane tension by opening nanoscale protein pores and allowing solutes to flow across the cell membrane in response to membrane tension <sup>199</sup>. Crystal structure of MscS from *E. coli* reveals a symmetric heptamer structure (Figure 4.1 A and 4.1 B) <sup>200, 201</sup>. It contains three transmembrane helices (THs), with the NT in the periplasm. After TH3, there is a large cytoplasmic domain. We cloned MscS with a CT *ssrA* tag. In addition, we added a histag followed by a flag tag at the NT of the protein, and flag tag is used as a marker for detection during degradation. Since degradation initiated by the *ssrA* tag is progressive, by putting the flag tag at the NT we can detect all partially degraded fragments if they exist. We found that the ClpXP complex can degrade *ssrA* tagged MscS both in *E. coli* cells and under reconstituted conditions in detergent solubilized state, but the presence of a periplasmic inter-subunit disulfide bond prevented complete degradation. Presence of

SspB is not essential, but greatly accelerated degradation. Finally, we developed a fluorescence polarization based method and quantitatively revealed the contribution from SspB, and the hampering effect of the disulfide bond.

## **4.2 Material and Method**

### **4.2.1 Plasmid construction**

Plasmid pET28-MscS encoding *E. coli* MscS bearing a NT histag was purchased from AddGene (Cambridge, MA). A flag tag was introduced at the N terminus right after the first Met, and the *ssrA* sequence was introduced at the C terminus, both using the fast cloning method<sup>202</sup>. The Quikchange Site Directed Mutagenesis kit (Agilent Genomics) was used to introduce point mutations including S9C, R128C, and A286C using the primer indicated in Table 4.1. All MscS constructs in this study contain the NT flag tag for detection and NT histag for purification. Plasmids and strains for expression of ClpX, ClpP, and SspB have been described in an earlier study<sup>198</sup>.

Table 4.1 Primer design for MscS mutants

Mutant	Primer forward	Primer reverse
S9C	gaagatttgaatgttgcgattgcataaacggcgcgg gaagc	gcttcccgcgccgttatgcaatcgacaacatt caaatcttc
R128C	gtgttacttgcatgttctgccccgttccgtgccggag	ctccggcacggaacgggcagaacatgacaa gtaacac
A286C	gtgaaagaagacaaagcttgcgcgccaacgatg aaaac	gttttcatcgttggccgcgcaagctttgtcttctt tcac

#### 4.2.2 Protein expression and purification

The *ssrA* tagged proteins were expressed in the ClpX knockout strain *DL41ΔclpX(DE3)*. The corresponding plasmid was transformed into *DL41ΔclpX(DE3)*, and a single colony was used to inoculate the overnight cell culture into 300 ml LB medium with 50 μg/ml Kanamycin. Cells were cultured at 37°C with shaking until absorbance at 600 nm reached 0.8, then induced with 1 mM IPTG for 4 hours. Cells were harvested by centrifugation, and the cell pellet was stored at -80°C overnight.

To purify MscS, cell pellet was resuspended in 30 ml HEPES buffer (50 mM HEPES, 0.2 M NaCl, pH 7.5) with 0.5 mM phenylmethanesulfonylfluoride (PMSF). The supernatant was removed and membrane pellet was resuspended in HEPES buffer with 2% Triton X-100. Fluorescein-5-maleimide, used for labeling with cysteine, was added to a final concentration of 1 mM when labeling was required. The mixture was incubated at 4°C with

shaking for 2 hours, followed by centrifugation at 10,000 rpm at 4°C for 20 minutes. The supernatant containing solubilized membrane proteins was mixed with Ni-NTA agarose resin (Qiagen Inc., Valencia, CA) for 40 min at 4°C. The protein-Ni-NTA suspension was packed in an empty column and washed with 20 column volumes of HEPES buffer containing 0.3% DDM (n-Dodecyl  $\beta$ -D-maltoside) and 40 mM imidazole. Finally, MscS was eluted using HEPES buffer with 0.3% DDM and 500 mM imidazole. Imidazole in the elution buffer was removed by dialysis in HEPES buffer containing 0.3% DDM overnight. For FITC-MLM labeled samples, ratio of labeling was calculated using absorbance at 280 nm and 490 nm, respectively, for protein and fluorescein, with extinction coefficient of 27,310 M<sup>-1</sup>cm<sup>-1</sup> and 68,000 M<sup>-1</sup>cm<sup>-1</sup>. The level of labeling is normally ~60%.

The expression and purification of proteases have been conducted as described with minor modifications <sup>198</sup>. Briefly, to express proteases ClpX, ClpP and SspB, *BL21(DE3)* cells containing the respective plasmid were cultured overnight. The overnight culture was used to inoculate 300 ml LB medium with proper antibiotics (ClpX and SspB are in pET28 vector with kanamycin resistance, and ClpP is in pQE70 vector with ampicillin resistance). ClpX was expressed at 28°C and induced when OD 600 nm reached 0.5 to 0.6 using 0.1 mM IPTG, followed by incubation for 5 hours. ClpP and SspB were expressed at 37°C and induced when OD 600 nm reached 0.6 to 0.8 with 0.5 mM IPTG followed by incubation for 3 hours. Cells were harvested by centrifugation at 15,300  $\times$ g and stored at -80°C overnight.

For purification, the cell pellet was re-suspended in 30 ml lysis buffer (50 mM Tris-Cl, 100 mM KCl, 2 mM MgCl<sub>2</sub>, 0.015% Triton, 10% glycerol, freshly add 0.15 g DTT to 100 ml



of lysis buffer, pH7.5). Cells were lysed using sonication (Sonicator FB120, Pittsburgh PA) in an ice-water bath for 10 min with amplitude of 75% and 5s on/ 5s off cycle. The cell lysate was centrifuged at  $15,300 \times g$ ,  $4^{\circ}C$  for 20 min. Supernatant was incubated with Ni-NTA resin for 40 min at  $4^{\circ}C$ . Resins were loaded into an empty column and drained, followed by wash with 30 column volume of dialysis buffer (300 mM KCl, 50 mM Tris HCl, 2 mM  $MgCl_2$ , 0.015% Triton, 30 % glycerol, 10 mM DTT, pH7.5) with 50 mM imidazole. Protein was eluted with dialysis buffer containing 500 mM imidazole. Finally, imidazole was removed through dialysis in dialysis buffer at  $4^{\circ}C$  overnight.

#### **4.2.3 Degradation in cells**

The proper plasmid was transformed into the indicated strain of *E. coli* cells. Cell cultures were induced at OD600 of ~0.8 with 1 mM IPTG and cultured for an additional 4 hours. To prepare samples for immunoblotting analysis, cell pellet was sonicated in dialysis buffer and centrifuged to collect membrane vesicles. The membrane vesicles were dissolved in SDS loading dye and analyzed using SDS PAGE followed by immunoblotting using an anti-flag tag antibody.

For stepwise degradation, pET28-MscS-ssrA and pBAD33-ClpX were co-transformed into *DL41ΔclpX(DE3)*. Cells were cultured in LB medium until OD600 reached 0.8, and induced with 1 mM IPTG for 2 hours. To induce the expression of ClpX, IPTG was first removed by changing the cells into fresh medium, and arabinose was added to 0.2% (w/v). At the indicated time, aliquots of cell culture were withdrawn and centrifuged to collect cells. Cell pellets were treated as described above and analyzed using anti-flag immunoblotting.

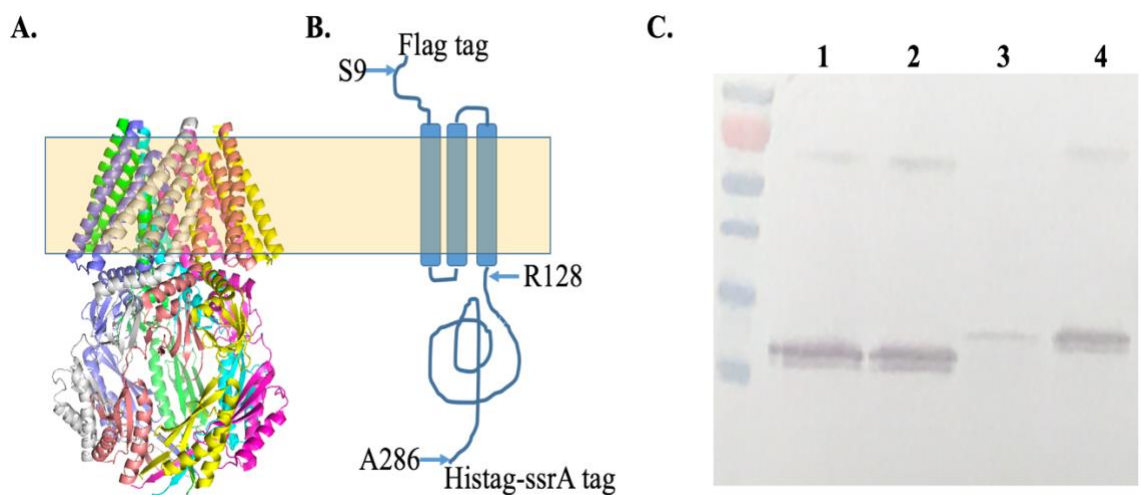
#### 4.2.4 Degradation in reconstituted system

Degradation *in vitro* was performed following published protocols at 30°C using the PD buffer (25 mM HEPES-KOH, pH7.6, 200 mM KCl, 10 mM MgCl<sub>2</sub>, 10% glycerol, 2 mM DTT, 5 mM ATP, 16 mM creatine phosphate, and 0.032 mg/mL creatine kinase) supplement with 0.03% DDM<sup>198</sup>. The purified protein was mixed with the ratio of (MscS-ssrA)<sub>1</sub>:(ClpX)<sub>6</sub>:(ClpP)<sub>14</sub>:(SspB)<sub>2</sub> = 1:3:6:3 unless otherwise noted. For the immunoblotting analysis, aliquots of samples were taken at the indicated time and put on ice. SDS loading dye was immediately added to fully stop the reaction. For FP assay, the indicated substrate was first added in PD buffer and incubated at 30°C for 2 min. The FP was then monitored at ex/em wavelengths of 490 and 520 nm until the reading stabilized, then proteases mixture of ClpXP with/without SspB (pre-incubated at 30°C) was added to initiate the degradation reaction. The degradation assay was monitored by fluorescence polarization using fluorescence spectrometer (PerkinElmer, Inc., Waltham, MA).

For the binding experiment, substrate was mixed with PD buffer and incubated at 30°C for 2 min. The ratio of ClpX to SspB was 1: 1, and kept constant. The ratio of substrate to ClpX varied from 3:1, 1:1, 1:3, to 1:10 while the concentration of the substrate was kept constant. Binding was monitored by fluorescence polarization at the excitation and emission wavelengths of 490 and 520 nm, respectively.

For quantitative measurement of degradation, to convert FP values into actual concentration of degraded MscS, we first established a calibration curve<sup>203</sup>. We obtained a commercial peptide corresponding to the last eight residues of MscS (RVKEDKAA) bearing a N-terminal FITC from Genscript. This peptide mimics the product of ClpXP

digestion, which lead to small peptide of 8-10 residues<sup>204</sup>. Concentration of FITC-RVKEDKAA and FITC-labeled full length MscS-R128C were measured using the absorbance of FITC at 490 nm with extinction coefficient of 68,000 ( $M^{-1} cm^{-1}$ ). They were then mixed at different molar ratios and the FP signal was recorded and plotted against the percentage of FITC-peptide fragment in the buffer, mimicking different stages of degradation.



**Figure 4.1** Crystal structure and expression of MscS-ssrA. **A.** Crystal structure of MscS created from 2oau.pdb<sup>201</sup>. Yellow frame marked the location of the cell membrane. **B.** Topology model of MscS. A flag tag was introduced the NT, and a histag followed by a ssrA tag were introduced at the CT. **C.** Anti-flag immunoblotting of MscS and MscS-ssrA in indicated cells. Introduction of CT ssrA tag lead to MscS degradation in *E. coli*. Removal of ClpX from the genome aborted the degradation. 1) MscS in DL41(DE3) 2) MscS in DL41 $\Delta$ ClpX (DE3) 3) MscS-ssrA in DL41(DE3) d) MscS-ssrA in DL41(DE3)  $\Delta$ ClpX. From top to bottom, bands in the molecular weight markers are 95, 72(red), 55, 43, 34, 26, and 17 KD.

### **4.3 Result and discussion**

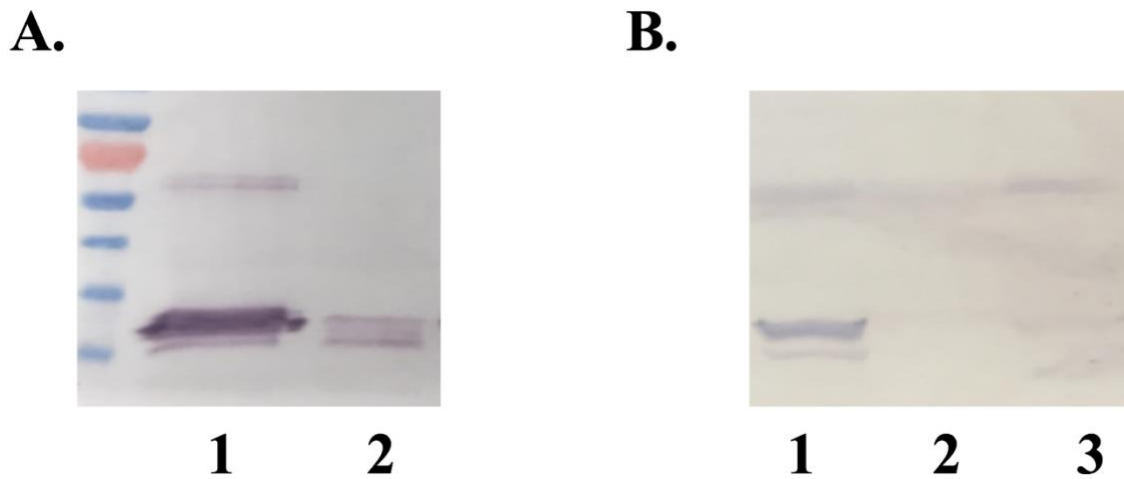
#### **4.3.1 The *ssrA* tag can facilitate degradation of full length MscS**

To investigate if the *ssrA*-tag could facilitate the degradation of MscS, we genetically introduced the *ssrA*-tag after a histag at the C-terminus of MscS to create MscS-*ssrA*. MscS-*ssrA* was completely degraded in *DL41(DE3)*, as revealed by the lack of detectable band in the anti-flag immunoblotting analysis (Figure 4.1 C). While the control construct, MscS, displayed a clear band. In the *clpX* knockout strain, both the control construct and the *ssrA*-tagged construct expressed at similar levels, indicating that the removal of ClpX in the cells is responsible for the lack of degradation of SsrA-tagged MscS. MscS-*ssrA* migrated slightly slower than MscS due to the presence of the extra 11 residues in the *ssrA* tag. This result indicate that MscS-*ssrA* can be fully degraded in *E. coli* cells, and ClpX plays a critical role in this process.

#### **4.3.2 MscS is membrane integrated**

We do not expect the presence of the flag tag and the *ssrA* tag to affect membrane integration of MscS. To confirm this speculation, we fractionated *DL41ΔclpX(DE3)* cells expressing MscS-*ssrA* into membrane fraction and soluble fraction, and detected the presence of MscS-*ssrA* in the fractions. As shown in Figure 4.2 A, the protein exists predominantly in membrane vesicles. Little could be detected in the soluble fraction. To demonstrate that ClpX and its partners could degrade membrane integrated MscS, we co-transformed pET28-MscS-*ssrA* and pBAD33-ClpX into *DL41ΔclpX(DE3)*. The

expression of MscS-ssrA was first induced using IPTG. After 2 hours, expression of ClpX was induced using arabinose. As shown in Figure 4.2 B, MscS-ssrA was expressed in the  $\Delta ClpX$  strain before ClpX expression. After the induction of the expression of ClpX, the protein disappears within 30 min. ClpX is clearly required for the degradation of MscS-ssrA.

