Purification Methods for Solid-State NMR Analysis of Influenza A M2[18-60]

KFKM05: Master thesis in Biophysical Chemistry | 2021-06-25

Maxine Gripberg

FMP Head Supervisor: Prof. Dr. Adam Lange
FMP Lab Supervisor: Dr. Sascha Lange

LTH Supervisor: Prof. Daniel Topgaard
LTH Examiner: Prof. Mikael Akke

ABSTRACT

With viral infections becoming an ever-increasing health problem in the world, the study of viral proteins is a cornerstone for understanding and preventing future pandemics. The following work shows the purification and structural characterization of the Influenza A M2 proton channel through solid-state NMR (ssNMR). Despite previously published structures of the M2 proton channel, the method of purification has not been well detailed and requires further optimization in order to produce samples to study at Leibniz-Forschungsinstitut für Molekulare Pharmakologie (FMP). A truncated version of M2 containing residues 18-60 was studied in this work, which is referred to as M2[18-60]. Purification of labeled and unlabeled M2[18-60] expressed at 22°C and 30°C degrees were performed with varied results. It was found that the expression of M2[18-60] produced the highest yield when expressed in M9 minimal media and that the temperature during expression has no significant impact on the purification results. In addition, there was no difference in purification yield of M2[18-60] using a C3 or C4 column. Both labeled and unlabeled M2[18-60] was able to be reconstituted into liposomes. The acquired 2D ^{13}C-^{13}C correlation spectrum confirmed that presence of M2[18-60] generated by the previous purification steps. The stability of the M2 protein was confirmed through overlay of previously published assigned spectra of a M2 S31N mutant. Further optimizations of the methods used in this work will be important in order to produce samples of M2[18-60] for future investigations into the design of antiviral drug candidates, ssNMR analysis of binding sites of M2-targeting drugs and assay development.

ACKNOWLEDGEMENTS

First of all, I would like to thank my supervisor at FMP, Prof. Dr. Adam Lange, for giving me the amazing opportunity to write my master thesis within his group. I would also like to thank all of my colleagues at FMP for their help during this time. Thank you, Kitty, Carl, Pascal and Swantje for helping me understand solid-state NMR and how to work with the different programs. Thank you, Henry for giving me so much support and guidance in my work. Thank you, Susanne and Dagmar, for showing me the ins and outs of the lab and helping me practice my German. Special thanks to Sascha for putting so much time and effort into this work and for helping me in the lab. In addition, I would like to thank my supervisor from LTH, Daniel Toppgard and examiner Mikael Akke, for the guidance they have given me. Finally, I wish to thank my friends and family for supporting me during this time, I am extremely thankful for you all.

TABLE OF CONTENT

LIST OF ABREVIATIONS

NMR	Nuclear magnetic resonance
ssNMR	Solid-state NMR
Mg	Membrane glycoprotein
MgM	Membrane glycoprotein M
DV	Dengue Virus
WNV	West Nile Virus
M2	Matrix 2
MBP	Maltose-binding protein
RF	Radiofrequency
FID	Free induction decay
TMS	Tetramethylsilane
DDS	Sodium salt of 2,2-dimethyl-2-silapentane-5-sulphonic acid
SNR	Signal-to-noise ratio
MAS	Magic angle spinning
PDSD	Proton-Driven Spin Diffusion
HA	Hemagglutinin
TM	Transmembrane
WT	Wild-type
IMAC	Immobilized metal affinity chromatography
RPC	Reversed-Phase Liquid Chromatography
CNBr	Cyanogen bromide
FA	Formic Acid
GRAVY	Grand average of hydropathicity
HPLC	High performance liquid chromatography
AU	Absorbance units
POPC	1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine
DPhPC	1,2-diphytanoyl-sn glycerol-3-phosphocholine
SDS-PAGE	Sodium dodecyl sulfate-plyacrylamide gel electrophoresis

1 INTRODUCTION

1.1 Project Motivation

Viral infections have become arguably the largest health problem in the world. The Covid-19 pandemic has highlighted a desperate need for the development of antiviral drugs and vaccinations. However, viral infections have always been an issue for mankind. Historically, the "Spanish flu", "Asian flu", and "Hong Kong flu" claimed the lives of millions of people all over the world in the early and late 20th century. More recently, SARS, MERS, and influenza A virus have been the cause of several outbreaks around the world, further supporting the importance of viral research and prevention.

The Lange group at the department of Molecular Biophysics at Leibniz Research Institute for Molecular Pharmacology (FMP) focusses on the biophysical characterization of membrane protein structure and dynamics, through the use of solid-state NMR (ssNMR) (1). This method allows proteins to be studied under close to native-like conditions. One of the newest projects within the group has been on viral proteins with ion channel activity that reside within the viral envelope, termed viroporins (2). Viroporins are a potential drug target due to their ion transport which is usually vital for viral infection and replication. Examples of viroporins are the membrane glycoprotein (Mg) of Dengue virus (DV), MgM of West Nile virus (WNV), and the Matrix 2 (M2) proton channel of influenza A virus (3). The M2 proton channel in particular plays a vital role in the infection of human cells by influenza A. Previously, amantadine has been used as an antiviral, due to it's ability to inhibit M2 activity, but this has been discontinued due to viral resistance through residue mutation in the M2 transmembrane region (4). One way for researchers to develop novel antiviral drugs for a protein target is to understand the structure and function of the protein. The structure of the protein allows researchers to synthesize possible drug candidates and knowledge of the function of the protein allows correct assays to be performed to evaluate the effectiveness of the drug. To study the structure of the protein, the protein first needs to be expressed and purified (5). This thesis will mainly cover the challenges of purifying M2 and the subsequent reconstitution into liposomes in order to study its structure using ssNMR. In addition, the importance of M2 as an antiviral drug target will be discussed.

