UNIVERSITY OF CALIFORNIA SAN DIEGO

Oriented Sample Solid-State NMR of the Mercury Detoxification Membrane Protein MerFt

A Thesis submitted in partial satisfaction of the requirements for the degree of

Master of Science

in

Chemistry

by

Jiaqian Wu

Committee in charge:

Professor Stanley J. Opella, Chair
Professor Joshua Figueroa
Professor Patricia A. Jennings

2019

The Thesis of Jiaqian Wu is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

Chair

University of California San Diego

2019

DEDICATION

To my parents, Hongwei Wu and Xiaojun Qian,

for their consistent support and endless love.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

^{1}H Proton

^{2}H Deuterium

^{13}C Carbon-13

^{15}N Nitrogen-15

^{31}P Phosphorus-31

6-O-PC 1,2-di-O-hexyl-sn-glycero-3-phosphocholine

ACN Acetonitrile

AIBN Azobisisobutyronitrile

AMS Ammonium sulfate

CNBr Cyanogen bromide

CP Cross polarization

CSA Chemical shift anisotropy

Cv Column volume

DC Dipolar coupling

DCM Dichloromethane

DHPC 1,2-dihexanoyl-sn-glycero-3-phosphocholine

DMF Dimethylformamide

DMPC 1,2-dimyristoyl-sn-glycero-3-phosphocholine

DMPG 1,2-dimyristoyl-sn-glycero-3-phospho-(1'-rac-glycerol)

DTT Dithiothreitol

EDTA Ethylenediamine tetracetic acid

FPLC Fast protein liquid chromatography

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

HPC Hexadecylphosphocholine

HPLC High performance liquid chromatography

HSQC Heteronuclear single quantum coherence

INEPT Insensitive nuclei enhanced polarization transfer

IPTG Isopropyl -D-1-thiogalactopyranoside

kD Kilo-dalton

LB Luria-Bertani broth

M9 Minimal media

MAS Magic angle spinning

MSP Membrane scaffolding protein

NBS N-Bromosuccinimide

Ni-NTA Nickelnitrolotriacetic acid

NOE Nuclear Overhauser effect NMR Nuclear magnetic resonance

PBS Phosphate buffered saline

PCS Pseudocontact shift

PDB Protein data bank

PMSF Phenylmethanesulfonylfluoride

PRE Paramagnetic relaxation enhancement

SDS Sodium dodecyl sulfate

SDS-PAGE Sodium dodecyl sulfate polyacrylamide gel electrophoresis

ssNMR Solid-state NMR

TCEP Tris(2-carboxyethyl) phosphine

TFA Trifluoroacetic acid

TM Transmembrane

TX-100 Triton X-100

UAA Unnatural amino acid

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to acknowledge Professor Stanley Opella for his support as chair of my committee. I would like to thank him for letting me conduct research in his laboratory and embrace my research result both good result and bad result. I would like to acknowledge Professor Joshua Figueroa and Professor Patricia A. Jennings who kindly served as my committee and give me advises.

I would also thank to all my colleagues in Opella lab. The whole lab was immersed in joyful atmosphere. Everyone has been so patient to me and give me so many useful advises. I would like to thank Zheng Long who guided me started in the wet lab. He patiently leaded me to use all the instrument in the wet lab and encourage me when I was trapped in difficulties. I would thank to Dr. Sang Ho Park who helped me troubleshot the protein purification process and really help me get start with this thesis. I also would like to thank Dr. Jasmina Radoicic for encouragement and all the nice conversation on both life and research. Dr. Luis Mansor Basso gave me so many practical suggestions and shared with me the information about protein purification. I would thank Dr. Thach Can for the collaboration with some lab instruments. Youjeoung Na and I are both lab mates and friends in life. I really enjoy talking to her about both research program and campus life.

I also appreciate Dr. Anna De Angelis for arranging the group meeting schedule and making everything running smoothly. I would like to say thank

you to Dr. Xuemei Huang who navigated me on solution NMR spectrometer instrument. Without you, the basement of natural science building will not be so much fun.

I would like to thank my instructor Dr. Carl Hoegor and Dr. Kim Albizati. It is very delightful to be a teaching assistant for them. I would like to thank my teaching colleagues for helping me to be a better teacher and bring so much fun during boring grading time.

Lastly, I would like to thank all my family members and friends for supporting me and encouraging me all the time. Without you, I will not purchase my goal in natural science. Big thanks to my parents, Hongwei Wu and Xiaojun Qian, who always believe in me and give me support. My roomies, also my best friends, Hui and Yahui bring me delight and joy. Thank you for listening to me and help me escape from stressful graduate student life.

Thank everyone for the past one year. Without their kindness and support this thesis cannot be completed.

VITA

2013 Bachelor of Science, University of Sheffield, Sheffield, the United Kingdom

2013 Bachelor of Science, University of Nanjing Tech University, Nanjing, China

2019 Teaching Assistant, Department of Chemistry and Biochemistry, University of California, San Diego

2019 Master of Science, University of California San Diego

ABSTRACT OF THE THESIS

Oriented Sample Solid-State NMR of the Mercury Detoxification Membrane Protein MerFt

by

Jiaqian Wu

Master of Science in Chemistry

University of California, San Diego, 2019

Professor Stanley J. Opella, Chair

Mer proteins are 6 proteins MerA, MerE, MerF, MerT, MerC and MerP performing bacterial mercury detoxification system. One of transmembrane protein MerFt with two a helix was mutated two sets of vicinal cysteines: C21, C22, C71, and C72 with Serine pairs. The truncated structure MerFt, as a nonmetal binding protein and ideal transmembrane model protein to

study solution NMR and solid-state NMR sample condition, was successfully expressed as fusion protein with Ketosteroid Isomerase as expression tag. CNBr (cyanogen bromide) chemical cleavage was performed to cleave KSI tag in different conditions. The protein was successfully purified with HPLC (High Performance Liquid Chromatography) and FPLC (Fast protein liquid chromatography) size exclusion. The structural studies of this membrane protein were developed on solution NMR (Nuclear Magnetic Resonance) and solid-state NMR.

Chapter 1: Introduction to the study of membrane protein MerFt by NMR

1.1 Introduction to membrane proteins

Membrane proteins are defined as a category of protein either can permanently anchored or part of the membrane of cells or organelles (integral membrane proteins) or only temporarily bind to the lipid bilayer or to other integral proteins (peripheral member proteins). For integral membrane proteins, two categories of protein were classified. One is transmembrane proteins which span across the membrane and the other one is integral monotonic proteins which merely associate with one side of biological membrane. As a curial class of proteins, membrane proteins functionalize in signaling, trafficking, transport, adhesion, and recognition and are targets of more than 50% of modern medicine.(Overington, Al-lazikani, & Hopkins, 2006) Mutations or improper folding of these proteins is associated with many known diseases such heart disease, cystic fibrosis, depression, obesity, cancer and many others.(Moraes, Evans, Sanchez-Weatherby, Newstead, & Stewart, 2014) However, structural information about these proteins is deficient due to difficulty on provide native membrane environment. (Ubarretxena-Belandia & Stokes, 2010)Three methods were developed to study structural biology of member protein- X ray crystallography, transmission electron cryomicroscopy and nuclear magnetic resonance. X ray crystallography is a method based on diffraction of protein crystal. By

repeatedly expose crystal protein to X-ray beam in different orientation, diffraction data points are collected and analyzed to calculate electron density of the protein molecule. The electron density provide information of location of atoms which is the fundamental information of 3-D structure of protein.(Mitzi, D.B., Feild, C.A., Harrison, W. T. A., Guloy, 1994) The limitation of this technique is in order to crystalize protein, mutations are made. The method has difficulties on direct determination of secondary structures and especially domain movements. For cyro-EM, the challenges including the low contrast and high noise level of imaging vitrified macromolecules. (Renaud et al., 2018) NMR is another structural biology analysis method which bases on signal observation of nuclear spin in external magnetic field. It is possible to study time dependent phenomena with NMR, such as intramolecular dynamics in macromolecules, reaction kinetics, molecular recognition or protein folding. However, the resolving power of NMR is less than X-ray crystallography and it is more complex to gain same amount of information comparing to x ray crystallography. According to devoted work of NMR specialist, the limitation and disadvantage are mitigating.

1.2 Introduction to protein purification

Protein purification is a set of methods that isolate target protein from a complex mixture. For those proteins of interest express in cells, the primary step is disruption of cells and extraction of protein. The disruption methods

includes sonication, homogenization by high pressure, permeabilization by detergents (e.g. Triton X-100) and/or enzymes (e.g. lysozyme). Centrifugation is another vita technique in protein purification utilizing centrifugal force to separate suspended particles and liquid.

There are three main properties of proteins are used to separate proteins. First, according to various molecular weight and size, size exclusion chromatogram and sys page gel was developed. The similarity of these two techniques is porous gel or column are served as immobile phase. Larger size or molecular weight protein elute first, and smaller protein can go through the porous gel and elute late. The difference is visualization of proteins. In size exclusion chromatogram, UV absorption was applied to column elution to plot out chromatogram. For SDS page gel, at the end of the electrophoretic separation, all proteins are sorted by size and can then be analyzed by other methods as staining with Coomassie staining or other staining solutions. (Rigaut et al., 1999) Second, polarity and hydrophobicity allow protein isolation with different gradient, the technique was called high performance liquid chromatogram. Thirdly, for those protein have various isoelectric points, running them through a pH graded gel or an ion exchange column can give resolution of protein mix.