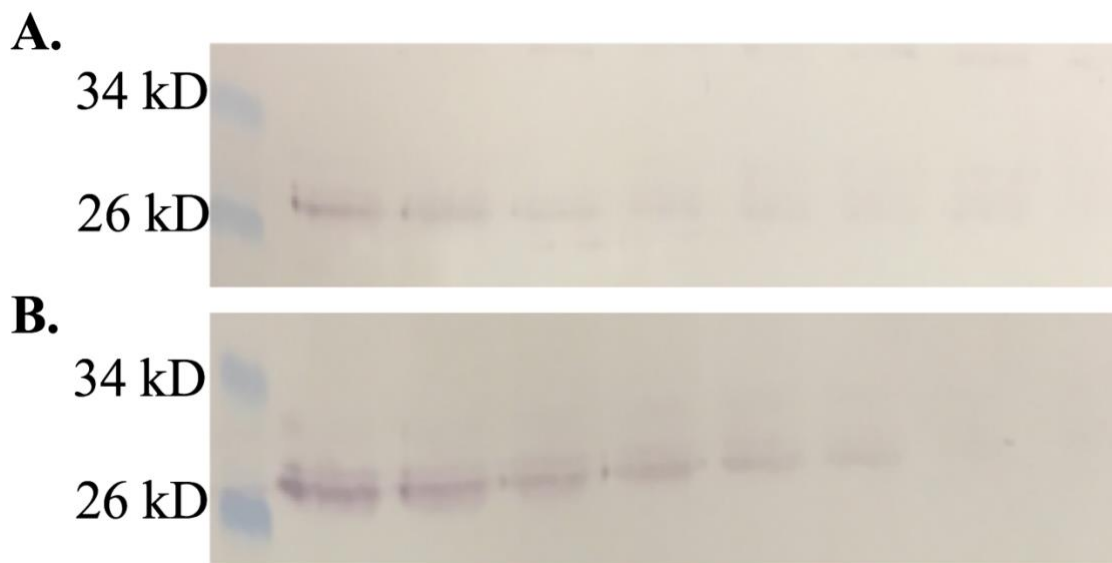


**Figure 4.2** Expression of MscS-ssrA. **A.** MscS-ssrA is present in the membrane vesicles (lane 1), not in cytosol (lane 2). From top to bottom, bands in the molecular weight markers are 95, 72(red), 55, 43, 34, and 26 kD. **B.** MscS-ssrA expressed after IPTG induction for 2 hours (lane 1), and then induced with 0.2% arabinose for 30 min (lane 2) and 1 hour (lane 3).

#### 4.3.3 Detergent solubilized MscS-ssrA can be degraded using purified ClpXP

To further investigate the mechanism of degradation, we expressed and purified MscS-ssrA in the detergent solubilized state and examined its degradation using purified ClpX and ClpP (Figure 4.3). In addition, we examined the effect of SspB, a known chaperone that enhances the interaction between ssrA tagged soluble protein substrate with ClpX. As

shown in Figure 4.3 A and 4.3 B, MscS could be completely degraded by ClpXP/SspB. The presence of SspB greatly increased the efficiency of degradation, similar to the case of soluble protein substrates.

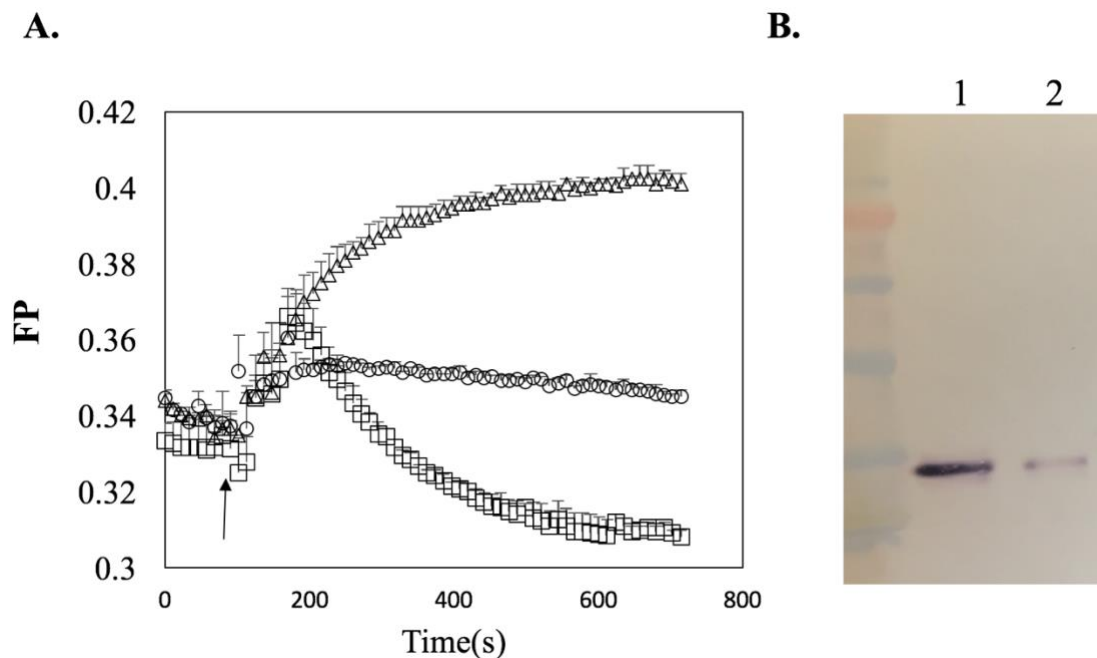


**Figure 4.3** Degradation of detergent solubilized MscS-ssrA using ClpXP. **A.** Degradation of detergent purified MscS-ssrA using purified enzymes, at ratio of 1:2:3:1. From left, the samples were taken at 0, 2.5, 5, 10, 15, 20, 30, 40 min. **B.** Same as A. except that SspB was not added during the degradation process.

#### **4.3.4 Development of a fluorescent polarization assay for real-time monitoring of degradation**

To quantitatively monitor the degradation process, we developed a FP based method.<sup>205</sup>  
<sup>206</sup> MscS does not contain Cys. We introduced a unique Cys into the sequence by site directly mutate R128 into C. We choose this site because it is well exposed and thus accessible for labeling according to the crystal structure. They are far apart in neighboring

subunits to avoid the potential formation of disulfide bond. And it is at the boundary between the transmembrane domain and cytoplasmic domain. MscS-R128C was labeled with fluorescein-5-maleimide. Upon digestion, the fluorescent labeled substrate becomes smaller, and the FP decreases with the shrinkage of the molecule until the peptide containing fluorescent labeled R128C was released from the protein. It has been reported that the ClpXP pore structure is very flexible and is able to accommodate multiple peptide chains simultaneously<sup>207</sup>. Therefore, we did not expect the labeling to affect degradation. As shown in Figure 4.4, we can clearly monitor a decrease of FP value. To confirm the protein is indeed degraded, we also took samples at different time points and conducted anti-flag western blot analysis. The reduction of FP signal correlates well with the decrease of WB band intensity, indicating that FP could be used to monitor protein degradation.



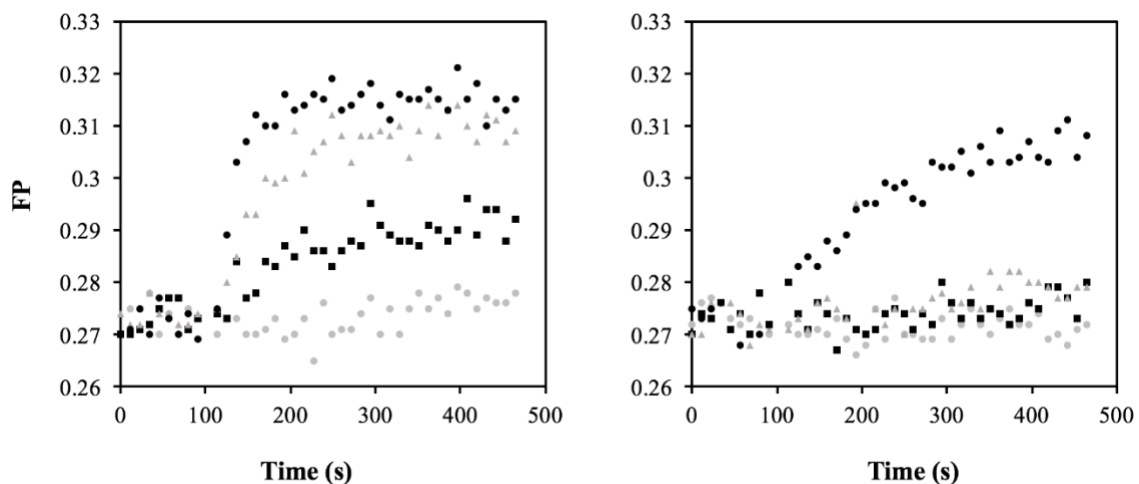
**Figure 4.4** Fluorescence polarization monitor the degradation of MscS-ssrA. **A.** Monitoring of MscS-R128C-ssrA degradation using fluorescent polarization in the

presence (squares) or absence (circles) of SspB. In the FP was monitored at ex/em wavelengths of 490 and 520 nm, respectively. In the absence of ClpP (triangles), binding to SspB and ClpX lead to the increase of FP. **B.** anti-flag WB of samples before digestion (lane 1) or 20 min (lane 2) into digestion in the presence of SspB.

#### **4.3.5 SspB increases the rate of association of ClpX with substrate**

SspB is known to enhance degradation efficiency of *ssrA* tagged substrate by ClpXP. The fluorescence polarization method offered us an opportunity to directly monitor the effect of SspB on the interaction rate between ClpX and fluorescein labeled substrate. As shown in Figure 4.5, binding of ClpX to fluorescent labeled MscS-R128C-*ssrA* lead to an increase of fluorescent polarization value. In the absence of SspB, little binding could be observed at low ClpX to MscS ratios. A slow increase of FP signal could be observed when ClpX was in 10 fold excess of MscS. In contrast, in the presence of SspB (kept at ClpX<sub>6</sub>:SspB<sub>2</sub>= 1:1 molar ratio), clear increase of FP could be observed for all concentrations tested. The initial binding rates at 10 fold excess of ClpX were increased by ~9 fold in the presence of SspB.





**Figure 4.5** SspB increases ClpX association with substrate. Monitoring of ClpX and substrate association using fluorescent polarization, with (left) or without (right) SspB. Ratio of MscS-R128C-ssrA to ClpX is 3:1 (grey circles), 1:1 (black squares), 1:3 (grey triangles), and 1:10 (black circles). When SspB was present, it was kept at a ratio of 1:1 to ClpX.

#### 4.3.6 Establishment of a calibration curve to measure degradation using FP

To convert the decrease of FP value into amounts of substrates degraded, we first established a calibration curve. As discussed in methods, a commercial peptide corresponding to the last 8 residues of MscS, bearing a NT fluorescein tag, was used to mimic the product of complete degradation. Full length MscS-R128C was labeled using fluorescent-MLM and used as the pre-induction species. The two molecules were mixed at different molar ratios to mimic different stages of degradation, and the FP signal was measured at each ratio. The increase of FP signal was plotted against the percentage of full length protein in the mixture (Figure 4.6). The linearity of the plot is very good. By

comparing the decrease of FP to the actual measurement during the degradation experiment, we can obtain the percentage of full length protein in the sample, and thus calculate the concentration of proteins degraded.

To quantitatively describe the degradation experiment, R128C is not the best site for labeling. Degradation of substrate by ClpXP occurs progressively, from the CT to the NT. In the process, the presence of partially degraded state affects the FP value. To eliminate this problem, we created a new mutant, A286C. A286 is the last residue in the structure of MscS, right before the *ssrA* tag. Therefore, it is degraded within 2 to 3 steps of degradations, and is released as a small fluorescein labeled peptide at the very early stage of the process. This construct has been created, and its degradation will continue to be studied in our lab.

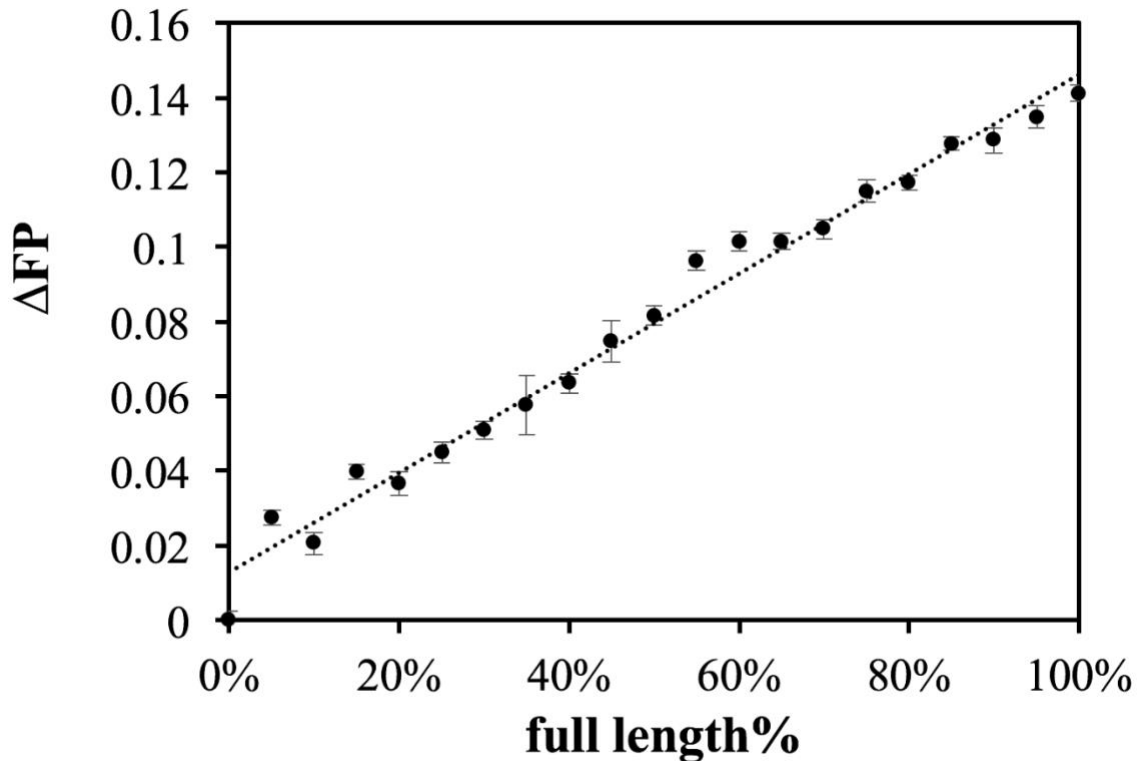


Figure 4.6 Calibration curve to convert change of FP into percent of full length protein left.

#### **4.3.7 S9C degradation-Formation of disulfide bond did not affect degradation *in vivo***

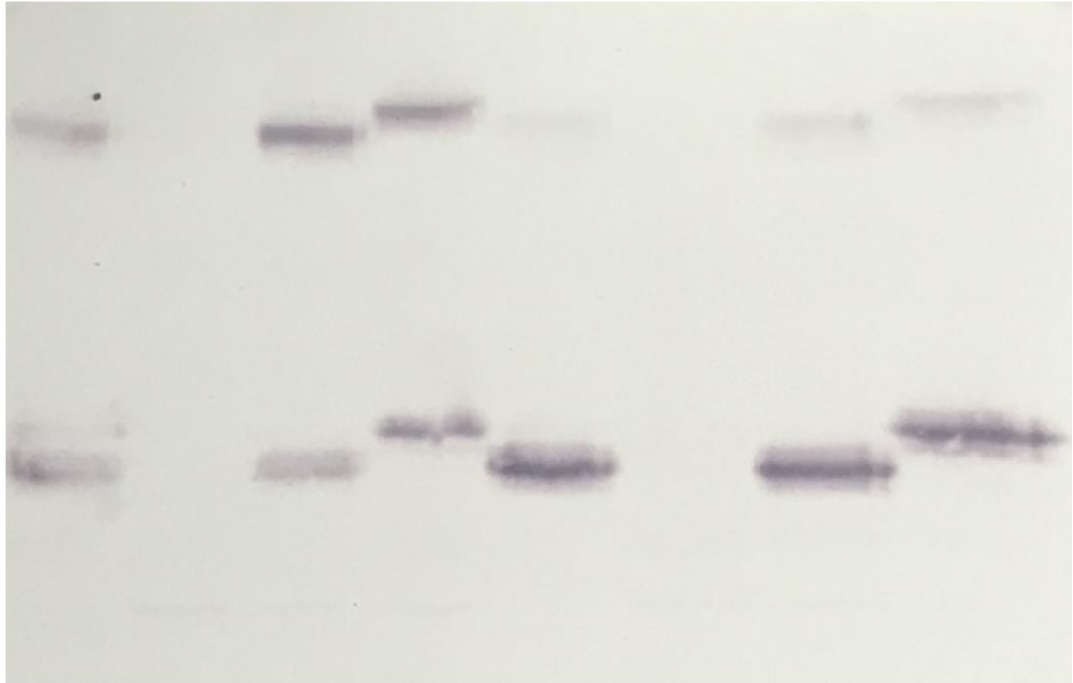
MscS S9 is on the periplasmic side of the membrane<sup>208, 209</sup> and the S9C mutant has been shown to form disulfide bond linked dimers.<sup>210</sup> As shown in Figure 4.7, S9C-ssrA has a clear dimer band. Similar as the wild type MscS, expression of S9C could only be detected in the  $\Delta clpX$  strain, indicating that the introduction of the disulfide bond did not affect the degradation of the ssrA-tagged protein in cells. Without the ssrA tag, both wild type and S9C MscS expressed to similar levels, indicating that the mutation and formation of disulfide bond did not affect the expression of the protein. When samples were pre-incubated with 10 mM DTT, the dimer band disappeared, further confirming that this high molecular weight band came from a disulfide linked MscS dimer.

We examined the degradation of detergent purified MscS-S9C-ssrA. Interestingly, we observed the accumulation of bands corresponding to partially degraded proteins (Figure 4.8). There is one band at a location higher than monomer MscS, with apparent molecular weight around 40 kD, 10 kD larger than that of the MscS monomer. We speculate that this correspond to a full length MscS linked to a fragment via disulfide bond. As expected, treatment with DTT before gel analysis eliminated this band. Next, we repeated the experiment with higher protease to substrate ratio, and observed that the 40kD band disappeared over time. When the samples were analyzed using 25% gels, a small fragment at ~ 20 kD could be observed. This species breaks down into smaller fragments of half of the size upon reduction. And this small fragment did not disappear over time even when we further increase the concentration of the protease complex. We speculate this correspond to the pieces of partially degraded proteins linked by the disulfide bond. We

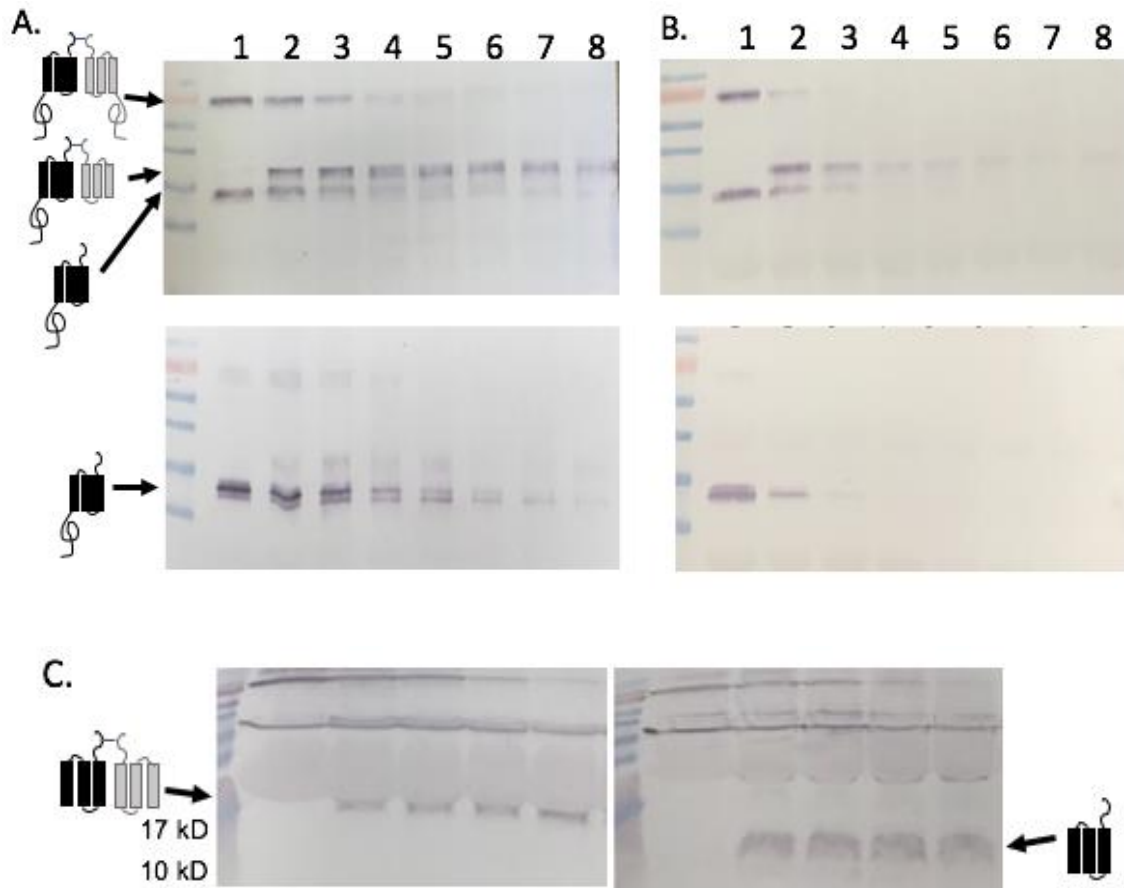
did not observe small fragments during the degradation of wild type MscS. This result indicate that the presence of a disulfide bond prevented complete degradation of the bonded part of the protein. This is different from previous observation that ClpXP degraded disulfied-cross-linked dimers of soluble proteins, which indicates the pore size of protease should be large enough for passage of two or more peptide chains.<sup>207</sup> the ClpXP system can handle disulfide-bonded soluble proteins <sup>207</sup>. We reason that the extra energy barrier required by stripping away detergent from the protein prevented the degradation.

Why we did not see small fragment similarly when we study the degradation in cells? There are two possibilities. First, the formation of the disulfide bond may not be efficient *in vivo*. Therefore, a large portion of the protein might actually exist as monomers. Second, there are other proteases in the cell, such as FtsH. Study have shown that FtsH is capable of degrading meta-stable protein substrate, including membrane proteins. Therefore, the MscS fragment, which does not have a well-folded stable structure, could have been degraded by FtsH and other proteases.

DTT: - - - - + + + +  
 1 2 3 4 1 2 3 4



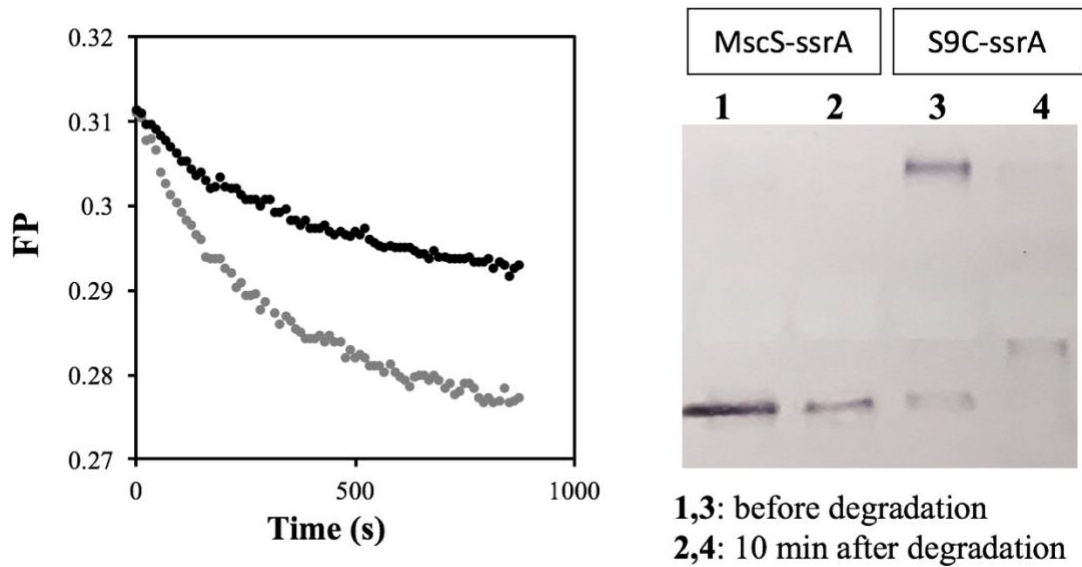
**Figure 4.7** Degradation of MscS containing an inter-subunit disulfide bond. 1. MscS-S9C constructs in DL41(DE3) 2. MscS-S9C-ssrA in DL41(DE3) 3. MscS-S9C in DL41(DE3)ΔClpX 4. MscS-ssrA in DL41(DE3)ΔClpX.



**Figure 4.8** Degradation of MscS-S9C in reconstituted system. **A.** At the low enzyme to substrate ratio of 1:2:3:1. DTT was left out of the PD buffer. Bottom is the same set of samples pretreated with 10 mM DTT to reduce disulfide bond before loading to the gel. **B.** Digestion at high enzyme to substrate ratio of 1:3:6:3. From 1 to 8, samples were taken as 0, 2.5, 5, 10, 15, 20, 30 and 40 min into digestion. **C.** Higher percent gel to reveal an additional small band at approximately 20 kD, which under the reduction condition is reduced into fragments of half of the size. Even at the higher enzyme ratio, this small fragment cannot be further reduced. Molecular weight of bands in markers are 95.72 (red), 55, 43, 34, 26 kD.

#### 4.3.8 Effect of substrate structure on the degradation with ClpXP

Next, I used FP to compare the degradation rate between S9C and wild type MscS. Similar as shown in the western blot result, S9C was degraded slower than the wild type MscS (Figure 4.9). S9C degradation measured in the presence of DTT is similar as that of the wild type MscS, as expected as a single point mutation is not expected to affect the efficiency of digestion (data not shown).



**Figure 4.9** Substrate structure effect on the degradation with ClpXP. FP experiment reveals that MscS-R128C(grey)/S9C-ssrA(black) was degraded slower than MscS-R128C-ssrA. After prolonged incubation, both proteins can be degraded.

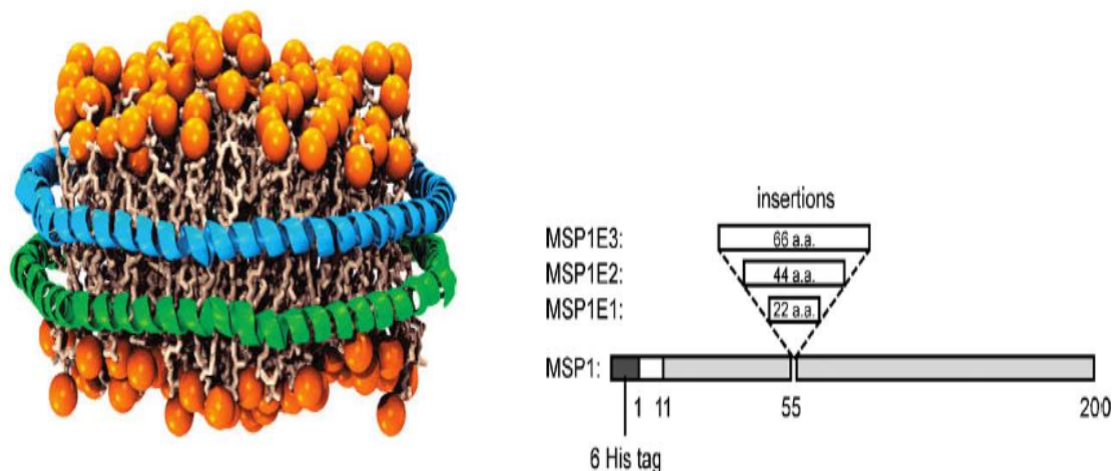
## Chapter V Construction of AcrB nanodisc

### 5.1 Introduction

Membrane scaffold protein (MSP) is an amphipathic  $\alpha$ -helical protein modified from the amino acid sequence of human apolipoprotein A-1 (apo A1)<sup>211</sup>. Apolipoprotein A1 is the major component of high density lipoprotein in plasma. Apo A1 has a 43 residues N-terminal globular domain and a 200 residues C-terminal lipid binding domain<sup>212</sup>. The C-terminal domain of apo A1 consists of repeated  $\alpha$ -helix<sup>213</sup>. MSP is based on apo A1 but without the globular N-terminal domain. A histidine tag is added on the N-terminal of MSP to facilitate the purification step<sup>214</sup>.

MSP tightly wraps around the phospholipid bilayer disc to form a highly homogeneous monodisperse sample<sup>215</sup>. Larger membrane protein complexes require larger phospholipid bilayer, thus extended versions of MSP that contain one, two or three additional 22-mer amphipathic helices inserted in the original MSP1 were created.<sup>211</sup> The extended MSP named MSP1E1, MSP1E2 and MSP1E3 have 22, 44 and 66 amino acid added respectively to the original MSP1 of around 200 amino acid residues<sup>211, 216</sup>, increasing the disk diameter from 9 nm to 20 nm. (Figure 5.1)





**Figure 5.1** Schematic of MSP. A. Structure of nanodisc. Orange circles are the hydrophobic head group of the lipid. Green and blue belts represent two copies of MSP. Adapted with permission from ref <sup>215</sup>. B. The additional amino acids were inserted at position 55 into MSP1. Reprinted (adapted) with permission from ref <sup>216</sup>, Copyright (2007) American Chemical Society.

## 5.2 Material and methods

### 5.2.1 Cloning, expression and purification of MSP1E3

Plasmid encoding the sequence of MSP1E3 was obtained from Addgene, in which the gene of MSP1E3 was cloned into vector pET28. The plasmid was transformed into the expression host *E. coli* BL21 (DE3) for expression.

For protein expression, 5 ml Luria Broth (LB) medium containing kanamycin (50 µg/ml) was inoculated with the indicated bacterium strain and cultured overnight, and then diluted into 300 ml sterilized LB medium containing 50 µg/ml kanamycin the next morning. The culture was incubated with shaking at 37°C until the absorbance at 600 nm reached 0.8,

then induced by the addition of 1 mM IPTG. The temperature was lowered to 28°C after induction and culture was incubated with shaking for 4 hours after induction. Cells are collected by centrifugation at 8000×g for 10 min, the cell can be stored at -80°C overnight.

To purify MSP1E3, cell pellets were re-suspended into 30ml lysis buffer (phosphate buffer pH 7.4, 1 mM PMSF, and 1% Triton X-100). The mixture was sonicated on ice to rupture the cells. The cell lysate was clarified by centrifugation at 10,000 rpm for 20 min. The supernatant solution was loaded onto a Ni-NTA column equilibrated with the same phosphate buffer. The column was first washed with 15 column volumes of wash buffer 1 (40 mM Tris/HCl, 0.3 M NaCl, 1% Triton X-100, pH 8.0), followed by the same volume of wash buffer 2 (40 mM Tris/HCl, 0.3 M NaCl, 50 mM cholate, 10 mM imidazole, pH 8.0), and last washed with same volume of wash buffer 3 (40 mM Tris/HCl, 0.3 M NaCl, 30 mM imidazole, pH 8.0). MSP1E3 was eluted with buffer 4 (40 mM Tris/HCl, 0.3 M NaCl, 0.5 M imidazole).

Purity of MSP1E3 was examined by running SDS-PAGE. Imidazole was removed through dialysis against dialysis buffer (20 mM Tris/HCl, 0.1 M NaCl, 0.5 mM EDTA, pH 7.4), then purified protein was stored at 4°C.