1.3 NMR applications to membrane proteins

There are two different NMR types known as solution state NMR and solid-state NMR. Both these two methods depend on NMR active isotopes such as $_{1}H$, $_{2}H$, $_{13}C$, $_{15}N$, $_{17}O$ extra. For protein, which is a long chain of amino acid residues, only 1H have the adorable property of existing at high natural abundance. Due to the low natural abundance of $_{13}C$ (1.07%), $_{15}N$ (0.36%), the NMR detection can be difficult. Thus protein are normally grown in isotopic media for example minimal media (M9) which allow for enhancement of labeling of the 1/2 spinning nuclei.(Clewell & Helinski, 1972) Although both of two method based on signal of nuclear spin in external magnetic field, there are several different part between these two methods such as sample preparation, sensitivity, resolution and compatible protein size. For soluble proteins, solution NMR was utilized to study structural biology and dynamic dimensions. The sample preparation of solution NMR is well-established. For member protein which have low solubility in regular solvent for soluble protein, SDS is used as detergent to soluble member protein. Since the chemical environment for membrane protein is critical for structure study, more natural lipid environment is discovered such as bicelle, nanodisces. However, In solution NMR molecules undergo Brownian motion or fast isotropic tumbling (Radoicic, Lu, & Opella, 2014)which averages out anisotropic interactions on the NMR time scale and results in sharp transition peaks. Because of this, solution NMR is limited to the study of proteins smaller than 70 KDa.

Decreasing in relaxation time which significantly broadens line widths and reduction of signal intensity also limit the micelle and bicelle size which have no benefit to stabilize the membrane protein. For solid state NMR, there is no size limitation (Opella, 2015) and membrane protein is reconstituted in Large bicelles (Opella, 2013) , nanodiscs(Ritchie et al., 2009), and macrodiscs such as large phospholipid bilayer or newly-developed SMA orientated sample environment (Radoicic, Park, & Opella, 2018). In addition, the proteins should not undergo isotropic tumbling. Instead their mobility is restricted enough to only undergo rotational diffusion about the bilayer normal(Jardetzky & Wade-Jardetzky, 1980). Thus, the nuclear spin systems in ssNMR depend on the anisotropic or directionally dependent interactions. The two main anisotropic interactions are chemical shift anisotropy (CSA) and internuclear dipolar coupling (DC). Since these interactions are directionally dependent, they normally result in broad spectral line. To solve this, the sample can be set at the magic angle (Hayase, Kim, & Kato, 1986) (54.74°) to average the DC to zero and the CSA to a non-zero value. The sample can also be oriented or aligned (either mechanically or magnetically) so the DC and CSA can be measured. In both NMR methods, optimization of the type of lipid and lipid environment are crucial in obtaining structures of membrane proteins in their near-native state.

1.4 Mer protein series from bacterial

To maintain living organism function, so-called bulk elements are essential O, C, H, N, and S. They are the building blocks of the compounds that constitute organs and muscles. Additionally, P is a key constituent of both DNA and RNA which the genetic building blocks of living organisms are, Na, K, Ca, Mg, Cl ions are vita in maintaining the osmotic pressure body fluids. Trace element such as Fe, Zn, Cu, Mn, Mo, Co etc. functionalize as redox reagents to perform special bioactivity in living organism. High concentration of these trance element can be toxic to the cell thus transport system was self-adapted to retain the intercellular component level. Due to exposure of toxic mercury ion, many bacterial organisms adapt additional pathway to detoxify mercury ions named mercury detoxification system.(Howell, Mesleh, & Opella, 2005)

Minamata disease is a neurological syndrome caused by severe mercury poisoning(Harada, 1995). Patient symptoms include ataxia, numbness in the hands and feet, general muscle weakness, loss of vision and damage to hearing and speech. Due to some human-generated pollution including industry waste and fertilizer mercury and organomercury compound can be absorbed by human body through direct contact with unprotective skin and food web. However, some bacterial would not be affect by mercury poisoning due to the mercury detoxification system.

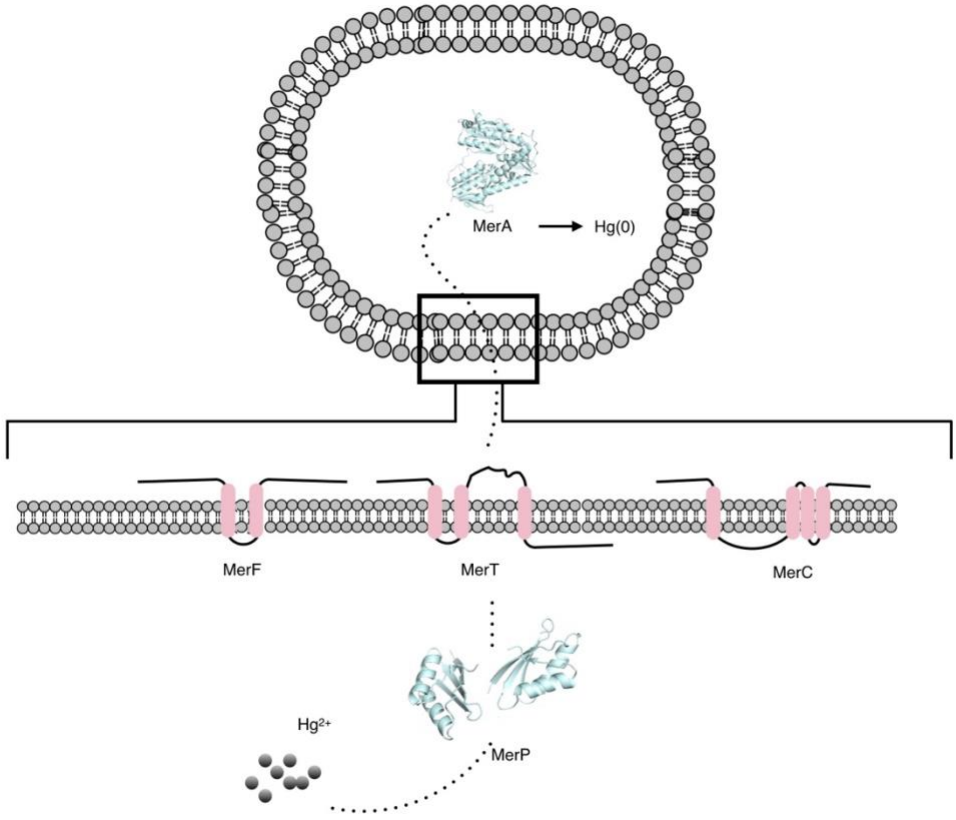

Figure 1.4.1: Scheme of mer operon detoxification process. Mercury ion diffused in bacterial cell and shuttled by MerP (periplasmic Hg(II) scavenging protein) . Membrane protein MerF, MerT and MerC transported the mercury ion to MerA and redued to Hg (0). Then the metallic mercury ia able to diffuse out of the cell.

From previous research, in liquid crystalline phospholipid bilayers environment, the structure of MerFt is determined by Rotationally Aligned (RA) solid-state NMR.(De Angelis, Howell, Nevzorov, & Opella, 2006) It has two long helices, which extend well beyond the bilayer, with a well-defined interhelical loop. The mercury binding site is ranging from the interface on the cytosolic side to the middle of the first transmembrane helix.

Chapter 2: Expression and purification of MerFt protein

2.1 Abstract

The sequence of MerFt was designed as mutated and truncated version of wild type MerFt. To eliminate aggregation, cysteine pairs were mutated to serine. Serine was chosen as a cysteine substitute so that the protein would continue to contain an isosteric sidechain with a polarity similar to the original cysteine but without the forming disulfide bonds. The truncation was designed to remain the hydrophobic domains of MerFt and reduce the spectral overlap caused by the terminal unstructured regions.

MerFt was successfully expressed and purified using a modified pET-31b(+) vector. Using this expression plasmid, MerFt is driven into inclusion bodies using an N-terminal ketosteroid isomerase (KSI) fusion partner. A histidine tag was used to separate the MerFt fusion from *E. coli* proteins. Yield of 15N uniformed labeled fusion protein is 80 mg per liter cell culture. This vector contains one methionine flanking each end of the MerFt for cyanogen bromide chemical cleavage from the fusion protein. Post chemical cleavage protein was further purified with high performance liquid chromatography (HPLC) and fast protein liquid chromatography (FPLC). Final yield of MerFt is 4 mg per liter cell culture.

2.2 Introduction

Expression of MerFt protein was performed based on the previous lab member, laboratory E.Coli C41 (λDE3) and E.ColiC43 (λDE3) were used and comparison of fusion protein yield was performed. Due to MerFt protein is a highly aggregated and insoluble protein, expression was targeted in inclusion body with ketosteroid isomerase (KSI) in N-terminal ((Palmer, 2004).). 6* histage on c terminal allow the targeted fusion protein separated from E coli protein by Ni-NTA affinity column. One of 2 Methiones was placed between KSI and MerFt, another one was placed between MerFt and C terminal his tag, allowing CNBr cleavage. (Christie & Gagnon, 2015) In addition, MerFt sequence does not contain any methionine residue which making the cyanogen bromide chemical cleavage a suitable method in this protein purification. Modification on the construct was done to simplify the purification protocol. 2 kinds of denatured reagent buffer were carried out in this purification. The purification efficiency will be discussed later. Size exclusion and HPLC were applied on further purification with different efficiency which will be clarified later in this chapter.