### **5.2.2 Assembly of nanodisc**

Nanodiscs self-assembled when detergents are removed from the mixture. Lipid-detergent mixture was prepared by adding 100mM sodium cholate solution into dried lipid film (1-Palmitoyl-2-Oleoyl-sn-Glycero-3-Phosphocholine, POPC) to final phospholipid concentration of 50 mM. When indicated, fluorescent lipid (1,2-dipalmitoyl-sn-glycero-3-phosphoethanolamine-N-(7-nitro-2-1,3-benzoxadiazol-4-yl), NBD-PE) was added to a

ratio of NBD-PE to POPC of 2:191. MSP1E3 solution was added to cholate-solubilized phospholipid to the ratio of NBD-PE: POPC: MSP1E3 2: 128:1. The final concentration of cholate was adjusted to the range of 12 mM to 40 mM. The mixture was incubated at 4°C for 20 min. To remove cholate, Biobeads SM-2 (Bio-Rad, Hercules, CA) was added to 0.5 - 0.8g per milliliter of the reconstruction mixture. The mixture was incubated with shaking on ice for 4 hours to yield nanodiscs.

### **5.2.3 Assembly of nanodisc containing AcrB trimer**

AcrB was expressed and purified as described previously.<sup>217</sup> MSP1E3 stock and cholate-solubilized lipid were mixed with AcrB to give an AcrB: POPC: MSP1E3 ratio of 1:45:1. Self-assembly of the nanodisc was conducted as described above. AcrB-containing nanodiscs were analyzed using SEC as described followed. The chromatogram and SDS-PAGE analysis of the indicated fractions were shown in Figure 5.4 and 5.5.

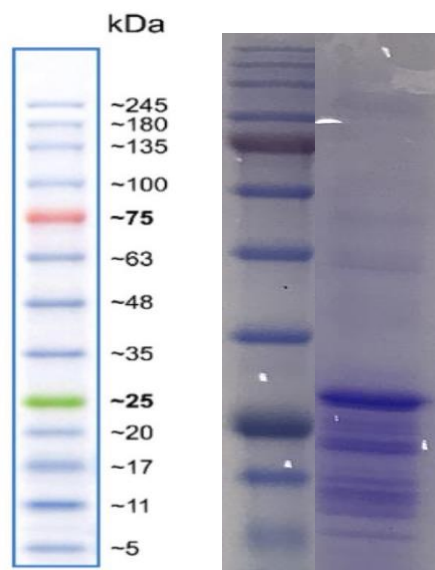
### **5.2.4 Nanodisc detection and identification**

The nanodisc sample was filtered through 0.22 µm filter and analyzed using size exclusion chromatography (Bio-Sil SEC 250 from Bio-rad), at a flow rate of 0.5 ml/min. The dialysis buffer was used as the mobile phase and absorbance of eluent was monitored at 280 nm. 0.5 ml fractions were collected. The elution chromatogram is shown in Figure 5.3. SDS-PAGE was used to examine the presence of MSP1E3, and fluorescence spectrometer was used to monitor the presence of fluorescent lipid when NBD-PE was used. The excitation wavelength of NBD-PE is 460 nm and emission wavelength is 535 nm. The presence of AcrB in the nanodisc was identified by western blot with anti-AcrB antibody.

## 5.3 Result and discussion

### 5.3.1 MSP1E3 expression and purification

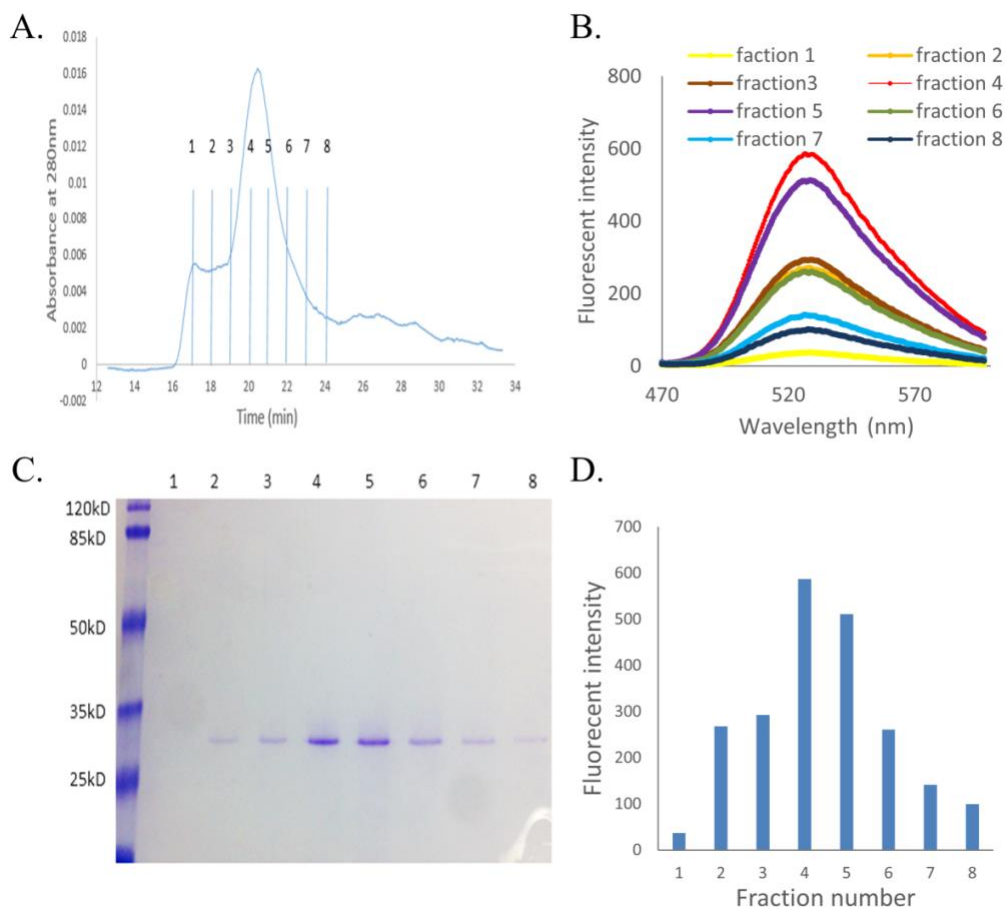
MSP1E3 has the molecular weight of 32.5 kD. The protein was purified using Ni-NTA, then SDS-PAGE was used to check the expression and purity. On the 12% SDS-PAGE (Figure 5.2), a dark band at around 33 kD indicates that MSP1E3 was purified. Though the purity is not high, the contamination bands are not involved in the formation of nanodisc and were later washed away (see Figure 5.3). MSP1E3 can be expressed in sufficient amount, there is an identical band of MSP1E3 at 33 KD even after five fold dilution of the purified sample (Figure 5.2). The BL21(DE3) strain can produce MSP1E3 up to 25mg per liter of culture.



**Figure 5.2** Expression of MSP1E3. With five fold dilution, a clear band at around 32.546 kD, the calculated molecular weight of MSP1E3.

### 5.3.2 Formation of the Nanodiscs with fluorescent lipids

The fluorescent nanodiscs were prepared from a ratio of MSP/POPC/NBD-PE 1:128:2. Nanodisc would only form with fixed number of phospholipids<sup>214, 216, 218</sup>, thus, the ratio between MSP and phospholipid is crucial. Size exclusion chromatography was used to optimize and analyze the formation of the nanodiscs (fluorescent lipid was used as indicated). In Figure 5.3 A, a sharp peak shown around 20 min corresponding to the fluorescent nanodiscs. No peak showed up when the ratio of MSP/phospholipid varies (data not shown). The vertical lines are the time points collected during elution with each fraction about 1 min in duration. The formation of fluorescent nanodisc was confirmed through fluorescence spectroscopy and SDS-PAGE. Figure 5.3 B is the fluorescent spectra of each fraction from A, the peak of fluorescent intensity in each fraction was used to draw the column diagram in Figure 5.3 D. The intensity of the peaks correlated well with the band intensity in Figure 5.3 C, which indicate that the fluorescent nanodiscs were formed since MSP1E3 and the lipid co- eluted at the same time.

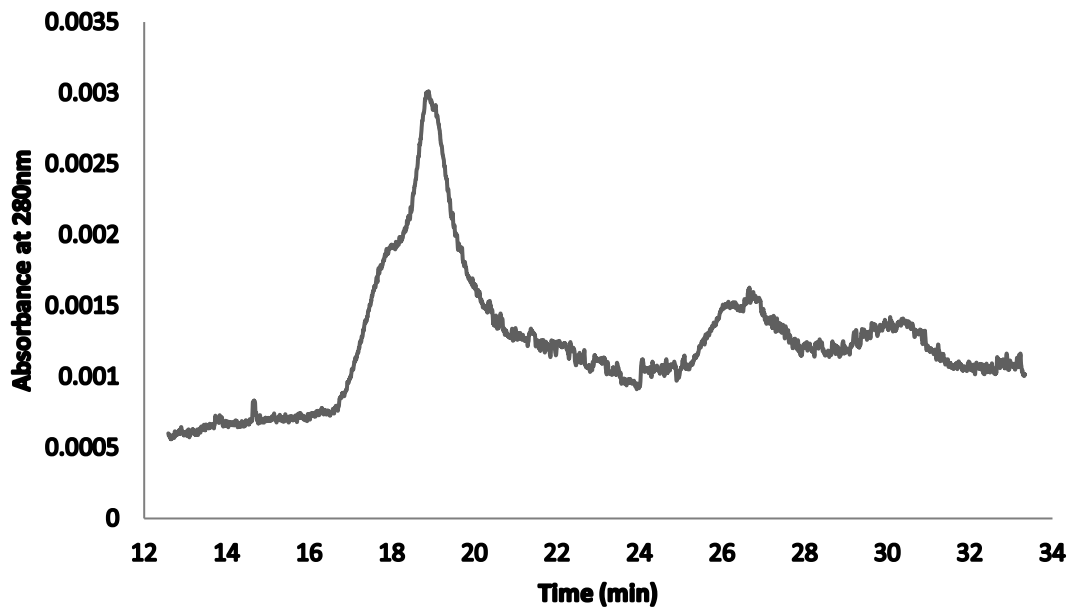


**Figure 5.3** Formation of empty nanodisc with fluorescent lipid. **A.** Size exclusion chromatogram of fluorescent nanodisc. **B.** Fluorescence spectroscopy of fractions as indicated in A. **C.** SDS-PAGE of the fractions as indicated in A. **D.** The peak intensity of the fluorescence spectra of each fraction as shown in B. The peak intensity of MSP in SDS-PAGE and fluorescence emission occurred at the same fractions (4 and 5), confirming the formation of fluorescent nanodisc.

### 5.3.3 Formation of AcrB nanodisc

In order to insert AcrB into the nanodisc, lipids were reduced so that AcrB could self-assemble into the lipid bilayer. The ratio of AcrB: POPC: MSP1E3 ratio is optimized to 1:45:1. The AcrB nanodisc was constructed following the procedure from material and

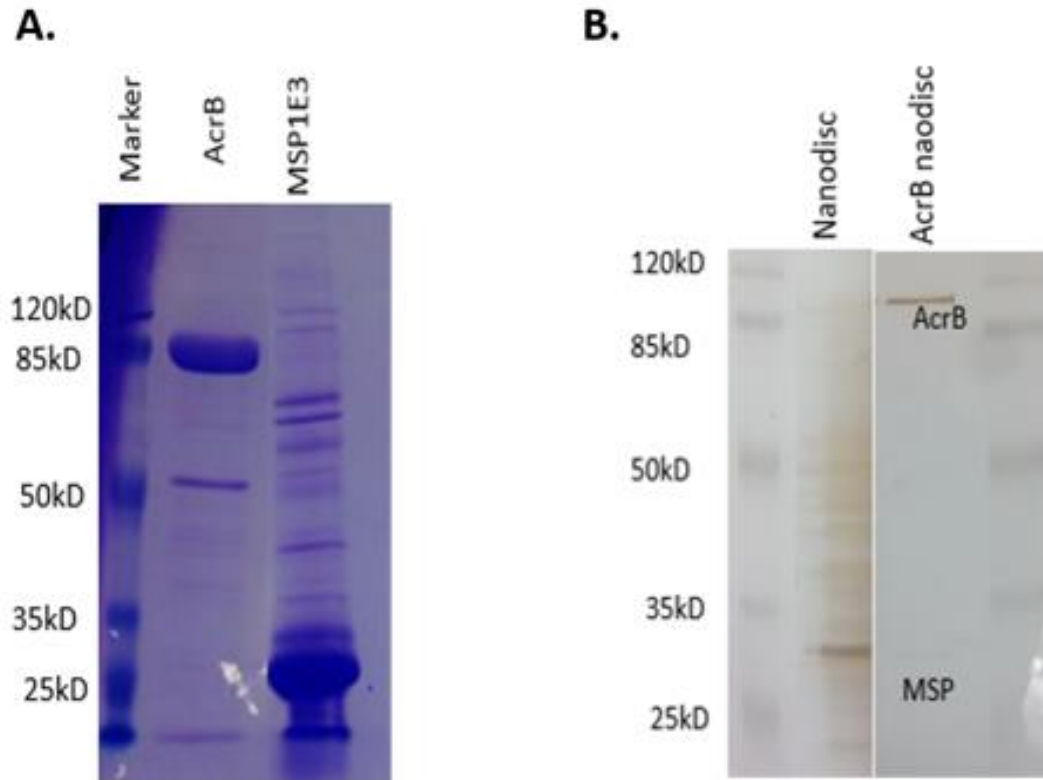
method. There is a major peak at around 19 min elution time (Figure 5.4), which is shorter than empty nanodisc shown at around 20min (Figure 5.3 A). The major peak is the AcrB nanodisc, which is confirmed in the following discussion. The ratio is critical for the formation of membrane protein inserted nanodisc. No distinct peak was observed if the ratio changed (data not shown).



**Figure 5.4** Chromatogram of AcrB-nanodisc. The peak fraction at ~19 min was collected and analysis using SDS-PAGE.

To confirm the formation of AcrB nanodisc, SDS-PAGE with silver stain was used. In Figure 5.5 A, two major bands showed the purified proteins used in the assembly, with AcrB monomer around 100kD, and MSP1E3 around 33kD. Figure 5.5 B is the silver stain of two nanodiscs. On the left, the distinct dark band is MSP, while, on the right, AcrB

nanodisc has two bands, one for AcrB and the other one for MSP. The MSP band is faint in the nanodisc because AcrB is much larger than MSP.



**Figure 5.5** Detection of AcrB-nanodisc. **A.** SDS-PAGE of purified AcrB and MSP1E3. **B.** Silver stain of AcrB-nanodisc peak fraction compared with empty nanodisc peak fraction.



## Chapter VI Conclusion

Protein engineering, using site-directed or random mutagenesis, alters the functions or structures of the target proteins to produce desired novel proteins. In the last 3 decades, protein engineering has been used in a wide range of areas, including industrial, biomedical and research applications.<sup>2</sup> In this thesis, functional tags were constructed and fused onto target proteins. In this way, the functional tag can be used for immobilization of active proteins. Also, protein modified with specific tag can help with the study of membrane protein degradation.

The dual-functional tag, histidine-AHA tag, could facilitate the covalent binding of the target protein through AHA as well as metal affinity purification through histidine. This AHA incorporated dual-functional tag seems less likely to disrupt the protein structure. The pyrophosphatase PpaC and green fluorescent protein sfGFP were successfully tagged with this functional tag without diminishing of protein activities. AHA as a reactive group in the protein could be used for other applications such as cross-linking or labeling. And this tag could be applied for other proteins in general.

The signaling peptides *ssrA*-tag was used to induce the degradation of soluble protein in *E.coli*. It is intriguing to know *ssrA* tagged membrane protein could also be degraded with proteases system in *E.coli*. The *ssrA* tagged membrane protein MscS was found to be degraded *in vivo*. One of the proteases system ClpXP was used for the *in vitro* degradation study of MscS, and dimer structure of MscS was found to lower down the degradation rate with ClpXP system. With the help of protein engineering, MscS was modified with *ssrA*-tag, so that it can be used for detailed, stepwise studying of protein degradation.

## Appendix

### Activity assay of membrane protein DGK

#### A.1 Introduction

Diacylglycerol kinase (DGK) in bacteria is a small membrane protein, which catalyzes the phosphorylation of diacylglycerol using ATP to produce phosphatidic acid and ADP.<sup>219-221</sup>

DGK in Gram-negative bacteria is an essential enzyme, critical for cellular response to environmental stress.<sup>222</sup> As a small membrane protein, DGK is a 121-residue kinase with a molecular weight of 13kD.<sup>221</sup> It is a homo-trimer with each monomer contains three transmembrane helices and an amino-terminal amphiphilic helix.<sup>223</sup>

DGK, being a small homo-trimer membrane protein, can be easily used as a model protein to study protein enzymology,<sup>224-226</sup> membrane protein folding processes,<sup>220, 227</sup> oligomer assembly,<sup>228, 229</sup> and protein stability.<sup>221, 230</sup>

#### A.2 Material and methods

##### A.2.1 DGK expression and purification

Gene of DGK was cloned into vector pET22 with N-terminal histag to facilitate purification. The plasmid was transformed into BL21(DE3) cell for expression. A single colony was used to start an overnight culture, which was used to inoculate 300 ml LB-Ampicillin medium the next morning. Cells were cultured at 37°C until OD 600 nm reached 0.8, and then induced with 1 mM IPTG followed by culturing for additional four hours. Finally, the cells were collected by centrifugation and stored at -80°C.

To purify DGK, the cells pellets were re-suspended in phosphate buffer (50 mM sodium phosphate, 0.3 M NaCl, pH 7.5) containing 1 mM PMSF. The cell suspension was sonicated on ice-water bath for 20 min with 5 s on/off intervals. Cell lysate was clarified by centrifugation at  $15,300 \times g$ ,  $4^{\circ}\text{C}$  for 20 min. Cell debris was collected and re-suspended in phosphate buffer containing 3% OG for 2 hours. The membrane protein suspension was clarified by centrifugation at  $15,300 \times g$  for 30 min at  $4^{\circ}\text{C}$ . The supernatant was mixed with Ni-NTA beads and incubated with beads at  $4^{\circ}\text{C}$  for 1 hour. The beads were loaded into a column and drained, followed by washing with 15 bed volumes of phosphate buffer containing 1.5 % OG and 0.03M imidazole, then with 5 bed volumes of phosphate buffer with 0.5% DM. At last, DGK was eluted with phosphate buffer with 0.5% DM and 0.25 M imidazole. Imidazole was removed from protein solution by dialysis against phosphate buffer with 0.5% DM.

### **A.2.2 DGK activity assay**

#### **Assay mixture preparation**

In order to test DGK activity, reaction buffers have to be prepared as followed precisely. Since DGK activity assay is based on a coupling system, which involves the use of detergent/lipid mixed micelles to solubilize both DGK and its substrate, for example, DOG (1,2-dioctanoyl-sn-glycerol). Thus, steps, as described followed, for making the assay mixture is critical. The assay mix was prepared using 75 mM PIPES with 0.1 mM EDTA, and NaOH was used to adjust the pH to 6.9. Then solid DM (Decyl  $\beta$ -D-maltopyranoside) was added into the mixture to the final concentration of 21 mM (or 1%). Cardiolipin was added to make the final concentration of 3 mol%. Substrate DOG was mixed with the assay

solution described above to the final concentration of 4.5 mol%. All detergents were dissolved by water-bath sonication until the solution was completely clear. Other reaction ingredients were added at last, to the final concentrations of 3 mM ATP, 15 mM Mg Acetate, 1 mM PEP, and 0.2 mM DTT from stocks of each solution with concentration of 50 mM, 1 M, 100 mM, and 100 mM respectively. The assay mix was allocated into small tubes for reactions. A stock NADH solution was prepared by adding NADH into the assay mix to the final concentration of 2.5mM. And this NADH solution was also allocated into small tubes to avoid contamination and multiple freeze-and-thaw process.

### **Activity assay**

The NADH solution was mixed with the assay mix to make the final concentration of NADH to be 0.25mM. Then PK (pyruvate kinase) and LDH (lactate dehydrogenase) was added to final concentrations of 14 and 22 unit respectively. To initiate the reaction, DGK was added into the mixture and absorbance at 340 nm was observed and recorded during the reaction using 60S UV-Visible Spectrophotometer (Thermo Fisher Scientific, Grand Island, NY).

**Table A.1** DGK assay mixture (add each of the following components by order)

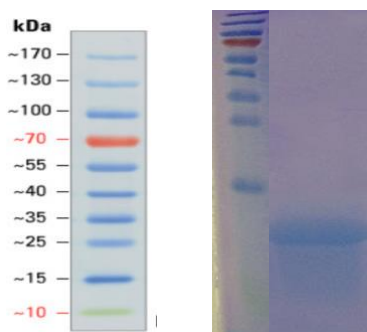
Order	Component	Final concentration	Stock concentration	Mixing guide
1	PIPES	75 mM		Adjust pH to 6.9 using NaOH
1	EDTA	0.1 mM	100 mM	
2	DM	21 mM or (1%)		
3	Cardiolipin	0.66 mM or ( 3mol%)		Dissolve by sonication in water bath
4	DOG	0.95 mM or (4.5mol%)		Dissolve by sonication in ice-water bath
5	ATP	3 mM	50 mM	
6	MgAcetate	15 mM	1 M	
7	PEP	1 mM	100 mM	
8	DTT	0.2 mM	100 mM	Add freshly
Total	Assay mix			Aliquot and store at -80°C

NADH solution	2.5 mM		Add freshly in to assay mix from above and aliquot, store at -80°C, do not re-use.
---------------	--------	--	--

## A.3 Result

### A.3.1 DGK expression

DGK was purified as described in the method. The monomer has a molecular weight of 13kD. There is a significant band in Figure 6.1. (on the right), has the molecular weight around 13kD, indicates that DGK can be expressed and purified at high concentration from the cell. N-terminal his-tag can be used for purification of DGK, while C-terminal his-tag lead to the inhibition of DGK expression. When his-tag was added at the C-terminus, no DGK was detected after purification. (data not shown)



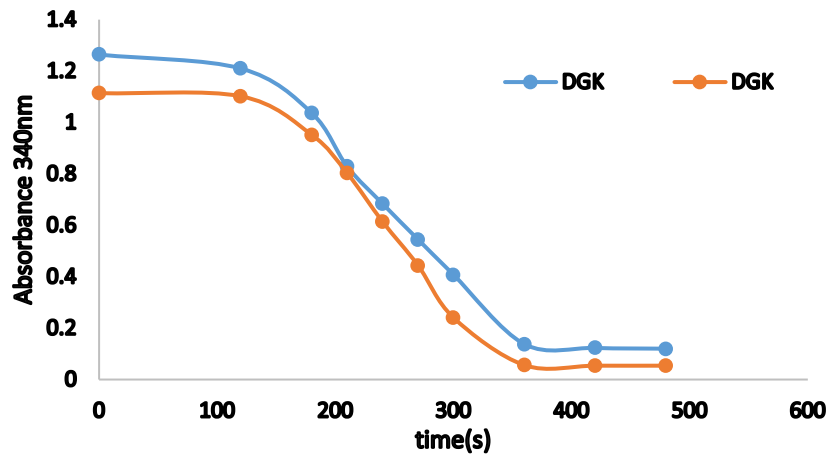
**Figure A.1** Expression of DGK. SDS-PAGE of purified DGK with 20% gel. Lane on the right is elution fraction of DGK, and lane on the left is protein ladder.

### A.3.2 DGK activity

The activity assay was conducted following the procedure described in method. NADH solution was mixed with assay mixture freshly, and enzyme LDH and PK was added and mixed well. Absorbance at 340 nm was monitored. DGK catalyzes the phosphorylation of diacylglycerol (here 1,2-dioctanoyl-sn-glycerol (DOG) was used) using ATP as phosphate source. This process produces phosphatidic acid and ADP. The product ADP then reacts with (phosphoenolpyruvate) PEP, leading to the production of pyruvate using PK. As pyruvate was formed, NADH can react with pyruvate using LDH to produce  $\text{NAD}^+$  and

lactate. For every one ADP produced by DGK, one molecule of NADH is converted to NAD<sup>+</sup>. NADH has a strong absorbance at 340 nm, but NAD<sup>+</sup> does not. Thus, the activity of DGK can be indirectly measured by monitoring the decreasing in absorbance at 340 nm. (Figure 6.2)

In the reaction, DGK was used as the final concentration of 0.09 mg/ml. The decreasing in absorbance was recorded.



**Figure A.2** Activity of DGK monitored by the consumption of NADH. Data from two repeats are shown.

The activity was calculated using the formula shown below:

$$\frac{\frac{\Delta \text{Absorbance}}{\text{time(s)}} * 60 \text{s/min}}{\frac{6110 \text{ au}}{\text{cm} \cdot \text{M}}} * \text{volume of reaction} \Bigg/ \text{Quantity of DGK (mg)}$$

Through the calculation, the rate of my DGK is around 55 unit/mg.

## References

1. Ulmer, K. M. (1983) Protein engineering, *Science* 219, 666-671.
2. Chica, R. A. (2015) Protein engineering in the 21st century, *Protein Science* 24, 431-433.
3. Sanger, F., and Tuppy, H. (1951) The amino-acid sequence in the phenylalanyl chain of insulin. 1. The identification of lower peptides from partial hydrolysates, *Biochemical Journal* 49, 463.
4. Kendrew, J., Dickerson, R., Strandberg, B., Hart, R., Davies, D., Phillips, D., and Shore, V. (1960) Structure of myoglobin: A three-dimensional Fourier synthesis at 2 Å. resolution, *Nature* 185, 422-427.
5. Meyer, E. E. (1997) The first years of the Protein Data Bank, *Protein Science* 6, 1591-1597.
6. Wagner, G., and Wüthrich, K. (1982) Sequential resonance assignments in protein <sup>1</sup>H nuclear magnetic resonance spectra: basic pancreatic trypsin inhibitor, *Journal of molecular biology* 155, 347-366.
7. Turanli-Yildiz, B., Alkim, C., and Cakar, Z. P. (2012) Protein engineering methods and applications, In *Protein Engineering*, InTech.
8. Altamirano, M. M., Blackburn, J. M., Aguayo, C., and Fersht, A. R. (2000) Directed evolution of new catalytic activity using the alpha/beta-barrel scaffold, *Nature* 403, 617.
9. Olempska-Bier, Z. S., Merker, R. I., Ditto, M. D., and DiNovi, M. J. (2006) Food-processing enzymes from recombinant microorganisms—a review, *Regulatory toxicology and Pharmacology* 45, 144-158.



10. Kuraishi, C., Yamazaki, K., and Susa, Y. (2001) Transglutaminase: its utilization in the food industry, *Food Reviews International* 17, 221-246.
11. Torres, E., Bustos-Jaimes, I., and Le Borgne, S. (2003) Potential use of oxidative enzymes for the detoxification of organic pollutants, *Applied Catalysis B: Environmental* 46, 1-15.
12. Le Borgne, S., and Quintero, R. (2003) Biotechnological processes for the refining of petroleum, *Fuel Processing Technology* 81, 155-169.
13. Vazquez, E., Ferrer-Miralles, N., Mangués, R., Corchero, J. L., Schwartz, J., and Villaverde, A. (2009) Modular protein engineering in emerging cancer therapies, *Current pharmaceutical design* 15, 893-916.
14. Bonin-Debs, A. L., Boche, I., Gille, H., and Brinkmann, U. (2004) Development of secreted proteins as biotherapeutic agents, *Expert opinion on biological therapy* 4, 551-558.
15. Pâques, F., and Duchateau, P. (2007) Meganucleases and DNA double-strand break-induced recombination: perspectives for gene therapy, *Current gene therapy* 7, 49-66.
16. Olafsen, T., and Wu, A. M. (2010) Antibody vectors for imaging, In *Seminars in nuclear medicine*, pp 167-181, Elsevier.
17. Rehm, B. H. (2006) Genetics and biochemistry of polyhydroxyalkanoate granule self-assembly: the key role of polyester synthases, *Biotechnology letters* 28, 207-213.
18. Antikainen, N. M., and Martin, S. F. (2005) Altering protein specificity: techniques and applications, *Bioorganic & medicinal chemistry* 13, 2701-2716.

19. Zheng, L., Baumann, U., and Reymond, J.-L. (2004) An efficient one-step site-directed and site-saturation mutagenesis protocol, *Nucleic acids research* 32, e115-e115.
20. Cirino, P. C., Mayer, K. M., and Umeno, D. (2003) Generating mutant libraries using error-prone PCR, *Directed Evolution Library Creation: Methods and Protocols*, 3-9.
21. Shimizu, Y., Inoue, A., Tomari, Y., Suzuki, T., Yokogawa, T., Nishikawa, K., and Ueda, T. (2001) Cell-free translation reconstituted with purified components, *Nature biotechnology* 19, 751.
22. Shimizu, Y., Kuruma, Y., Ying, B. W., Umekage, S., and Ueda, T. (2006) Cell-free translation systems for protein engineering, *The FEBS journal* 273, 4133-4140.
23. Chalfie, M. (1995) Green fluorescent protein, *Photochemistry and photobiology* 62, 651-656.
24. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W., and Prasher, D. C. (1994) Green fluorescent protein as a marker for gene expression, *Science*, 802-805.
25. Kanda, T., Sullivan, K. F., and Wahl, G. M. (1998) Histone–GFP fusion protein enables sensitive analysis of chromosome dynamics in living mammalian cells, *Current Biology* 8, 377-385.
26. Kostrzynska, M., Leung, K. T., Lee, H., and Trevors, J. T. (2002) Green fluorescent protein-based biosensor for detecting SOS-inducing activity of genotoxic compounds, *Journal of microbiological methods* 48, 43-51.

27. Straight, A. F., Belmont, A. S., Robinett, C. C., and Murray, A. W. (1996) GFP tagging of budding yeast chromosomes reveals that protein–protein interactions can mediate sister chromatid cohesion, *Current Biology* 6, 1599-1608.
28. Truong, K., and Ikura, M. (2001) The use of FRET imaging microscopy to detect protein–protein interactions and protein conformational changes in vivo, *Current opinion in structural biology* 11, 573-578.
29. Jackson, S. E., Craggs, T. D., and Huang, J.-r. (2006) Understanding the folding of GFP using biophysical techniques, *Expert review of proteomics* 3, 545-559.
30. Terpe, K. (2003) Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems, *Applied microbiology and biotechnology* 60, 523-533.
31. Hedhammar, M., Gräslund, T., and Hober, S. (2005) Protein engineering strategies for selective protein purification, *Chemical engineering & technology* 28, 1315-1325.
32. Chong, S., Mersha, F. B., Comb, D. G., Scott, M. E., Landry, D., Vence, L. M., Perler, F. B., Benner, J., Kucera, R. B., and Hirvonen, C. A. (1997) Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element, *Gene* 192, 271-281.
33. Arnau, J., Lauritzen, C., Petersen, G. E., and Pedersen, J. (2006) Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins, *Protein expression and purification* 48, 1-13.
34. Sassenfeld, H. M., and Brewer, S. J. (1984) A polypeptide fusion designed for the purification of recombinant proteins, *Nature Biotechnology* 2, 76-81.

35. Fuchs, S. M., and Raines, R. T. (2005) Polyarginine as a multifunctional fusion tag, *Protein science* 14, 1538-1544.
36. Nock, S., Spudich, J. A., and Wagner, P. (1997) Reversible, site-specific immobilization of polyarginine-tagged fusion proteins on mica surfaces, *FEBS letters* 414, 233-238.
37. Jeong, J., Ha, T. H., and Chung, B. H. (2006) Enhanced reusability of hexa-arginine-tagged esterase immobilized on gold-coated magnetic nanoparticles, *Analytica Chimica Acta* 569, 203-209.
38. Porath, J., Carlsson, J., Olsson, I., and Belfrage, G. (1975) Metal chelate affinity chromatography, a new approach to protein fractionation, *Nature* 258, 598-599.
39. Hochuli, E., Döbeli, H., and Schacher, A. (1987) New metal chelate adsorbent selective for proteins and peptides containing neighbouring histidine residues, *Journal of Chromatography A* 411, 177-184.
40. Hochuli, E., Bannwarth, W., Döbeli, H., Gentz, R., and Stüber, D. (1988) Genetic approach to facilitate purification of recombinant proteins with a novel metal chelate adsorbent, *Nature biotechnology* 6, 1321-1325.
41. Westra, D. F., Welling, G. W., Koedijk, D. G., Scheffer, A. J., The, T. H., and Welling-Wester, S. (2001) Immobilised metal-ion affinity chromatography purification of histidine-tagged recombinant proteins: a wash step with a low concentration of EDTA, *Journal of Chromatography B: Biomedical Sciences and Applications* 760, 129-136.