2.3 Material and experimental

2.3.1 Construct design, expression and growth of MerFt

Two constructs were used in MerFt protein purification process. The original plasmid pET-31(+)b construct consists of the MerFt protein spanned by two CNBr cleavable methionine. The protein is fused to KSI on the N-terminal side, and a 6-His tag on the C-terminal side. The modified construct was designed by inserting a second 6-His-tag between KSI and MerFt, providing possible extra means for KSI removal in later steps of the protein purification.

Protein production was carried out using the C43 (DE3) strain from Lucigen (www.lucigen.com). Uniformly ^{15}N enriched NMR sample growth was carried out in ^{15}N-labeled ammonium sulfate enriched M9 media(1x M9 salts (6 g/L Na2HPO4, 3 g/L KH2PO4, and 0.5 g/L NaCl ,1 g/L ammonium sulfate ams, 50 ml 10% glucose, 25 ml post media (1x MEM vitamin solution (Thermo Fisher Scientific), trace elements, 1 mM MgSO4, 0.1 mM CaCl2, 0.03 mM thiamine), 50 mg/L carbenicillin). 14N test-proposed protein was carried out in Luria Bertani (LB) media with 50 mg/L carbenicillin. LB agar plates were used to grow fresh transformations of MerFt into C43 (DE3) competent cells (Lucigen). Single colony was inoculated in LB media (12.5 ml, 37oc, 8hr) shaken at 300 rpm. Overnight LB cell culture was inoculated in M9

media (500 ml, 37oc, 4hr) and shaken 300 rpm. Protein production was induced by 1mM ITPG upon reaching a cell density of 0.6 OD. Cell culture was allow to grow for 4 hr with final O.D. 4.0.

2.3.2 Chemical cleavage and purification of MerFt

Upon pelleting the cells in 5000 rpm centrifuge and cell lysis with sonication in lysis buffer (30ml/half liter culture pellet, 40mM tris base, 12% glycol, 10 mg lysozyme), the KSI-MerFt fusion protein remains in the inclusion body fraction of the cell lysate. Binding buffer (SDS binding buffer: 1% SDS, 40 mM Tris-base, 10mM Imidazole, 250 mM NaCl, pH 8.0, Guanidinium chloride binding buffer: 6M Guanidinium chloride, 40 mM Tris-base, 10mM imidazole, 250 mM NaCl, pH 8.0) was applied on dissolving inclusion body. Fusion protein isolation was carried out with Ni-NTA resin (10 ml for half liter cell culture wet pellet) (Qiagen) shaken for 4 hrs, elution buffer (SDS buffer: 1% SDS, 40 mM Tris-base, 500 mM Imidazole, 250 mM NaCl, pH 8.0. Guanidinium chloride buffer: : 6M Guanidinium chloride, 40 mM tris base, 500 mM imidazole, 250 mM NaCl, pH 8.0) and washing buffer(SDS buffer: 1% SDS, 40 mM tris base, 250 mM NaCl, pH 8.0 Guanidinium chloride buffer: 6M Guanidinium chloride, 40 mM Tris-base, 250 mM NaCl, pH 8.0). SDS and Guanidinium chloride elution fraction (20 ml) was subsequently removed by extensive dialysis against di h2o. The protein was lyophilized, dissolved and cleaved with CNBr in 70% formic acid (30 mg fusion protein in 2ml 70% formic

acid). The result solution was diluted to 10-fold dilution. CNBr and formic acid was dialyzed out against di h2o. A Sephacryl S-200HR column was used to isolate the cleaved MerFt protein in FPLC SDS buffer (4mM SDS, 1.2 mM NaH_2PO_4, 18.8 mM Na_2HPO_4, 1 mM EDTA, 1 mM NaN_3, pH 8.2). C4 reversible column HPLC was used to isolated the cleaved protein with C4 reversible column (buffer A: 94.9% Di-water, 3% isopropanol, 2% acetonitrile, 0.1% trifluoracetic acid; buffer B: 47% isopropanol, 28% acetonitrile, 20% 2,2,2-trifluoroethanol, 5% Di-water, 0.1% Trifluoracetic acid; MerFt elution gradient: 98% buffer B) The result protein was determined by SDS page gel, solution NMR, solid state NMR. Experiment run was 1H detected HSQC experiment, with 15N in the indirect dimension. Refer to NMR spectrum of previous paper, MerFt is successfully expressed and purified.

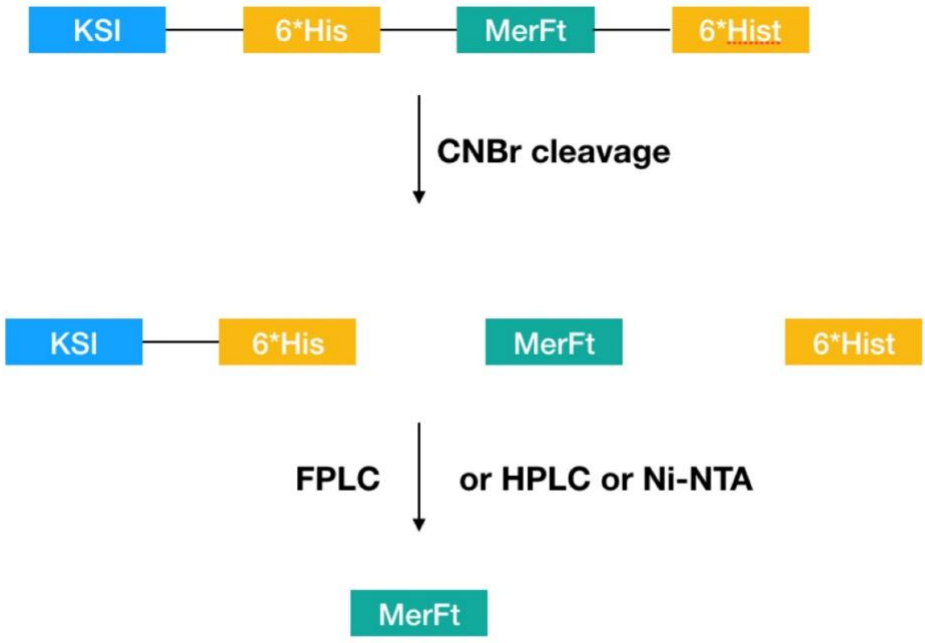

Figure 2.3.2.1: Purification scheme of MerFt. MerFt is expressed as a fusion with KSI. Two 6-histidine tag is insertedin fusion protein, one is between KSI and MerFt, the other one is in terminal. A single methionine was added right before the MerFt gene to allow for CNBr cleave of KSI and MerFt, another methionine was added between MerFt and his tage. The Ni-NTA column yields pure fusion protein. The fusion is cleaved by CNBr. Uncleaved fusion protein, KSI, and MerFt are separated by FPLC, HPLC or Ni-NTA, and MerFt is eluted out as a single peak for NMR studies.

2.4 Result and discussion

Initially, expression and purification on MerFt protein was not optimized due to several reasons. First, the E coli. Cell was found by induction test. According to pervious group member, MerFt can be successfully expressed in both E Coli. C41 (DE3) and E coli. C43 (Lamdar DE3) (Lucigen). Induction test was conducted to test the yield of better competent cell. From Figure

2.4.1, both of them can express the target protein. However, the yields of LB media growth protein was different, 48 mg per liter cell culture and 39 mg per liter cell culture in C43 and C41 respectively. Thus, E Coli C43(DE3) was selected to carry the later protein expression. Secondly, cell growth conditions were examined especially in M9 media. Glucose and glycerol can be used as carbon source in M9 media, yielding 17 mg/L culture and 7 mg/L culture respectively. Thus, glucose was selected as carbon source in M9 media. The growth protocol was test. The cell culture was grown in small overnight first (100ml M9) to make E. coli accommodate M9 media environment. However, the accommodation step was tested as a main reason of low fusion protein yield. Thirdly, SDS and guanidinium chloride are two denatured solution which normally used in membrane protein purification. Comparing the SDS page gel

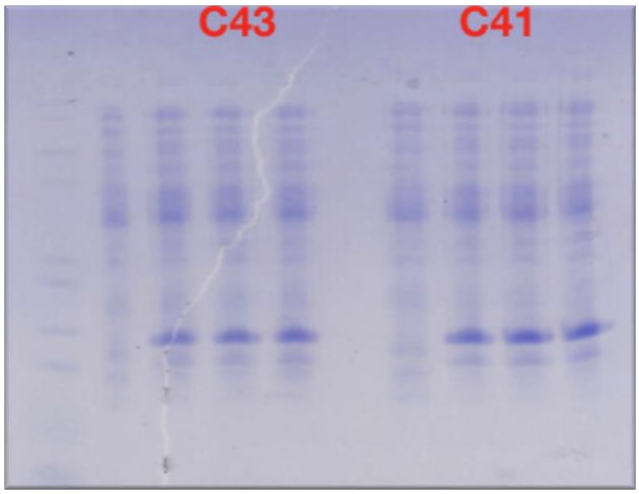

Figure 2.4.1 Induction test with two different competed cell E Coli C41 (DE3) and E Coli C43(DE3). From left lane to right lane, they are before inducing with IPTG, after IPTG induction 1hour, 2 hour, 4 hour

for purified fusion protein (Figure. 2.4.6), the clear one band for fusion protein