42. Hefti, M. H., Dixon, R., and Vervoort, J. (2001) A novel purification method for histidine-tagged proteins containing a thrombin cleavage site, *Analytical biochemistry* 295, 180-185.
43. Wu, J., and FUtowiczK, M. (1999) Hexahistidme (Hise> tag dependent protein dimerization: A cautionary tale, *Acta Biochim. Pol.* 46, 591-599.
44. Halliwell, C. M., Morgan, G., Ou, C.-P., and Cass, A. E. (2001) Introduction of a (poly) histidine tag in L-lactate dehydrogenase produces a mixture of active and inactive molecules, *Analytical biochemistry* 295, 257-261.
45. Schmidt, T. G., and Skerra, A. (1993) The random peptide library-assisted engineering of a C-terminal affinity peptide, useful for the detection and purification of a functional Ig Fv fragment, *Protein Engineering, Design and Selection* 6, 109-122.
46. Laitinen, O., Hytönen, V., Nordlund, H., and Kulomaa, M. (2006) Genetically engineered avidins and streptavidins, *Cellular and Molecular Life Sciences CMLS* 63, 2992-3017.
47. Schmidt, T. G., and Skerra, A. (2007) The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins, *Nature protocols* 2, 1528.
48. Skerra, A., and Schmidt, T. G. (1999) Applications of a peptide ligand for streptavidin: the Strep-tag, *Biomolecular engineering* 16, 79-86.
49. Ostermeier, C., Harrenga, A., Ermler, U., and Michel, H. (1997) Structure at 2.7 Å resolution of the *Paracoccus denitrificans* two-subunit cytochrome c oxidase

- complexed with an antibody FV fragment, *Proceedings of the National Academy of Sciences* 94, 10547-10553.
50. Hopp, T. P., Prickett, K. S., Price, V. L., Libby, R. T., March, C. J., Cerretti, D. P., Urdal, D. L., and Conlon, P. J. (1988) A short polypeptide marker sequence useful for recombinant protein identification and purification, *Biotechnology* 6, 1204-1210.
  51. Prickett, K. S., Amberg, D., and Hopp, T. (1989) A calcium-dependent antibody for identification and purification of recombinant proteins, *Biotechniques* 7, 580-589.
  52. Knappik, A., and Plückthun, A. (1994) An improved affinity tag based on the FLAG peptide for the detection and purification of recombinant antibody fragments, *Biotechniques* 17, 754-761.
  53. Brizzard, B. L., Chubet, R., and Vizard, D. (1994) Immunoaffinity purification of FLAG epitope-tagged bacterial alkaline phosphatase using a novel monoclonal antibody and peptide elution, *Biotechniques* 16, 730-735.
  54. Schuster, M., Wasserbauer, E., Einhauer, A., Ortner, C., Jungbauer, A., Hammerschmid, F., and Werner, G. (2000) Protein expression strategies for identification of novel target proteins, *Journal of biomolecular screening* 5, 89-97.
  55. Bucher, M. H., Evdokimov, A. G., and Waugh, D. S. (2002) Differential effects of short affinity tags on the crystallization of *Pyrococcus furiosus* maltodextrin-binding protein, *Acta Crystallographica Section D: Biological Crystallography* 58, 392-397.

56. Evan, G. I., Lewis, G. K., Ramsay, G., and Bishop, J. M. (1985) Isolation of monoclonal antibodies specific for human c-myc proto-oncogene product, *Molecular and cellular biology* 5, 3610-3616.
57. Kipriyanov, S. M., Kupriyanova, O. A., Little, M., and Moldenhauer, G. (1996) Rapid detection of recombinant antibody fragments directed against cell-surface antigens by flow cytometry, *Journal of immunological methods* 196, 51-62.
58. Mckern, N. M., Frenkel, M. J., Verkuylen, A., Bentley, J. D., Lovrecz, G. O., Ivancic, N., Elleman, T. C., Cosgrove, L. J., Lou, M., and Garrett, T. P. (1997) Crystallization of the first three domains of the human insulin-like growth factor-1 receptor, *Protein science* 6, 2663-2666.
59. Karpeisky, M. Y., Senchenko, V., Dianova, M., and Kanevsky, V. Y. (1994) Formation and properties of S-protein complex with S-peptide-containing fusion protein, *FEBS letters* 339, 209-212.
60. Connelly, P. R., Varadarajan, R., Sturtevant, J. M., and Richards, F. M. (1990) Thermodynamics of protein-peptide interactions in the ribonuclease S system studied by titration calorimetry, *Biochemistry* 29, 6108-6114.
61. Lellouch, A. C., and Geremia, R. A. (1999) Expression and Study of Recombinant ExoM, a  $\beta$ 1-4 Glucosyltransferase Involved in Succinoglycan Biosynthesis in *Sinorhizobium meliloti*, *Journal of bacteriology* 181, 1141-1148.
62. Kelemen, B. R., Klink, T. A., Behike, M. A., Eubanks, S. R., Leland, P. A., and Raines, R. T. (1999) Hypersensitive substrate for ribonucleases, *Nucleic acids research* 27, 3696-3701.

63. Strange, R. C., Spiteri, M. A., Ramachandran, S., and Fryer, A. A. (2001) Glutathione-S-transferase family of enzymes, *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* 482, 21-26.
64. Yip, Y. L., Smith, G., and Ward, R. L. (2001) Comparison of phage pIII, pVIII and GST as carrier proteins for peptide immunisation in Balb/c mice, *Immunology letters* 79, 197-202.
65. McTigue, M. A., Williams, D. R., and Tainer, J. A. (1995) Crystal Structures of a Schistosomal Drug and Vaccine Target: Glutathione S-Transferase from *Schistosoma japonica* and its Complex with the Leading Antischistosomal Drug Praziquantel, *Journal of molecular biology* 246, 21-27.
66. Smith, D. B., and Johnson, K. S. (1988) Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase, *Gene* 67, 31-40.
67. Sassenfeld, H. M. (1990) Engineering proteins for purification, *Trends in biotechnology* 8, 88-93.
68. Phizicky, E., Bastiaens, P. I., Zhu, H., Snyder, M., and Fields, S. (2003) Protein analysis on a proteomic scale, *Nature* 422, 208.
69. Vikis, H. G., and Guan, K.-L. (2004) Glutathione-S-transferase-fusion based assays for studying protein-protein interactions, *Protein-protein interactions: methods and applications*, 175-186.
70. Saaem, I., Papasotiropoulos, V., Wang, T., Soteropoulos, P., and Libera, M. (2007) Hydrogel-based protein nanoarrays, *Journal of nanoscience and nanotechnology* 7, 2623-2632.



71. Tomme, P., Warren, R. A. J., Miller Jr, R. C., Kilburn, D. G., and Gilkes, N. R. (1995) Cellulose-binding domains: classification and properties, ACS Publications.
72. Sugimoto, N., Igarashi, K., and Samejima, M. (2012) Cellulose affinity purification of fusion proteins tagged with fungal family 1 cellulose-binding domain, *Protein expression and purification* 82, 290-296.
73. Tomme, P., Boraston, A., McLean, B., Kormos, J., Creagh, A. L., Sturch, K., Gilkes, N. R., Haynes, C. A., Warren, R. A. J., and Kilburn, D. G. (1998) Characterization and affinity applications of cellulose-binding domains, *Journal of Chromatography B: Biomedical Sciences and Applications* 715, 283-296.
74. Xu, Z., Bae, W., Mulchandani, A., Mehra, R. K., and Chen, W. (2002) Heavy metal removal by novel CBD-EC20 sorbents immobilized on cellulose, *Biomacromolecules* 3, 462-465.
75. Stofko-Hahn, R. E., Carr, D. W., and Scott, J. D. (1992) A single step purification for recombinant proteins Characterization of a microtubule associated protein (MAP 2) fragment which associates with the type II cAMP-dependent protein kinase, *FEBS letters* 302, 274-278.
76. Blumenthal, D. K., Takio, K., Edelman, A. M., Charbonneau, H., Titani, K., Walsh, K. A., and Krebs, E. G. (1985) Identification of the calmodulin-binding domain of skeletal muscle myosin light chain kinase, *Proceedings of the National Academy of Sciences* 82, 3187-3191.
77. Head, J. F. (1992) A better grip on calmodulin, *Current Biology* 2, 609-611.

78. Zheng, C.-F., Simcox, T., Xu, L., and Vaillancourt, P. (1997) A new expression vector for high level protein production, one step purification and direct isotopic labeling of calmodulin-binding peptide fusion proteins, *Gene* 186, 55-60.
79. Duplay, P., and Hofnung, M. (1988) Two regions of mature periplasmic maltose-binding protein of *Escherichia coli* involved in secretion, *Journal of bacteriology* 170, 4445-4450.
80. Duplay, P., Bedouelle, H., Fowler, A., Zabin, I., Saurin, W., and Hofnung, M. (1984) Sequences of the malE gene and of its product, the maltose-binding protein of *Escherichia coli* K12, *Journal of Biological Chemistry* 259, 10606-10613.
81. Ford, C. F., Suominen, I., and Glatz, C. E. (1991) Fusion tails for the recovery and purification of recombinant proteins, *Protein expression and purification* 2, 95-107.
82. Kapust, R. B., and Waugh, D. S. (1999) *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused, *Protein Science* 8, 1668-1674.
83. Cronan, J. E. (1990) Biotination of proteins in vivo. A post-translational modification to label, purify, and study proteins, *Journal of Biological Chemistry* 265, 10327-10333.
84. Nilsson, B., Abrahmsen, L., and Uhlen, M. (1985) Immobilization and purification of enzymes with staphylococcal protein A gene fusion vectors, *The EMBO journal* 4, 1075.
85. Schatz, P. J. (1993) Use of peptide libraries to map the substrate specificity of a peptide-modifying enzyme: a 13 residue consensus peptide specifies biotinylation in *Escherichia coli*, *Nature Biotechnology* 11, 1138-1143.

86. Tucker, J., and Grisshammer, R. (1996) Purification of a rat neurotensin receptor expressed in *Escherichia coli*, *Biochemical Journal* 317, 891-899.
87. Studier, F. W., and Moffatt, B. A. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes, *Journal of molecular biology* 189, 113-130.
88. Dulebohn, D., Choy, J., Sundermeier, T., Okan, N., and Karzai, A. W. (2007) Trans-translation: the tmRNA-mediated surveillance mechanism for ribosome rescue, directed protein degradation, and nonstop mRNA decay, *Biochemistry* 46, 4681-4693.
89. Wiegert, T., and Schumann, W. (2001) SsrA-mediated tagging in *Bacillus subtilis*, *Journal of Bacteriology* 183, 3885-3889.
90. Muto, A., Fujihara, A., Ito, K. i., Matsuno, J., Ushida, C., and Himeno, H. (2000) Requirement of transfer-messenger RNA for the growth of *Bacillus subtilis* under stresses, *Genes to Cells* 5, 627-635.
91. Withey, J. H., and Friedman, D. I. (2002) The biological roles of trans-translation, *Current opinion in microbiology* 5, 154-159.
92. Williams, K. P., and Bartel, D. P. (1996) Phylogenetic analysis of tmRNA secondary structure, *Rna* 2, 1306-1310.
93. Felden, B., Himeno, H., Muto, A., Atkins, J., and Gesteland, R. (1996) Structural organization of *Escherichia coli* tmRNA, *Biochimie* 78, 979-983.
94. KoMINE, Y., Kitabatake, M., Yokogawa, T., Nishikawa, K., and Inokuchi, H. (1994) A tRNA-like structure is present in 10Sa RNA, a small stable RNA from *Escherichia coli*, *Proceedings of the National Academy of Sciences* 91, 9223-9227.

95. Nameki, N., Tadaki, T., Muto, A., and Himeno, H. (1999) Amino acid acceptor identity switch of *Escherichia coli* tmRNA from alanine to histidine in vitro, *Journal of molecular biology* 289, 1-7.
96. Tu, G.-F., Reid, G. E., Zhang, J.-G., Moritz, R. L., and Simpson, R. J. (1995) C-terminal extension of truncated recombinant proteins in *Escherichia coli* with a 10Sa RNA decapeptide, *Journal of Biological Chemistry* 270, 9322-9326.
97. Keiler, K. C., Waller, P. R., and Sauer, R. T. (1996) Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA, *Science-AAAS-Weekly Paper Edition* 271, 990-993.
98. Gottesman, S., Roche, E., Zhou, Y., and Sauer, R. T. (1998) The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the SsrA-tagging system, *Genes & development* 12, 1338-1347.
99. Herman, C., Thévenet, D., Bouloc, P., Walker, G. C., and D'Ari, R. (1998) Degradation of carboxy-terminal-tagged cytoplasmic proteins by the *Escherichia coli* protease HflB (FtsH), *Genes & development* 12, 1348-1355.
100. Silber, K. R., Keiler, K. C., and Sauer, R. T. (1992) Tsp: a tail-specific protease that selectively degrades proteins with nonpolar C termini, *Proceedings of the National Academy of Sciences* 89, 295-299.
101. Keiler, K. C., Silber, K. R., Sauer, R. T., Downard, K. M., Papayannopoulos, I. A., and Biemann, K. (1995) C-terminal specific protein degradation: activity and substrate specificity of the Tsp protease, *Protein Science* 4, 1507-1515.

102. Silber, K. R., and Sauer, R. T. (1994) Deletion of the *prc* (*tsp*) gene provides evidence for additional tail-specific proteolytic activity in *Escherichia coli* K-12, *Molecular and General Genetics MGG* 242, 237-240.
103. Tomoyasu, T., Yamanaka, K., Murata, K., Suzaki, T., Bouloc, P., Kato, A., Niki, H., Hiraga, S., and Ogura, T. (1993) Topology and subcellular localization of FtsH protein in *Escherichia coli*, *Journal of bacteriology* 175, 1352-1357.
104. Tomoyasu, T., Yuki, T., Morimura, S., Mori, H., Yamanaka, K., Niki, H., Hiraga, S., and Ogura, T. (1993) The *Escherichia coli* FtsH protein is a prokaryotic member of a protein family of putative ATPases involved in membrane functions, cell cycle control, and gene expression, *Journal of Bacteriology* 175, 1344-1351.
105. Tomoyasu, T., Gamer, J., Bukau, B., Kanemori, M., Mori, H., Rutman, A. J., Oppenheim, A. B., Yura, T., Yamanaka, K., and Niki, H. (1995) *Escherichia coli* FtsH is a membrane-bound, ATP-dependent protease which degrades the heat-shock transcription factor sigma 32, *The EMBO Journal* 14, 2551.
106. Akiyama, Y., Kihara, A., and Ito, K. (1996) Subunit a of proton ATPase F<sub>0</sub> sector is a substrate of the FtsH protease in *Escherichia coli*, *FEBS letters* 399, 26-28.
107. Herman, C., Prakash, S., Lu, C. Z., Matouschek, A., and Gross, C. A. (2003) Lack of a robust unfoldase activity confers a unique level of substrate specificity to the universal AAA protease FtsH, *Molecular cell* 11, 659-669.
108. Kihara, A., Akiyama, Y., and Ito, K. (1995) FtsH is required for proteolytic elimination of uncomplexed forms of SecY, an essential protein translocase subunit, *Proceedings of the National Academy of Sciences* 92, 4532-4536.

109. Ito, K., and Akiyama, Y. (2005) Cellular functions, mechanism of action, and regulation of FtsH protease, *Annu. Rev. Microbiol.* 59, 211-231.
110. Hari, S. B., and Sauer, R. T. (2016) The AAA+ FtsH protease degrades an ssrA-tagged model protein in the inner membrane of Escherichia coli, *Biochemistry* 55, 5649-5652.
111. Wickner, S., Maurizi, M. R., and Gottesman, S. (1999) Posttranslational quality control: folding, refolding, and degrading proteins, *Science* 286, 1888-1893.
112. Porankiewicz, J., Wang, J., and Clarke, A. K. (1999) New insights into the ATP-dependent Clp protease: Escherichia coli and beyond, *Molecular microbiology* 32, 449-458.
113. Maurizi, M. R., Clark, W. P., Katayama, Y., Rudikoff, S., Pumphrey, J., Bowers, B., and Gottesman, S. (1990) Sequence and structure of Clp P, the proteolytic component of the ATP-dependent Clp protease of Escherichia coli, *Journal of Biological Chemistry* 265, 12536-12545.
114. Kirstein, J., Hoffmann, A., Lilie, H., Schmidt, R., Rübsamen-Waigmann, H., Brötz-Oesterhelt, H., Mogk, A., and Turgay, K. (2009) The antibiotic ADEP reprogrammes ClpP, switching it from a regulated to an uncontrolled protease, *EMBO molecular medicine* 1, 37-49.
115. Gottesman, S., Wickner, S., Maurizi, M. R., Beals, C. R., Clipstone, N. A., Ho, S. N., Crabtree, G. R., Zamir, I., Zhang, J., and Lazar, M. A. (1997) Protein quality control: triage by chaperones and proteases 815, *Genes Dev* 11, 815-823.
116. Wang, J., Hartling, J. A., and Flanagan, J. M. (1997) The structure of ClpP at 2.3 Å resolution suggests a model for ATP-dependent proteolysis, *Cell* 91, 447-456.

117. Hanson, P. I., and Whiteheart, S. W. (2005) AAA+ proteins: have engine, will work, *Nature reviews. Molecular cell biology* 6, 519.
118. Wojtyra, U. A., Thibault, G., Tuite, A., and Houry, W. A. (2003) The N-terminal zinc binding domain of ClpX is a dimerization domain that modulates the chaperone function, *Journal of biological chemistry* 278, 48981-48990.
119. Singh, S. K., Rozycki, J., Ortega, J., Ishikawa, T., Lo, J., Steven, A. C., and Maurizi, M. R. (2001) Functional domains of the ClpA and ClpX molecular chaperones identified by limited proteolysis and deletion analysis, *Journal of biological chemistry* 276, 29420-29429.
120. Martin, A., Baker, T. A., and Sauer, R. T. (2007) Distinct static and dynamic interactions control ATPase-peptidase communication in a AAA+ protease, *Molecular cell* 27, 41-52.
121. Baker, T. A., and Sauer, R. T. (2012) ClpXP, an ATP-powered unfolding and protein-degradation machine, *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research* 1823, 15-28.
122. Levchenko, I., Seidel, M., Sauer, R. T., and Baker, T. A. (2000) A specificity-enhancing factor for the ClpXP degradation machine, *Science* 289, 2354-2356.
123. Wah, D. A., Levchenko, I., Baker, T. A., and Sauer, R. T. (2002) Characterization of a specificity factor for an AAA+ ATPase: assembly of SspB dimers with ssrA-tagged proteins and the ClpX hexamer, *Chemistry & biology* 9, 1237-1245.
124. Park, E. Y., Lee, B.-G., Hong, S.-B., Kim, H.-W., Jeon, H., and Song, H. K. (2007) Structural basis of SspB-tail recognition by the zinc binding domain of ClpX, *Journal of molecular biology* 367, 514-526.

125. Wah, D. A., Levchenko, I., Rieckhof, G. E., Bolon, D. N., Baker, T. A., and Sauer, R. T. (2003) Flexible linkers leash the substrate binding domain of SspB to a peptide module that stabilizes delivery complexes with the AAA+ ClpXP protease, *Molecular cell* 12, 355-363.
126. McGinness, K. E., Baker, T. A., and Sauer, R. T. (2006) Engineering controllable protein degradation, *Molecular cell* 22, 701-707.
127. Sauer, R. T., Bolon, D. N., Burton, B. M., Burton, R. E., Flynn, J. M., Grant, R. A., Hersch, G. L., Joshi, S. A., Kenniston, J. A., and Levchenko, I. (2004) Sculpting the proteome with AAA+ proteases and disassembly machines, *Cell* 119, 9-18.
128. Zabel, U., Häusler, C., Weeger, M., and Schmidt, H. H. (1999) Homodimerization of Soluble Guanylyl Cyclase Subunits DIMERIZATION ANALYSIS USING A GLUTATHIONES-TRANSFERASE AFFINITY TAG, *J. Biol. Chem.* 274, 18149-18152.
129. Schmidt, T. G., and Skerra, A. (1993) The random peptide library-assisted engineering of a C-terminal affinity peptide, useful for the detection and purification of a functional Ig Fv fragment, *Protein Eng.* 6, 109-122.
130. Skerra, A., and Schmidt, T. G. (2000) [18] Use of the Strep-tag and streptavidin for detection and purification of recombinant proteins, *Methods Enzymol.* 326, 271-304.
131. Green, N., Alexander, H., Olson, A., Alexander, S., Shinnick, T. M., Sutcliffe, J. G., and Lerner, R. A. (1982) Immunogenic structure of the influenza virus hemagglutinin, *Cell* 28, 477-487.



132. Wang, Y., Shao, Q., Yu, X., Kong, W., Hildreth, J. E., and Liu, B. (2011) N-terminal hemagglutinin tag renders lysine-deficient APOBEC3G resistant to HIV-1 Vif-induced degradation by reduced polyubiquitination, *Journal of virology* 85, 4510-4519.
133. Einhauer, A., and Jungbauer, A. (2001) The FLAG™ peptide, a versatile fusion tag for the purification of recombinant proteins, *J. Biochem. Bioph. Methods* 49, 455-465.
134. Gottschling, D., Seliger, H., Tarrasón, G., Piulats, J., and Eritja, R. (1998) Synthesis of oligodeoxynucleotides containing N 4-mercaptoethylcytosine and their use in the preparation of oligonucleotide-peptide conjugates carrying c-myc Tag-sequence, *Bioconjugate Chem.* 9, 831-837.
135. Honegger, B. K., Mocikat, R., and Pliickthun, A. (1997) Specific detection of his-tagged proteins with recombinant anti-His tag scFv-phosphatase or scFv-phage fusions, *Biotechniques* 22, 140-149.
136. Shah, N. H., and Muir, T. W. (2014) Inteins: nature's gift to protein chemists, *Chemical science* 5, 446-461.
137. Reddington, S. C., and Howarth, M. (2015) Secrets of a covalent interaction for biomaterials and biotechnology: SpyTag and SpyCatcher, *Curr. Opin. Chem. Biol.* 29, 94-99.
138. Zakeri, B., Fierer, J. O., Celik, E., Chittock, E. C., Schwarz-Linek, U., Moy, V. T., and Howarth, M. (2012) Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin, *Proceedings of the National Academy of Sciences* 109, E690-E697.

139. Zakeri, B., and Howarth, M. (2010) Spontaneous intermolecular amide bond formation between side chains for irreversible peptide targeting, *J. Am. Chem. Soc.* *132*, 4526-4527.
140. Wang, J., Yu, Y., and Xia, J. (2013) Short peptide tag for covalent protein labeling based on coiled coils, *Bioconjugate Chem.* *25*, 178-187.
141. Lai, Y.-T., Chang, Y.-Y., Hu, L., Yang, Y., Chao, A., Du, Z.-Y., Tanner, J. A., Chye, M.-L., Qian, C., Ng, K.-M., Li, H., and Sun, H. (2015) Rapid labeling of intracellular His-tagged proteins in living cells, *Proceedings of the National Academy of Sciences* *112*, 2948-2953.
142. Uchinomiya, S., Nonaka, H., Wakayama, S., Ojida, A., and Hamachi, I. (2013) In-cell covalent labeling of reactive His-tag fused proteins, *Chem Commun (Camb)* *49*, 5022-5024.
143. Uttamapinant, C., White, K. A., Baruah, H., Thompson, S., Fernández-Suárez, M., Puthenveetil, S., and Ting, A. Y. (2010) A fluorophore ligase for site-specific protein labeling inside living cells, *Proceedings of the National Academy of Sciences* *107*, 10914-10919.
144. Lin, C.-W., and Ting, A. Y. (2006) Transglutaminase-catalyzed site-specific conjugation of small-molecule probes to proteins in vitro and on the surface of living cells, *J. Am. Chem. Soc.* *128*, 4542-4543.
145. Antos, J. M., Chew, G.-L., Guimaraes, C. P., Yoder, N. C., Grotenbreg, G. M., Popp, M. W.-L., and Ploegh, H. L. (2009) Site-specific N- and C-terminal labeling of a single polypeptide using sortases of different specificity, *J. Am. Chem. Soc.* *131*, 10800-10801.

146. Carrico, I. S., Carlson, B. L., and Bertozzi, C. R. (2007) Introducing genetically encoded aldehydes into proteins, *Nature chemical biology* 3, 321-322.
147. Zhou, Z., Cironi, P., Lin, A. J., Xu, Y., Hrvatin, S., Golan, D. E., Silver, P. A., Walsh, C. T., and Yin, J. (2007) Genetically encoded short peptide tags for orthogonal protein labeling by Sfp and AcpS phosphopantetheinyl transferases, *ACS chemical biology* 2, 337-346.
148. Kiick, K. L., Saxon, E., Tirrell, D. A., and Bertozzi, C. R. (2002) Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation, *Proceedings of the National academy of Sciences* 99, 19-24.
149. Lang, K., Davis, L., Torres-Kolbus, J., Chou, C., Deiters, A., and Chin, J. W. (2012) Genetically encoded norbornene directs site-specific cellular protein labelling via a rapid bioorthogonal reaction, *Nature chemistry* 4, 298-304.
150. Chin, J. W., Santoro, S. W., Martin, A. B., King, D. S., Wang, L., and Schultz, P. G. (2002) Addition of p-Azido-l-phenylalanine to the Genetic Code of *Escherichia coli*, *J. Am. Chem. Soc.* 124, 9026-9027.
151. Exner, M. P., Kuenzl, T., To, T. M., Ouyang, Z., Schwagerus, S., Hoesl, M. G., Hackenberger, C. P., Lensen, M. C., Panke, S., and Budisa, N. (2017) Design of S-Allylcysteine in Situ Production and Incorporation Based on a Novel Pyrrolysyl-tRNA Synthetase Variant, *ChemBiochem* 18, 85-90.
152. Kramer, G., Kasper, P. T., de Jong, L., and de Koster, C. G. (2011) Quantitation of newly synthesized proteins by pulse labeling with azidohomoalanine, *Gel-Free Proteomics: Methods and Protocols*, 169-181.

153. Martell, J., and Weerapana, E. (2014) Applications of copper-catalyzed click chemistry in activity-based protein profiling, *Molecules* *19*, 1378-1393.
154. Avti, P. K., Maysinger, D., and Kakkar, A. (2013) Alkyne-azide “click” chemistry in designing nanocarriers for applications in biology, *Molecules* *18*, 9531-9549.
155. Hatzenpichler, R., Scheller, S., Tavormina, P. L., Babin, B. M., Tirrell, D. A., and Orphan, V. J. (2014) In situ visualization of newly synthesized proteins in environmental microbes using amino acid tagging and click chemistry, *Environ. Microbiol.* *16*, 2568-2590.
156. tom Dieck, S., Kochen, L., Hanus, C., Heumüller, M., Bartnik, I., Nassim-Assir, B., Merk, K., Mosler, T., Garg, S., and Bunse, S. (2015) Direct visualization of newly synthesized target proteins in situ, *Nature methods* *12*, 411-414.
157. Link, A. J., and Tirrell, D. A. (2003) Cell Surface Labeling of Escherichia coli via Copper (I)-Catalyzed [3+ 2] Cycloaddition, *J. Am. Chem. Soc.* *125*, 11164-11165.
158. Simon, M., Stefan, N., Borsig, L., Plückthun, A., and Zangemeister-Wittke, U. (2014) Increasing the antitumor effect of an EpCAM-targeting fusion toxin by facile click PEGylation, *Molecular cancer therapeutics* *13*, 375-385.
159. McClatchy, D. B., Ma, Y., Liu, C., Stein, B. D., Martínez-Bartolomé, S., Vasquez, D., Hellberg, K., Shaw, R. J., and Yates III, J. R. (2015) Pulsed Azidohomoalanine Labeling in Mammals (PALM) detects changes in Liver-specific LKB1 knockout mice, *Journal of proteome research* *14*, 4815-4822.
160. Dieterich, D. C., Hodas, J. J., Gouzer, G., Shadrin, I. Y., Ngo, J. T., Triller, A., Tirrell, D. A., and Schuman, E. M. (2010) In situ visualization and dynamics of

- newly synthesized proteins in rat hippocampal neurons, *Nature neuroscience* 13, 897-905.
161. Deal, R. B., Henikoff, J. G., and Henikoff, S. (2010) Genome-wide kinetics of nucleosome turnover determined by metabolic labeling of histones, *Science* 328, 1161-1164.
162. Ma, Y., Biava, H., Contestabile, R., Budisa, N., and di Salvo, M. L. (2014) Coupling bioorthogonal chemistries with artificial metabolism: intracellular biosynthesis of azidohomoalanine and its incorporation into recombinant proteins, *Molecules* 19, 1004-1022.
163. Agard, N. J., Baskin, J. M., Prescher, J. A., Lo, A., and Bertozzi, C. R. (2006) A comparative study of bioorthogonal reactions with azides, *ACS Chemical Biology* 1, 644-648.
164. Gajadeera, C. S., Zhang, X., Wei, Y., and Tsodikov, O. V. (2015) Structure of inorganic pyrophosphatase from *Staphylococcus aureus* reveals conformational flexibility of the active site, *Journal of structural biology* 189, 81-86.
165. Zhong, M., Fang, J., and Wei, Y. (2010) Site specific and reversible protein immobilization facilitated by a DNA binding fusion tag, *Bioconjugate Chem.* 21, 1177-1182.
166. Li, C., Wen, A., Shen, B., Lu, J., Huang, Y., and Chang, Y. (2011) FastCloning: a highly simplified, purification-free, sequence-and ligation-independent PCR cloning method, *BMC biotechnology* 11, 1.

167. Baykov, A., Evtushenko, O., and Avaeva, S. (1988) A malachite green procedure for orthophosphate determination and its use in alkaline phosphatase-based enzyme immunoassay, *Analytical biochemistry* 171, 266-270.
168. Lahti, R., and Jokinen, M. (1985) Kinetic model for the action of the inorganic pyrophosphatase from *Streptococcus faecalis*, *Biochemistry* 24, 3526-3530.
169. Bundy, B. C., and Swartz, J. R. (2010) Site-specific incorporation of p-propargyloxyphenylalanine in a cell-free environment for direct protein– protein click conjugation, *Bioconjugate Chem.* 21, 255-263.
170. Pédelacq, J.-D., Cabantous, S., Tran, T., Terwilliger, T. C., and Waldo, G. S. (2006) Engineering and characterization of a superfolder green fluorescent protein, *Nat. Biotechnol.* 24, 79-88.
171. Baykov, A., Cooperman, B., Goldman, A., and Lahti, R. (1999) Cytoplasmic inorganic pyrophosphatase, In *Inorganic Polyphosphates*, pp 127-150, Springer.
172. Kornberg, A. (1962) On the metabolic significance of phosphorolytic and pyrophosphorolytic reactions, *Horizons in biochemistry*, 251-264.
173. Lahti, R. (1983) Microbial inorganic pyrophosphatases, *Microbiological reviews* 47, 169.
174. Heikinheimo, P., Lehtonen, J., Baykov, A., Lahti, R., Cooperman, B. S., and Goldman, A. (1996) The structural basis for pyrophosphatase catalysis, *Structure* 4, 1491-1508.
175. Kankare, J., Neal, G. S., Salminen, T., Glumhoff, T., Cooperman, B. S., Lahti, R., and Goldman, A. (1994) The structure of *E. coli* soluble inorganic pyrophosphatase at 2.7 Å resolution, *Protein Engineering, Design and Selection* 7, 823-830.