was shown in SDS purification indicate that SDS give more specific binding to Ni-NTA column. Fourthly, CNBr chemical cleavage conditions were tested by before cleavage sample preparation, cleavage media (1% SDS, 70% TFA, 70% formic acid cleavage media will be discussed in next paragraph), cleavage time length, after cleavage naturalization. From Niemann and his colleagues publication,(Niemann & Mole, 1982) the efficient CNBr cleavage was conducted by 70% formic acid. Initially, the sample (fusion protein powder) was dissolved in 88% formic acid then diluted to 70% formic acid. From SDS page gel, smear lane which indicate protein aggregation was shown. Low pH harsh condition can be a reason that protein degradation. Then 70% formic was prepared before dissolve the protein. However, the problem was not solved. Neutralization could be a reason that protein aggregated due to adding high concentration of NaOH cause local temperature increasing. However, there are two reasons that neutralization was necessary. One is upon pH adjusted to neutral, the CNBr cleavage was stopped(Rodríguez, Wong, & Jennings, 2003)Another one is while CNBr was removed by dialyzing, the pH ranges that dialysis membrane (Spectrum) can tolerate is pH 2-12. 70% formic acid was pH 1.29 which not compatible with dialysis bag. According to Werner Reutrmann paper(Christoph, Schmidt, Steinberner, Dilla, & Karinen, 2012), 56.9 kilojoule was generated per mole formic acid was neutralized. In this case 10 ml 70% formic acid was used to dissolve 40 mg fusion protein. For neutralize purpose, 15 ml of 1M NaOH was

added in after cleavage mixture and approximately 5 Celsius temperature increase immediately resulting aggregation and on SDS page gel the band is smear (Figure 2.4.2). Instead of naturalization, the after-cleavage aliquot was diluted to stop the cleavage and avoiding low recovery percentage in dialysis bag. From Figure 2.4.3, aggregation problem was solved in this step. Cleavage time length was determined by taking cleavage time point and analyzed by SDS page gel. From the SDS page gel (Figure 2.4.4), after 2 hours the cleavage was not conducted further. Initially, size exclusion chromatography was applied on after cleavage protein to separate target protein from KSI and other impurity. The chromatograph (Figure 2.4.5) show that resolution is questionable. High performance liquid chromatography (HPLC) was then applied on post-cleavage protein. HPLC injection sample preparation was worth-mention. The post-cleavage protein powder was initially dissolved in pre-made organic solvent included 50% TFE, 49.9% Di-water and 0.1% TFA. The chromatography was gained and from SDS page gel analysis while buffer A (aqueous phase) ran in the first 10 minutes. The injection did not bind to C4 column. This problem was solved by optimizing injection sample preparation. TFE was added first to fully dissolve protein powder and make protein exposure to organic solvent which is hydrophobic. Then equivalent volume of di water and 1% TFA mixture was added. The

chromatography (Figure 2.4.6) indicated the protein properly binds on HPLC C4 column.

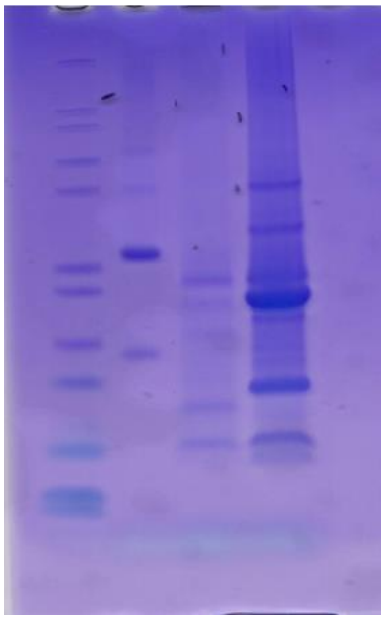

Figure 2.4.2 SDS page gel of neutralized after cleavage sample, from left to right lanes are fusion protein, after CNBr cleavage fraction 1 and 2

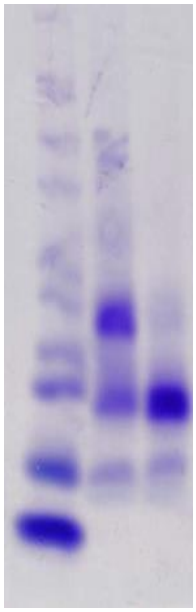

Figure 2.4.3 SDS page gel of no neutralization after cleavage aliquot from left to right lanes are cleaved fusion protein without DTT and cleaved fusion protein with DTT

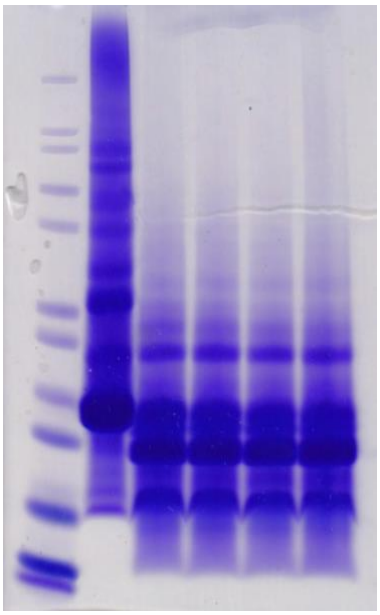

Figure 2.4.4 SDS page gel of CNBr cleavage time point from left to right lans are before cleavage, 1 hour cleavage time point, 2 hour cleavage time point, 3 hour cleavage time point, 4 hour cleavage time point

.

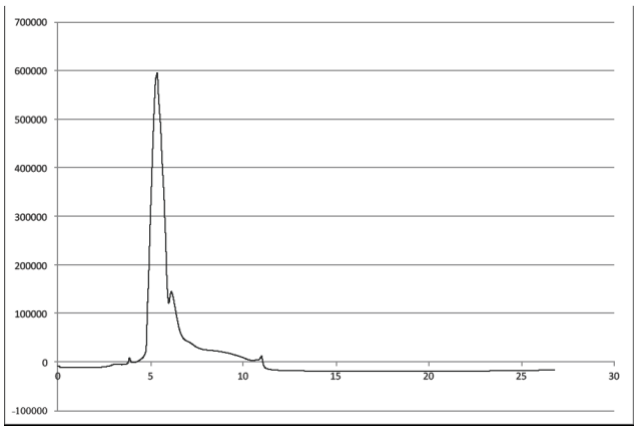

Figure 2.4.5 HPLC without optimized sample preparation

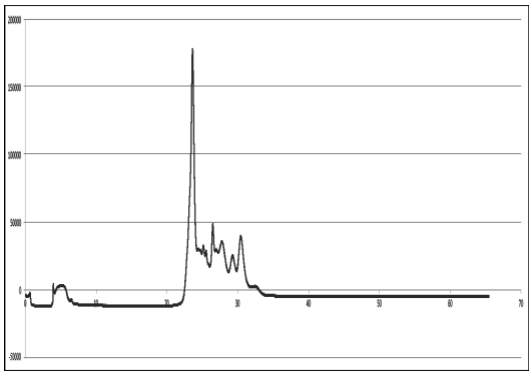

Figure 2.4.7 HPLC with optimized sample preparation

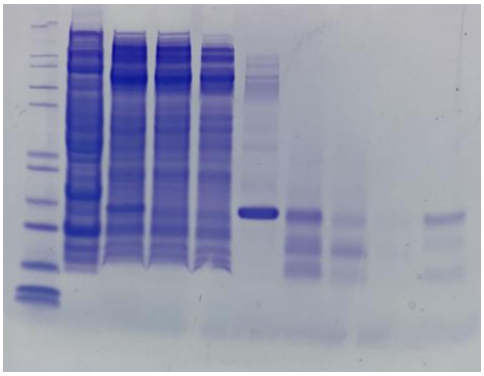

Figure 2.4.8 SDS page gel of CNBr cleavage with 1% SDS rom left to right lanes are supernatant of inclusion body, inclusion body, flowthrough, Ni-NTA wash 1 , Ni-NTA wash2, post-cleavage, flowthrough 2nd Ni-NTA column, washing 2nd Ni-NTA column, elution 2nd Ni-NTA column.

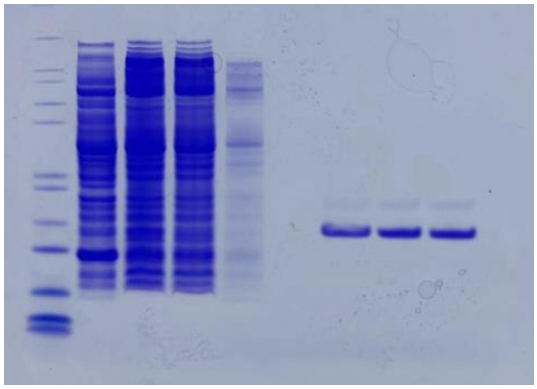

Figure 2.4.6 SDS page gel of SDS purified fusion protein. From left to right lanes are supernatant of inclusion body, inclusion body, flow through from Ni-NTA column, wash1, wash2
elution1, fusion protein in dialysis process

Sodium dodecyl sulfate, an anionic surfactant used as detergent in daily life clean product, is also utilized in MerFt protein purification. (Andersen et al., 2009)The Guanidinium chloride purification protocol was followed in SDS purification except that 6M Guanidinium chloride was switched to 1% SDS. Compare to the fusion protein purified with Guanidinium chloride, SDS give specific binding with Ni column and relatively purer fusion protein was gained (Fig 2.4.6). Since 1% SDS form micelle which size around 20 Kda(Yoshii & Okazaki, 2006). It is time-consuming to remove SDS and once SDS was

removed protein have chance to aggregate. Thus 1% SDS was selected as CNBr cleavage media. Since from Kell K. Andersen and coworker paper(Andersen et al., 2009), efficient chemical cleavage have to be conducted under acid environment. HCl was used to adjust pH to 3, the cleavage efficient indicated that 1% SDS can be used as CNBr cleavage media (Figure. 2.4.8). The purification process after cleavage based on the construct designed by adding 6* histidine between KSI and MerFt. The new construct allowed KSI- 6* histidine attached on Ni resin (Figure. 2.4.8). The purification of MerFt is not successful through this construct. 70% TFA was also used as CNBr cleavage media, the SDS page gel of FPLC fractions were shown that two peak which have similar molecular weight and difficult to separated are MerFt (Figure 2.4.9). However upper band in SDS page gel which correspond to second last peak in FPLC chromatography is misfolded. The solution NMR HSQC experiment spectrums of them are discussed in chapter 3. Thus, HPLC purification technique was applied to separate the 2 similar size species. From the HPLC chromatogram and corresponding SDS page gel. The good resolution of last two peaks indicate the 2 similar size spices are successfully separated. The two lanes before mark 12 in SDS page gel (Figure 2.4.10) show the separation. Both of them were packed in NMR sample subject to HSQC solution NMR spectroscopy.

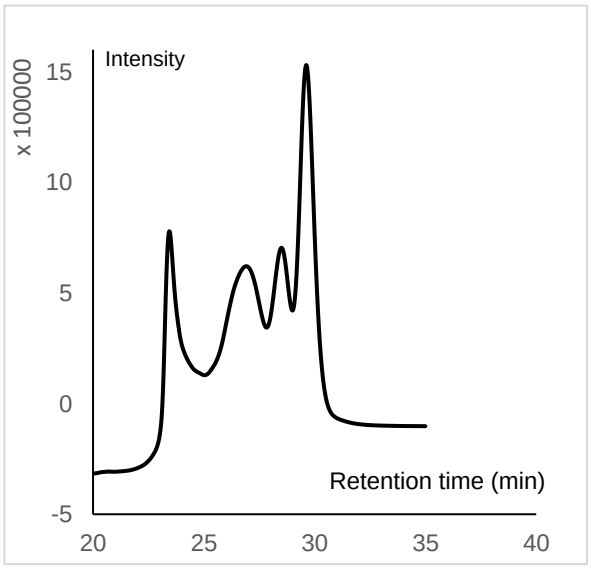

Fig. 2.4.9 HPLC chromatogram

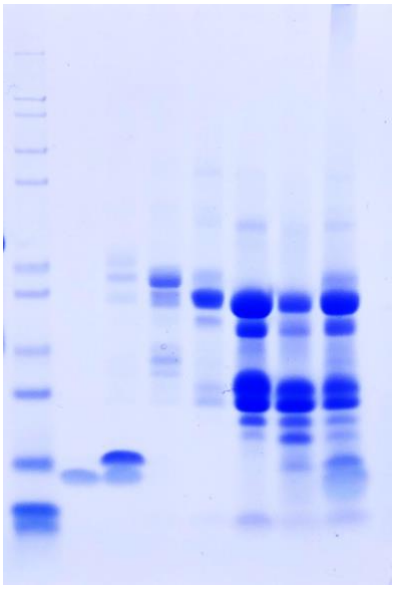

Fig. 2.4.10 SDS page gel of HPLC factions

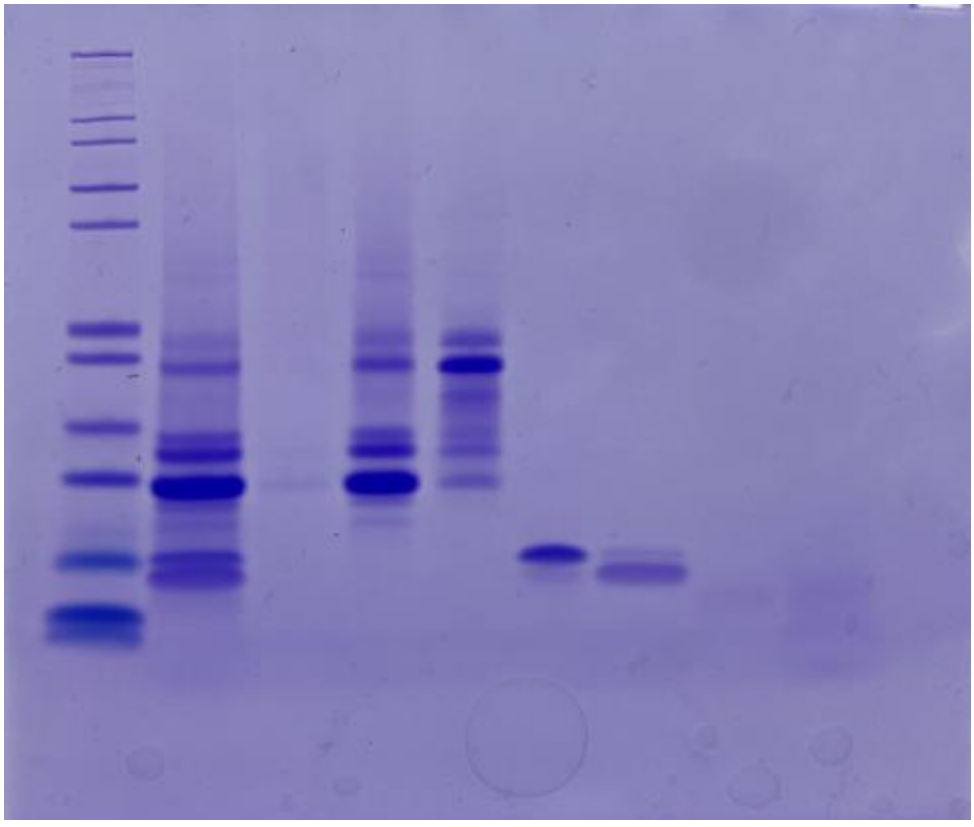

Figure 2.4.11 FPLC of resulting after cleavage mixture with 70%TFA as cleavage media

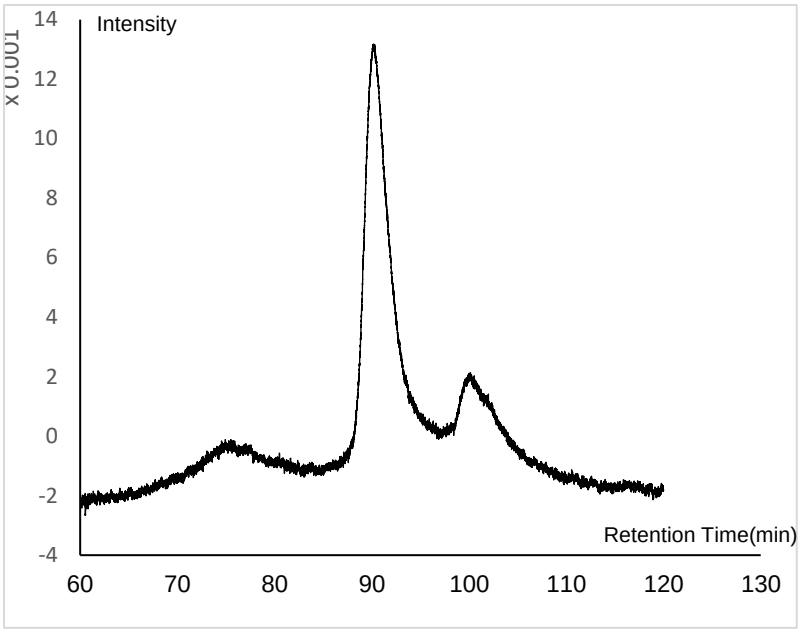

Fig. 2.4.12 FPLC chromatography of post CNBr cleavage fusion protein

2.5 Conclusion

In conclusion, MerFt protein purification was conducted by two denature chemical reagent guanidine chloride and sodium dodecyl sulfate. The CNBr chemical cleavage was conducted in three media, 70% formic acid, 1% SDS, 70% TFA respectively. SDS purification was easier to monitor by SDS page gel, avoid aggregation and misfolded target protein. The detailed MerFt protein expression, growth and purification protocol is developed. The yield of final pure MerFt is 4 mg per liter cell culture.

Chapter 3: Heteronuclear single quantum coherence spectroscopy experiment on MerFt protein

3.1 Abstract

The solution Heteronuclear single quantum coherence spectroscopy solution NMR spectroscopy was successfully gained with proper conformation of MerFt protein. Comparing to the previously published spectroscopy, the spectrum has good resolution, reasonable line board, correct chemical shifts and number of peaks.

3.2 Introduction of HSQC solution-NMR experiment

Heteronuclear single quantum coherence experiment, normally abbreviated as HSQC, is used in protein NMR spectroscopy.(Bodenhausen & Ruben, 1980) HSQC spectrum is 2-dimensional spectrum. One of the two axis is for proton ($_{1}H$) and another axis is for heteronuclear. In protein NMR, since amino acid was the basic residue which are containing amine (NH_2) and carboxyl (COOH) functional groups, the heteronuclear axis normally is $_{13}C$ or $_{15}N$ to indicate amino acids. For small organic compounds, HSQC can be performed with natural abundance of either $_{13}C$ or $_{15}N$ isotope. However, for protein HSQC, isotopically labelled protein was necessary to gain the residual signal. In $_{1}H$—$_{15}N$ HSQC, the side chain NH_2 peaks
of asparagine and glutamine appear as doublets and a smaller peak may appear on top of each peak because of deuterium exchange in D_2O

containing NMR sample. These special sidechains give the sidechain a distinguish appearance. The backbone amide peaks of glycine also have fingerprint position near the top of the spectrum. HSQC spectra are recorded at a series of time points while the hydrogen is exchanging with the deuterium. The major drawback is that it requires a prior assignment of the spectrum for the protein in question. This method is labor intensive and is limited by size of protein. Usually the protein size is less than 20KDa. Since the molecular weight of MerFt is 6.8 KDa. HSQC is suitable to analysis the protein structure.