176. Shintani, T., Uchiumi, T., Yonezawa, T., Salminen, A., Baykov, A. A., Lahti, R., and Hachimori, A. (1998) Cloning and expression of a unique inorganic pyrophosphatase from *Bacillus subtilis*: evidence for a new family of enzymes, *FEBS letters* 439, 263-266.
177. Sivula, T., Salminen, A., Parfenyev, A. N., Pohjanjoki, P., Goldman, A., Cooperman, B. S., Baykov, A. A., and Lahti, R. (1999) Evolutionary aspects of inorganic pyrophosphatase, *FEBS letters* 454, 75-80.
178. Young, T. W., Kuhn, N. J., Wadeson, A., Ward, S., Burges, D., and Cooke, G. D. (1998) *Bacillus subtilis* ORF yybQ encodes a manganese-dependent inorganic pyrophosphatase with distinctive properties: the first of a new class of soluble pyrophosphatase?, *Microbiology* 144, 2563-2571.
179. Kajander, T., Kellosalo, J., and Goldman, A. (2013) Inorganic pyrophosphatases: one substrate, three mechanisms, *FEBS letters* 587, 1863-1869.
180. Ahn, S., Milner, A. J., Fütterer, K., Konopka, M., Ilias, M., Young, T. W., and White, S. A. (2001) The “open” and “closed” structures of the type-C inorganic pyrophosphatases from *Bacillus subtilis* and *Streptococcus gordonii*, *Journal of molecular biology* 313, 797-811.
181. Fabrichniy, I. P., Lehtiö, L., Salminen, A., Zyryanov, A. B., Baykov, A. A., Lahti, R., and Goldman, A. (2004) Structural studies of metal ions in family II pyrophosphatases: the requirement for a Janus ion, *Biochemistry* 43, 14403-14411.
182. Fabrichniy, I. P., Lehtiö, L., Tammenkoski, M., Zyryanov, A. B., Oksanen, E., Baykov, A. A., Lahti, R., and Goldman, A. (2007) A trimetal site and substrate

- distortion in a family II inorganic pyrophosphatase, *Journal of Biological Chemistry* 282, 1422-1431.
183. Merckel, M. C., Fabrichniy, I. P., Salminen, A., Kalkkinen, N., Baykov, A. A., Lahti, R., and Goldman, A. (2001) Crystal structure of *Streptococcus mutans* pyrophosphatase: a new fold for an old mechanism, *Structure* 9, 289-297.
184. Chaudhuri, R. R., Allen, A. G., Owen, P. J., Shalom, G., Stone, K., Harrison, M., Burgis, T. A., Lockyer, M., Garcia-Lara, J., and Foster, S. J. (2009) Comprehensive identification of essential *Staphylococcus aureus* genes using Transposon-Mediated Differential Hybridisation (TMDH), *BMC genomics* 10, 291.
185. Halonen, P., Tammenkoski, M., Niiranen, L., Huopalahti, S., Parfenyev, A. N., Goldman, A., Baykov, A., and Lahti, R. (2005) Effects of active site mutations on the metal binding affinity, catalytic competence, and stability of the family II pyrophosphatase from *Bacillus subtilis*, *Biochemistry* 44, 4004-4010.
186. Parfenyev, A. N., Salminen, A., Halonen, P., Hachimori, A., Baykov, A. A., and Lahti, R. (2001) Quaternary structure and metal ion requirement of family II pyrophosphatases from *Bacillus subtilis*, *Streptococcus gordonii*, and *Streptococcus mutans*, *Journal of Biological Chemistry* 276, 24511-24518.
187. Rantanen, M. K., Lehtiö, L., Rajagopal, L., Rubens, C. E., and Goldman, A. (2007) Structure of the *Streptococcus agalactiae* family II inorganic pyrophosphatase at 2.80 Å resolution, *Acta Crystallographica Section D: Biological Crystallography* 63, 738-743.



188. Kehres, D. G., and Maguire, M. E. (2003) Emerging themes in manganese transport, biochemistry and pathogenesis in bacteria, *FEMS microbiology reviews* 27, 263-290.
189. Gur, E., Biran, D., and Ron, E. Z. (2011) Regulated proteolysis in Gram-negative bacteria—how and when?, *Nature Reviews Microbiology* 9, 839-848.
190. Olivares, A. O., Baker, T. A., and Sauer, R. T. (2016) Mechanistic insights into bacterial AAA+ proteases and protein-remodelling machines, *Nature Reviews Microbiology* 14, 33-44.
191. Gerdes, F., Tatsuta, T., and Langer, T. (2012) Mitochondrial AAA proteases—towards a molecular understanding of membrane-bound proteolytic machines, *Biochimica Et Biophysica Acta (BBA)-Molecular Cell Research* 1823, 49-55.
192. Sauer, R. T., and Baker, T. A. (2011) AAA+ proteases: ATP-fueled machines of protein destruction, *Annual review of biochemistry* 80, 587-612.
193. Gur, E., and Sauer, R. T. (2008) Recognition of misfolded proteins by Lon, a AAA+ protease, *Genes & development* 22, 2267-2277.
194. Zolkiewski, M. (2006) A camel passes through the eye of a needle: protein unfolding activity of Clp ATPases, *Molecular microbiology* 61, 1094-1100.
195. Flynn, J. M., Levchenko, I., Seidel, M., Wickner, S. H., Sauer, R. T., and Baker, T. A. (2001) Overlapping recognition determinants within the ssrA degradation tag allow modulation of proteolysis, *Proceedings of the National Academy of Sciences* 98, 10584-10589.

196. Lies, M., and Maurizi, M. R. (2008) Turnover of endogenous SsrA-tagged proteins mediated by ATP-dependent proteases in *Escherichia coli*, *Journal of Biological Chemistry* 283, 22918-22929.
197. Keiler, K. C., and Feaga, H. A. (2014) Resolving nonstop translation complexes is a matter of life or death, *Journal of bacteriology* 196, 2123-2130.
198. Chai, Q., Wang, Z., Webb, S. R., Dutch, R. E., and Wei, Y. (2016) The ssrA-Tag Facilitated Degradation of an Integral Membrane Protein, *Biochemistry* 55, 2301-2304.
199. Haswell, E. S., Phillips, R., and Rees, D. C. (2011) Mechanosensitive channels: what can they do and how do they do it?, *Structure* 19, 1356-1369.
200. Bass, R. B., Strop, P., Barclay, M., and Rees, D. C. (2002) Crystal structure of *Escherichia coli* MscS, a voltage-modulated and mechanosensitive channel, *Science* 298, 1582-1587.
201. Steinbacher, S., Bass, R., Strop, P., and Rees, D. C. (2007) Structures of the prokaryotic mechanosensitive channels MscL and MscS, *Current topics in membranes* 58, 1-24.
202. Li, C., Wen, A., Shen, B., Lu, J., Huang, Y., and Chang, Y. (2011) FastCloning: a highly simplified, purification-free, sequence-and ligation-independent PCR cloning method, *BMC biotechnology* 11, 92.
203. Leissring, M. A., Lu, A., Condrón, M. M., Teplow, D. B., Stein, R. L., Farris, W., and Selkoe, D. J. (2003) Kinetics of amyloid  $\beta$ -protein degradation determined by novel fluorescence-and fluorescence polarization-based assays, *Journal of Biological Chemistry* 278, 37314-37320.

204. Jennings, L. D., Lun, D. S., Médard, M., and Licht, S. (2008) ClpP hydrolyzes a protein substrate processively in the absence of the ClpA ATPase: mechanistic studies of ATP-independent proteolysis, *Biochemistry* 47, 11536-11546.
205. Checovich, W. J., Bolger, R. E., and Burke, T. (1995) Fluorescence polarization-- a new tool for cell and molecular biology, *Nature* 375, 254-256.
206. Heyduk, T., Ma, Y., Tang, H., and Ebright, R. H. (1996) Fluorescence anisotropy: rapid, quantitative assay for protein-DNA and protein-protein interaction, *Methods in enzymology* 274, 492-503.
207. Burton, R. E., Siddiqui, S. M., Kim, Y. I., Baker, T. A., and Sauer, R. T. (2001) Effects of protein stability and structure on substrate processing by the ClpXP unfolding and degradation machine, *The EMBO journal* 20, 3092-3100.
208. Miller, S., Bartlett, W., Chandrasekaran, S., Simpson, S., Edwards, M., and Booth, I. R. (2003) Domain organization of the MscS mechanosensitive channel of Escherichia coli, *The EMBO journal* 22, 36-46.
209. McLaggan, D., Jones, M. A., Gouesbet, G., Levina, N., Lindey, S., Epstein, W., and Booth, I. R. (2002) Analysis of the kefA2 mutation suggests that KefA is a cation-specific channel involved in osmotic adaptation in Escherichia coli, *Molecular microbiology* 43, 521-536.
210. Miller, S., Edwards, M. D., Ozdemir, C., and Booth, I. R. (2003) The closed structure of the MscS mechanosensitive channel cross-linking of single cysteine mutants, *Journal of Biological Chemistry* 278, 32246-32250.

211. Denisov, I. G., Grinkova, Y. V., Lazarides, A. A., and Sligar, S. G. (2004) Directed self-assembly of monodisperse phospholipid bilayer nanodiscs with controlled size, *J Am Chem Soc* 126, 3477-3487.
212. Segrest, J. P., Garber, D. W., Brouillette, C. G., Harvey, S. C., and Anantharamaiah, G. M. (1994) The Amphipathic Alpha-Helix - a Multifunctional Structural Motif in Plasma Apolipoproteins, *Adv Protein Chem* 45, 303-369.
213. Nolte, R. T., and Atkinson, D. (1992) Conformational-Analysis of Apolipoprotein a-I and E-3 Based on Primary Sequence and Circular-Dichroism, *Biophys J* 63, 1221-1239.
214. Bayburt, T. H., Grinkova, Y. V., and Sligar, S. G. (2002) Self-assembly of discoidal phospholipid bilayer nanoparticles with membrane scaffold proteins, *Nano Lett* 2, 853-856.
215. Shih, A. Y., Freddolino, P. L., Sligar, S. G., and Schulten, K. (2007) Disassembly of nanodiscs with cholate, *Nano Lett* 7, 1692-1696.
216. Bayburt, T. H., Grinkova, Y. V., and Sligar, S. G. (2006) Assembly of single bacteriorhodopsin trimers in bilayer nanodiscs, *Arch Biochem Biophys* 450, 215-222.
217. Lu, W., Zhong, M., and Wei, Y. N. (2011) Folding of AcrB Subunit Precedes Trimerization, *J Mol Biol* 411, 264-274.
218. Denisov, I., Grinkova, Y., Lazarides, A., and Sligar, S. (2004) Directed self-assembly of monodisperse phospholipid bilayer Nanodiscs with controlled size, *Journal of the American Chemical Society* 126, 3477-3487.

219. Badola, P., and Sanders, C. R. (1997) Escherichia coli diacylglycerol kinase is an evolutionarily optimized membrane enzyme and catalyzes direct phosphoryl transfer, *Journal of Biological Chemistry* 272, 24176-24182.
220. Sanders, C. R., Czernski, L., Vinogradova, O., Badola, P., Song, D., and Smith, S. O. (1996) Escherichia coli diacylglycerol kinase is an  $\alpha$ -helical polytopic membrane protein and can spontaneously insert into preformed lipid vesicles, *Biochemistry* 35, 8610-8618.
221. Zhou, Y., and Bowie, J. U. (2000) Building a thermostable membrane protein, *Journal of Biological Chemistry* 275, 6975-6979.
222. Yamashita, Y., Takehara, T., and Kuramitsu, H. (1993) Molecular characterization of a *Streptococcus mutans* mutant altered in environmental stress responses, *Journal of bacteriology* 175, 6220-6228.
223. Li, D., Lyons, J. A., Pye, V. E., Vogeley, L., Aragão, D., Kenyon, C. P., Shah, S. T., Doherty, C., Aherne, M., and Caffrey, M. (2013) Crystal structure of the integral membrane diacylglycerol kinase, *Nature* 497, 521-524.
224. Van Horn, W. D., and Sanders, C. R. (2012) Prokaryotic diacylglycerol kinase and undecaprenol kinase, *Annual review of biophysics* 41, 81-101.
225. Lau, F. W., Chen, X., and Bowie, J. U. (1999) Active sites of diacylglycerol kinase from Escherichia coli are shared between subunits, *Biochemistry* 38, 5521-5527.
226. Lahiri, S., Brehs, M., Olschewski, D., and Becker, C. F. (2011) Total chemical synthesis of an integral membrane enzyme: diacylglycerol kinase from Escherichia coli, *Angewandte Chemie International Edition* 50, 3988-3992.

227. Gorzelle, B. M., Nagy, J. K., Oxenoid, K., Lonzer, W. L., Cafiso, D. S., and Sanders, C. R. (1999) Reconstitutive refolding of diacylglycerol kinase, an integral membrane protein, *Biochemistry* 38, 16373-16382.
228. Vinogradova, O., Badola, P., Czerski, L., Sönnichsen, F., and Sanders, C. (1997) Escherichia coli diacylglycerol kinase: a case study in the application of solution NMR methods to an integral membrane protein, *Biophysical journal* 72, 2688-2701.
229. Nagy, J. K., Lau, F. W., Bowie, J. U., and Sanders, C. R. (2000) Mapping the oligomeric interface of diacylglycerol kinase by engineered thiol cross-linking: homologous sites in the transmembrane domain, *Biochemistry* 39, 4154-4164.
230. Lau, F. W., and Bowie, J. U. (1997) A method for assessing the stability of a membrane protein, *Biochemistry* 36, 5884-5892.

## Vita

### Personal information

Name	Xinyi Zhang
Place of Birth	China
Nationality	China

### **EDUCATION BACKGROUND:**

- 2007-2011 **B.S.** in Biotechnology
- China Pharmaceutical University, Nanjing, China

### **PUBLICATION:**

- Zhang, Xinyi, Wei Lu, Kevin Kwan, Dibakar Bhattacharyya, and Yinan Wei. "Dual-Functional-Tag-Facilitated Protein Labeling and Immobilization." *ACS Omega* 2, no. 2 (2017): 522-528.
- Sebastián Hernández, Cassandra Porter, Xinyi Zhang, Yinan Wei and Dibakar Bhattacharyya. "Layer-by-layer assembled membranes with immobilized porins" *RSC Adv.*, 2017,7, 56123-56136
- Priyesh Wagh, Xinyi Zhang, Yinan Wei, Isabel Escobar. "Alignment and Immobilization of Aquaporins on Polybenzimidazole Nanofiltration Membranes" (submitted)
- Wang, Zhaoshuai, Cui Ye, Xinyi Zhang, and Yinan Wei. "Cysteine residue is not essential for CPM protein thermal-stability assay." *Analytical and bioanalytical chemistry* 407, no. 13 (2015): 3683-3691.

- Gajadeera, Chathurada S., Xinyi Zhang, Yinan Wei, and Oleg V. Tsodikov. "Structure of inorganic pyrophosphatase from *Staphylococcus aureus* reveals conformational flexibility of the active site." *Journal of structural biology* 189, no. 2 (2015): 81-86.
- Chai, Qian, Brent Ferrell, Meng Zhong, Xinyi Zhang, Cui Ye, and Yinan Wei. "Diverse sequences are functional at the C-terminus of the *E. coli* periplasmic chaperone SurA." *Protein Engineering Design and Selection* (2014): gzu003.

**HONORS AND AWARDS:**

- PhD Student Fellow Award, Center of Membrane Sciences, Aug 2016
- Max Steckler Fellowship, University of Kentucky, Aug 2016
- Elias Klein Travel Supplement, North American Membrane Society (NAMS), Feb. 2016
- 3<sup>rd</sup> place in poster section, Naff Symposium, March 2015
- 2<sup>nd</sup> place in the oral competition, Kentucky Academy of Science meeting, Nov. 2014