3.3 Material and experiment

The HSQC solution NMR sample was prepared in standard protein NMR sample manner. The lyophilized MerFt (~0.5 mg) was dissolved in NMR sample buffer (500 mM SDS, 10 mM sodium phosphate, 90%H_2O/10%D_2O, 500 uL)

For NMR spectroscopy, the NMR spectroscopy was obtained on on a Bruker (www.bruker-biospin.com) DMX 600 MHz spectrometer using a triple-resonance $^{1}H/^{13}C/^{15}N$ probe equipped with three-axis pulsed field gradients. All NMR experiments were performed at 60 ^{o}C using a 1.5 s recycle delay. The chemical shifts are referenced to the $^{1}H_2O$ resonance set to its expected position of 4.3935 ppm at 60 ^{o}C. The standard HSQC experiment was used for isotropic samples with 2048 points in $t2$ and 128 in $t1$. Acquisition time of

F2 (1H) axis 0.154 sec and of F1 axis (15N) is 0.026 sec. Numbers of scans for HPLC second to last peak is 256 scans and it for HPLC last peak is 64 scans, respectively.

3.4 Discussion

Two-dimensional ^{1}H-^{15}N HSQC spectra was used to check the protein behavior for different CNBr cleavage media and purification method. Since the good-resolved quality spectra was obtained by former group member(fig) with detailed presented sample condition(Howell et al., 2005), the spectra was reproducible with the same condition on various purification method MerFt. For MerFt which purified with guanidine hydrochloride and CNBr cleavage in 70% formic acid, from HPLC chromatography from chapter 2, 2 peaks with the similar MerFt protein molecular weight bands were shown in SDS page. HSQC solution spectrometry was applied on these two similar molecular weight proteins resulting two spectrums. Comparing these two spectrums, the protein eluted in last peak properly folded and have better resolution. 57 and 53 peaks were gained respectively from publicized and last peak spectra which means 93% peaks can be reproducible. For MerFt which purified with guanidine hydrochloride and CNBr cleavage in 70% trifluoroacetic acid, FPLC size exclusion purification have two peaks and HSQC spectra were shown in Fig 3.4.1 and Fig 3.4.2 respectively. The last peak protein spectra have better resolution with 58 peaks.

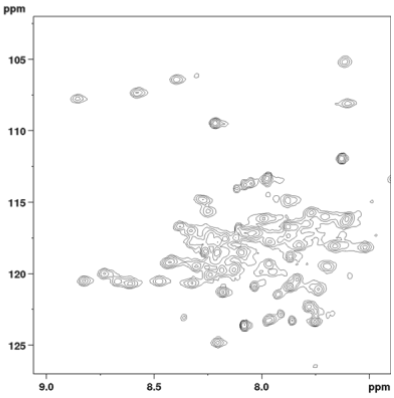

Fig 3.4.1 HSQC spectrum of uniformly 15N-labeled MerFt HPLC second to the last peak

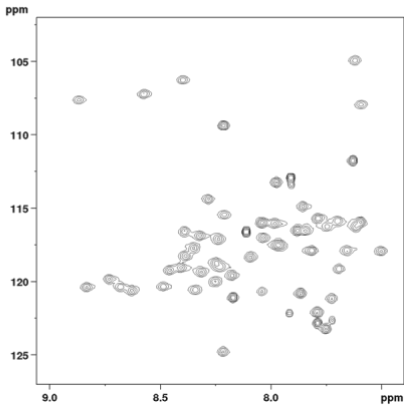

Fig 3.4.2 HSQC spectrum of uniformly 15N-labeled MerFt HPLC last peak

3.3 Conclusion

In conclusion, the HSQC solution NMR spectra can be successfully reproducible with guanidine hydrochloride and SDS purification and 70%TFA, 70% formic acid and 1% SDS CNBr cleavage media. The proper folded and behavior protein is ready for solid state nuclear magnetic resonance spectroscopy. For detergent-free phosphate lipid mrocodisc sample, further optimization should be researched including pH, protein polymer interaction and lipid compositions.

Chapter 4: Solid-state nuclear magnetic resonance spectroscopy on MerFt protein

4.1 Introduction of solid-state nuclear magnetic resonance

Solid-state NMR spectroscopy is a method that fully capable to study structural biology for membrane protein in their biological native lipid environment. During past three decades, this method was largely developed and utilized to do research on molecular dynamic, anisotropy, internuclear distance and other protein structure determination aspects. There are two mainstreams of this approach. One is magic angle spinning solid-state NMR. The other one is oriented-sample solid-state NMR. In this thesis, the later one is the main focus. Oriented sample Solid-state NMR allows three-dimensional protein structural determination with aligned samples. In this chapter, PISEMA, short for polarization inversion spin exchange at the magic angle, was utilized(Marassi & Opella, 2000). This method provides high resolution of the correlation and accurate chemical shift. As the name meaning, PISEMA is a combination of polarization inversion and spin exchange at magic angle (54.74º). Heteronuclear two-dimensional PISEMA spectra provides protein secondary structure because the characteristic wheel-like pattern was observed in the spectra indicate the helix.

4.2 Introduction of solid-state sample preparation

4.2.1 Bicelle sample

Preparation of bilayer protein-containing sample is important for obtain high-quality NMR spectroscopy. Bicelles are mimics of native phospholipid bilayer environment and beneficial for solid-state NMR study. It is a mixture of long chain lipid and short chain lipid which capable align spontaneously in certain lipid ratio, temperature and high magnetic field. The parameter q, molar ratio of long-and short chain lipids, is well-studied. With small q value ($q<0.5$), the isotropic bicelle is not capable to align. In the contrast, with higher q ($q>2.5$), the bicelles form a liquid crystalline phase showing alignment in magnetic field in certain temperature. MerFt aligned phospholipid bilayer solid-state NMR spectrum was obtained and the quality was largely improved. The first reason, the protein purity was largely improved by HPLC purification. Secondly, the instrument is 900 MHz magnetic NMR spectroscopy. The sensitivity of the signal is relatedl to the strength of the magnetic field which contributes to the shape peaks.

4.2.2 Macrodiscs Comprising SMALPs

To provide a detergent-free lipid bilayer environment, styrene-maleic acid (SMA) was introduced in solid-state NMR sample preparation. Styrene-maleic acid (SMA) is a co-polymer with alternating styrene and maleic acid (Figure 4.2.2.1). There are two types of commercially available SMA, 2:1(N:M) and 3:1(N:M), based on the styrene to maleic acid ratio. The polymer is pH-

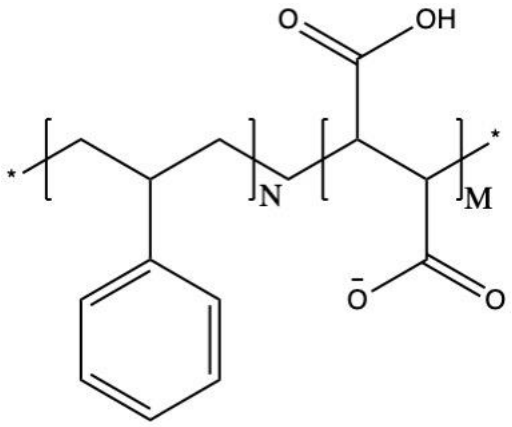

Figure 4.2.2.1 Styrene-maleic acid

sensitive due to the high pKa of the carbonyl groups on styrene groups. The term SMALPs is from the function of SMA as extraction lipid discs from membrane to make SMA lipid particles.(Teo et al., 2019)

4.3 Material and experiments

Bicelle sample was prepared by glass film method(De Angelis & Opella, 2007). Molar ratio of long- and short chain lipids (q) equals to 3.2. The long and short chain lipids were 614-0-PC (DMPC) and 6-O- PC (DHPC) respectively. 3 mg purified MerFt protein was dissolved in 0.5 ml TFE and sonicated for 5 min in bath sonicator. The organic solvent was dried under nitrogen gas. An aqueous solution of DHPC, prepared by addition of 100 uL

Di-H_2O in 9.5 mg DHPC, was added to protein film. The protein film was completely soluble in micelles. An aqueous solution of DMPC, prepared by addition of 80 uL of water to 46.5 mg DMPC, was added in protein micelles followed by vertexing and cold/heat cycles (ice/42$_o$C) until the opaque dispersion turned clear. The pH of the sample was adjusted to 6.0. The sample was packed in short, flat-bottomed tube with 5 mm diameter.

SMALPs sample was prepared with similar method as bicelles sample.

For NMR Experiment, all samples were equilibrated in the magnetic field at 40 °C for 20 min before NMR experiments. The common parameters for both the one-dimensional ^{1}H-^{15}N cross-polarization (CP) and two-dimensional separated local field experiments were 1 ms crosspolarization mix time and radio frequency field strengths of 50-55 kHz for ^{1}H irradiation. The two-dimensional PISEMA and SAMMY spectra of uniformly ^{15}N-labeled samples resulted from a total of 128 t_1 increments and 512 t_2 complex points.

4.4 Conclusion

The 31P spectrum of both bicelle samples and SMALPs indicate that the sample is aligned from 303K to 313K. The spectrum of uniformly labelled $_1$H-$_{15}$N of bicelle sample.

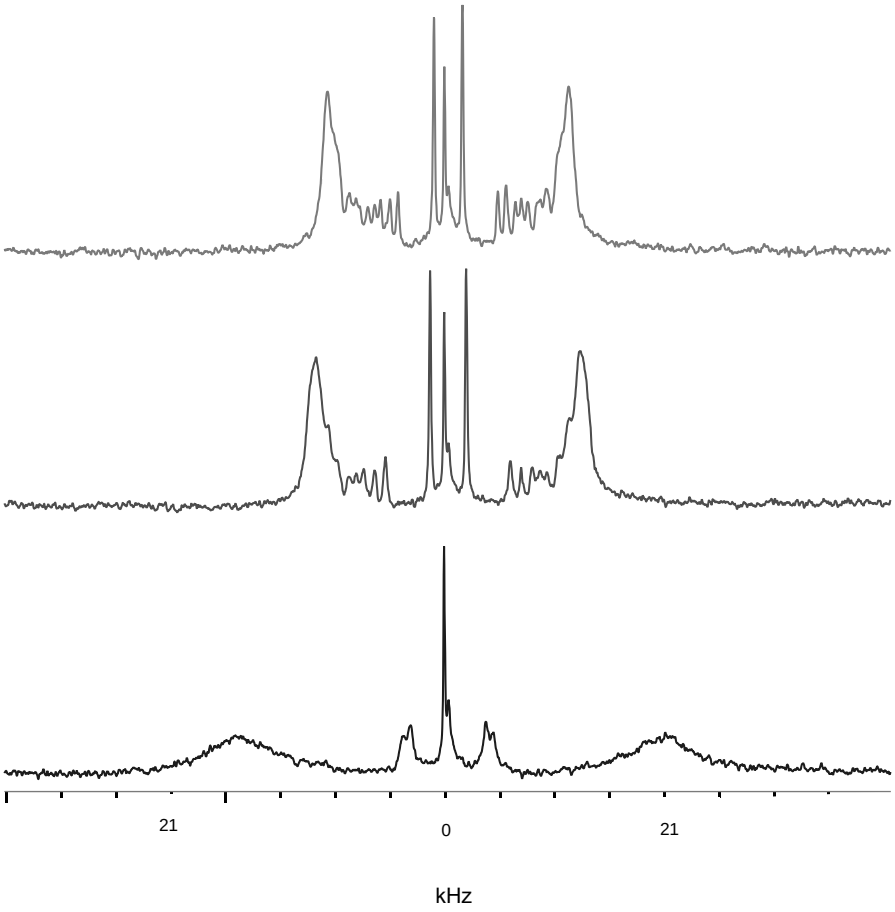

Figure 4.4.1 ^{2}H ns=64, 10% DMPC w/v, 3:1 SMA, Lipid: SMA = 10:3, 50mM HEPES pH 8

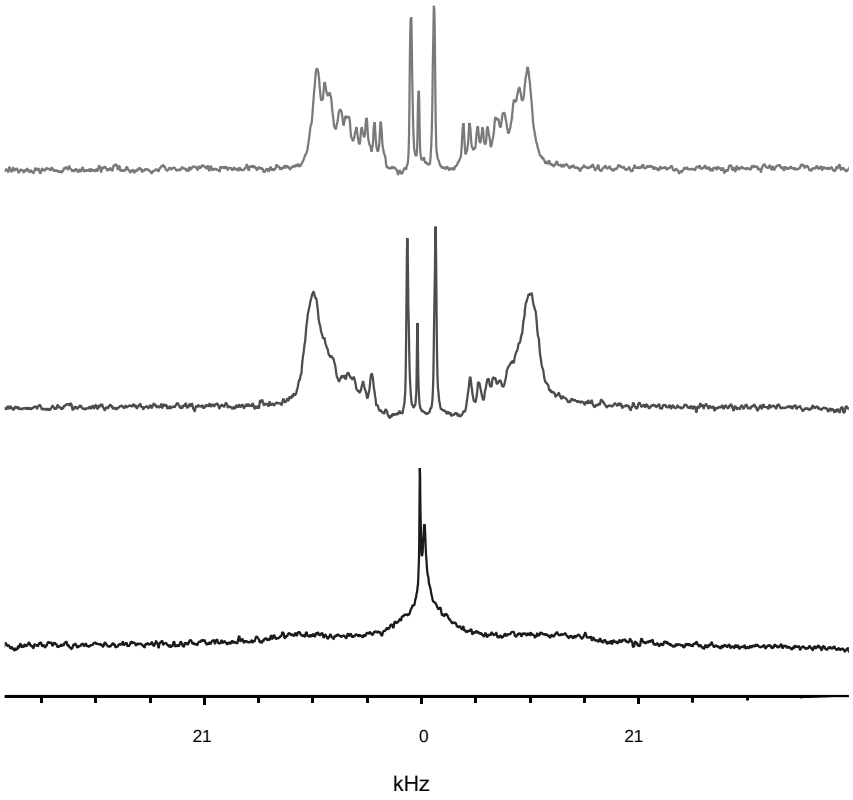

Figure 4.4.2 ^{2}H ns=128, 15N MerFt(2mg), DMPC/DHPC = 3.2, 20% DMPC, pH=5

For 15N NMR spectroscopy, in PISEMA spectrum, 2 wheels clearly shown up indicate 2 helixes. The quality of spectrum is improved compared to previous publication (De Angelis et al., 2006)mainly contributed by two reasons. First, A 900 MHz NMR instrument with a 21.1 T magnet was applied. It largely improves the sensitivity of the spectrum. Second, HPLC purification largely increases the purity of MerFt protein based on hydrophobicity. FPLC was used before but difficult to separate the similar molecular weight spices.

The preliminary 15N solid-state NMR spectroscopy of aligned macrodisc (SMALPs) is obtained (Fig 4.4.3). The sample condition is lipid combination of DMPC/DMPG(1:1 ratio), SMA 3:1, lipid to polymer ratio is 10:3(w/v). The quality of spectrum was not as good as that in bicelle sample according to the resolution and linewidth. Optimizations on SMALP samples are necessary in future research.

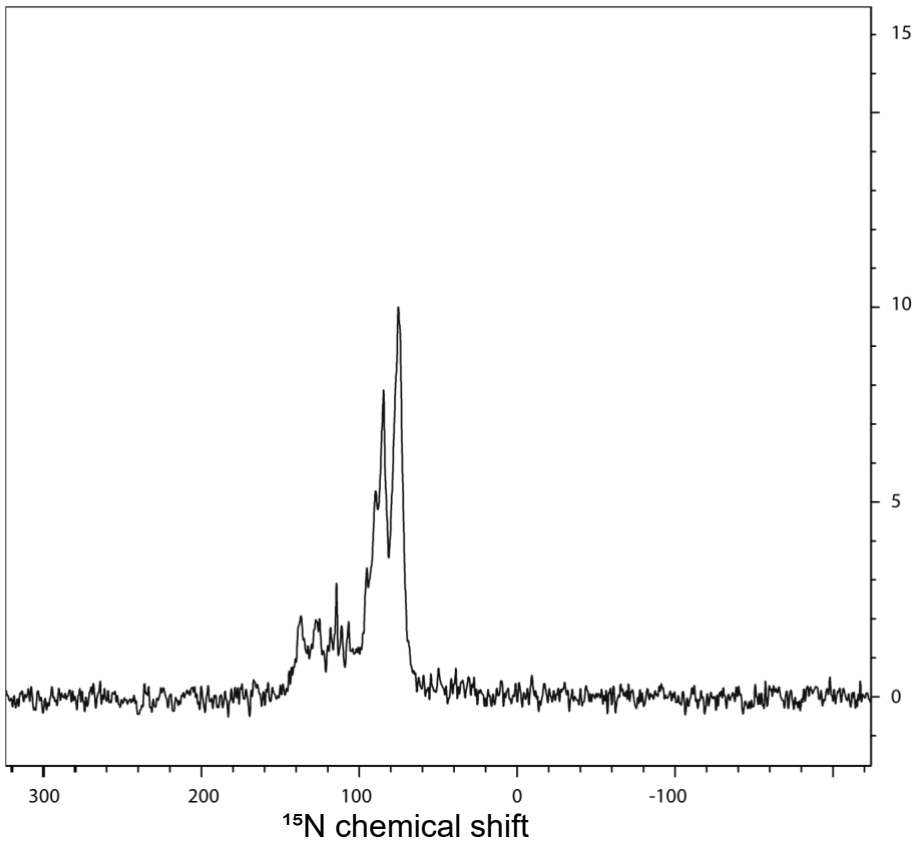

Fig 4.4.3 $_{15}$N solid-state NMR spectroscopy of aligned macrodisc (SMALPs)

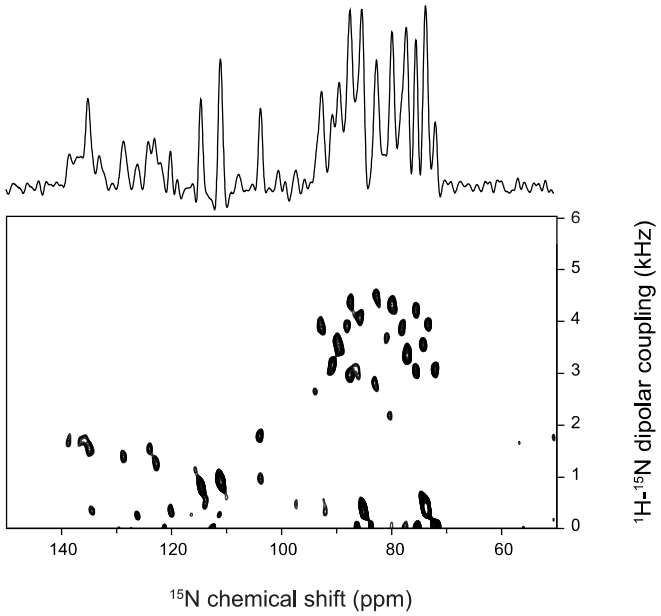

Fig 4.4.4 PISEMA Solid-state NMR of uniformly $_{15}$N-labeled MerFt in lipid bilayer

Chapter 5: Conclusion and future prospective

In conclusion, the MerFt protein is successfully expressed with E coil. (C43) strain and purified with nickel affinity column in SDS. The ^{14}N and ^{15}N uniformly labeled fusion protein yield are 55mg and 20 mg respectively. CNBr is used to cleave fusion protein followed by HPLC. The final yield of ^{14}N and uniformly ^{15}N labeled MerFt are 7 mg and 3 mg respectively. Solution NMR is used to confirm the folding and conformation of MerFt. In solid state NMR, bicelle sample is successfully made and reproduced. Detergent-free solid-state sample preparation was utilized on this protein.

For future plan, improvement on solid-state NMR spectroscopy is the main goal on this mutated and truncated protein. The detergent-free solid-state sample will be fully optimized to study furthermore on other membrane proteins. MerF full length will be expressed and purified as a critical membrane protein in mercury transportation system. Other than MerF, MerA, MerP, MerC and MerT will be used to study the protein binding in this system with biologically natural detergent-free environment.

REFERENCES

Andersen, K. K., Oliveira, C. L., Larsen, K. L., Poulsen, F. M., Callisen, T. H., Westh, P., Otzen, D. (2009). The Role of Decorated SDS Micelles in Sub-CMC Protein Denaturation and Association. *Journal of Molecular Biology*, *391*(1), 207–226. https://doi.org/10.1016/j.jmb.2009.06.019

Bodenhausen, G., & Ruben, D. J. (1980). and David J. RUBEN. *Chemical Physics Letters*, *69*(1), 185–189.

Christie, D. L., & Gagnon, J. (2015). Isolation, characterization and N - terminal sequences of the CNBr-cleavage peptides from human complement Factor B. Localization of a free thiol group and a sequence defining the site cleaved by factor D . *Biochemical Journal*, *201*(3), 555–567. https://doi.org/10.1042/bj2010555

Christoph, R., Schmidt, B., Steinberner, U., Dilla, W., & Karinen, R. (2012). Formic Acid. *Ullmann's Encyclopedia of Industrial Chemistry*, 67–82. https://doi.org/10.1002/14356007.a12

Clewell, D. B., & Helinski, D. R. (1972). Effect of growth conditions on the formation of the relaxation complex of supercoiled ColE1 deoxyribonucleic acid and protein in Escherichia coli. *Journal of Bacteriology*, *110*(3), 1135–1146.

De Angelis, A. A., Howell, S. C., Nevzorov, A. A., & Opella, S. J. (2006). Structure determination of a membrane protein with two trans-membrane helices in aligned phospholipid bicelles by solid-state NMR spectroscopy. *Journal of the American Chemical Society*, *128*(37), 12256–12267. https://doi.org/10.1021/ja063640w

De Angelis, A. A., & Opella, S. J. (2007). Bicelle samples for solid-state nmr of membrane proteins. *Nature Protocols*, *2*(10), 2332–2338. https://doi.org/10.1038/nprot.2007.329

Harada, M. (1995). Minamata disease: Methylmercury poisoning in Japan caused by environmental pollution. *Critical Reviews in Toxicology*, *25*(1), 1–24. https://doi.org/10.3109/10408449509089885

Hayase, F., Kim, S. B., & Kato, H. (1986). Analyses of the chemical structures of melanoidins by13c nmr,13c and15n cp-mas nmr spectrometry. *Agricultural and Biological Chemistry*, *50*(8), 1951–1957. https://doi.org/10.1080/00021369.1986.10867690

Howell, S. C., Mesleh, M. F., & Opella, S. J. (2005). NMR structure determination of a membrane protein with two transmembrane helices in micelles: MerF of the bacterial mercury detoxification system. *Biochemistry*, *44*(13), 5196–5206. https://doi.org/10.1021/bi048095v

Jardetzky, O., & Wade-Jardetzky, N. G. (1980). Comparison of protein structures by high resolution solid state and solution NMR. *FEBS Letters*, *110*(1), 133–135. https://doi.org/10.1016/0014-5793(80)80041-X

Marassi, F. M., & Opella, S. J. (2000). A Solid-State NMR Index of Helical Membrane Protein Structure and Topology. *Journal of Magnetic Resonance*, *144*(1), 150–155. https://doi.org/10.1006/jmre.2000.2035

Mitzi, D.B., Feild, C.A., Harrison, W. T. A., Guloy, A. M. (1994). The X-ray crystal structure of the membrane protein prostaglandin H2 synthase-1. *Nature*, *367*(6463), 532–538. https://doi.org/10.1038/350055a0

Moraes, I., Evans, G., Sanchez-Weatherby, J., Newstead, S., & Stewart, P. D. S. (2014). Membrane protein structure determination - The next generation. *Biochimica et Biophysica Acta - Biomembranes*, *1838*(1 PARTA), 78–87. https://doi.org/10.1016/j.bbamem.2013.07.010

Niemann, M. A., & Mole, J. E. (1982). Characterization of the Cnbr Peptides Generated from the Factor B Cleavage Fragments, Ba and Bb, by Molecular Exclusion High Performance Liquid Chromatography. *Immunological Communications*, *11*(1), 47–58. https://doi.org/10.3109/08820138209050723

Opella, S. J. (2013). Structure Determination of Membrane Proteins in Their Native Phospholipid Bilayer Environment by Rotationally Aligned Solid-State NMR Spectroscopy. *Account of Chemical Research*, *18*(1), 2145–2153. https://doi.org/10.1021/ar400067z

Opella, S. J. (2015). Solid-state NMR and membrane proteins. *Journal of Magnetic Resonance*, *8*, 129–137. https://doi.org/10.1016/j.jmr.2014.11.015

Overington, J. P., Al-lazikani, B., & Hopkins, A. L. (2006). *How many drug targets are there? 5*(December), 993–996. https://doi.org/10.1021/pr100570j

Palmer, I. (2004). Preparation and Extraction of Insoluble (Inclusion-Body)

Proteins from Escherichia coli. *Current Protocols in Protein Science*, *6.3.1-6.3.*, 1–18.

Radoicic, J., Lu, G. J., & Opella, S. J. (2014). NMR structures of membrane proteins in phospholipid bilayers. *Quarterly Reviews of Biophysics*, *47*(03), 249–283. https://doi.org/10.1017/S0033583514000080

Radoicic, J., Park, S. H., & Opella, S. J. (2018). Macrodiscs Comprising SMALPs for Oriented Sample Solid-State NMR Spectroscopy of Membrane Proteins. *Biophysical Journal*, *115*(1), 22–25. https://doi.org/10.1016/j.bpj.2018.05.024

Renaud, J. P., Chari, A., Ciferri, C., Liu, W. T., Rémigy, H. W., Stark, H., & Wiesmann, C. (2018). Cryo-EM in drug discovery: Achievements, limitations and prospects. *Nature Reviews Drug Discovery*, *17*(7), 471–492. https://doi.org/10.1038/nrd.2018.77

Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., & Séraphin, B. (1999). *A generic protein purification method for protein complex characterization and proteome exploration*. *17*(October), 7–9. https://doi.org/10.1038/13732

Ritchie, T. K., Grinkova, Y. V., Bayburt, T. H., Denisov, I. G., Zolnerciks, J. K., Atkins, W. M., & Sligar, S. G. (2009). Chapter 11 Reconstitution of Membrane Proteins in Phospholipid Bilayer Nanodiscs. In *Methods in Enzymology* (Vol. 464). https://doi.org/10.1016/S0076-6879(09)64011-8

Rodríguez, J. C., Wong, L., & Jennings, P. A. (2003). The solvent in CNBr cleavage reactions determines the fragmentation efficiency of ketosteroid isomerase fusion proteins used in the production of recombinant peptides. *Protein Expression and Purification*, *28*(2), 224–231. https://doi.org/10.1016/S1046-5928(02)00700-3

Teo, A. C. K., Lee, S. C., Pollock, N. L., Stroud, Z., Hall, S., Thakker, A., Pitt, A. R, Dafforn. T. R, Spickett. C. M., Roper, D. I. (2019). Analysis of SMALP co-extracted phospholipids shows distinct membrane environments for three classes of bacterial membrane protein. *Scientific Reports*, *9*(1), 1–10. https://doi.org/10.1038/s41598-018-37962-0

Ubarretxena-Belandia, I., & Stokes, D. L. (2010). Membrane protein structure determination by electron crystallography. In *Current Opinion in Structural Biology* (1st ed., Vol. 22). https://doi.org/10.1016/j.sbi.2012.04.003

Yoshii, N., & Okazaki, S. (2006). A molecular dynamics study of surface structure of spherical SDS micelles. *Chemical Physics Letters*, *426*(1-3), 66–70. https://doi.org/10.1016/j.cplett.2006.05.038