1.2 Previous work at FMP

Previous work on protein expression and purification of the Mg and MgM proteins of DV and WNV at FMP was performed according to the method used in the article by Gan *et al.* (6). The proteins were expressed using a vector with a maltose-binding protein (MBP) fusion partner with a His_6 tag at the N-terminus. After affinity purification, Mg and MgM could be cleaved from the fusion partner using *Tobacco Etch Virus* (TEV) protease that cleaves at Gln/Ser in the amino acid sequence ENLYFQS (7). The complete constructs for the Mg and MgM proteins of DV and WNV can be seen in Figure 1.

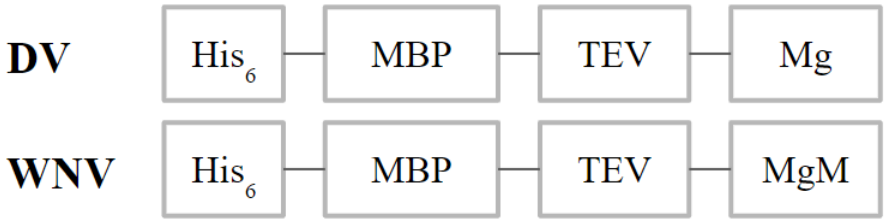

Figure 1: Constructs of Mg and MgM proteins of DV (top) and WNV (bottom). The constructs contain a His_6 tag, MBP fusion partner, and a TEV cleavage site.

The method yielded no detectable expression of the Mg protein of DV but was successful for the MgM protein of WNV. Since the method by Gan *et al.* (6) yielded insufficient results, the protocol by Fu *et al.* (8) was used to produce a sample of the M2 proton channel of Influenza A that could be analyzed using ssNMR.

1.3 Aim

To be able to study the M2 proton channel of influenza A using ssNMR, the protein requires isotopic labelling and reconstitution into liposomes. For successful reconstitution of M2 into liposomes, the isotopically labeled protein first needs to be purified after expression, which involves several important steps. There are a couple of publications that briefly mention the expression of M2 (9, 10). However, these publications do not explicitly detail the purification process, therefore, for application at the FMP, additional optimization steps are required. In previously mentioned publications, different temperatures of expression (22°C and 30°C) were used as well as different columns (C3 and C4) for liquid chromatography. The primary goal for this thesis was therefore to analyze and evaluate the different purification methods and subsequent ssNMR results. In addition, the importance of M2 as an antiviral drug target as well as possible future investigations that can be made regarding drug discovery and potential assay development was discussed.

Limitations

This project was limited to the expression, purification, and ssNMR analysis of M2[18-60]. Therefore, assay development and drug candidate synthesis were not performed. In addition, because the purification of M2[18-60] took longer time than expected, the ssNMR analysis was limited and was only broadly analyzed by overlaying previously assigned peaks to confirm the presence of M2[18-60] and the success of the purification. The purification methods were limited to methods previously published regarding the purification of transmembrane domains and M2[18-60].

2 INFLUENZA A M2 PROTON CHANNEL

2.1 Influenza A

Influenza A is an enveloped, negative-sense RNA virus within the Orthomyxoviridae family (11). Influenza A viruses are responsible for causing the seasonal flu and have also been the cause of several historic pandemics (12). The virus mutates frequently, producing new strains that can be transmissible between animals and humans and can become more lethal as it mutates. Currently, vaccines have been the primary method to prevent viral infections. Vaccines that are designed to give immunity against the seasonal flu need to be produced every year by predicting which strains will be most prevalent that year. However, this strategy cannot be used for novel viruses since it is impossible to predict what mutations may occur in novel variants. In order to prevent or slow down influenza outbreaks, antiviral drugs can be used before a vaccine has been produced. Antiviral drugs can inhibit a protein that is important for the viral life-cycle or the virus's ability to infect cells. For example, amantadine was used to inhibit the M2 proton channel of influenza A viruses but was discontinued due to resistance. Due to the limitations in vaccine production, and the increase in anti-viral resistance in recent years, the development of novel ways to inhibit influenza A viruses is of huge importance to prevent future pandemics (5).

2.2 M2 proton channel

One of the proteins within the lipid envelope of the influenza A virus is the M2 proton channel. The channel is pH-gated, which means that the open and closed conformation is pH-dependent. The unidirectional proton conduction of the M2 channels is responsible for several vital steps in the viral lifecycle. For example, during viral infection, the pH (~5) of the endosomal epithelial cells of the respiratory tract activates the M2 channel which lowers the pH within the viral envelope. This mediates the fusion of the viral envelope to the endosomal membrane and the subsequent release of ribonucleoprotein. After viral replication within the infected cell, M2 is transported to the cell surface by the trans-Golgi network. During the transportation, M2 regulates the pH at the Golgi lumen, preventing surrounding hemagglutinins (HA) to undergo conformational rearrangement before reaching the plasma membrane of the cell (11). Therefore, M2 is important for both the release of the viral genome and viral replication.

2.3 Structure

The M2 channel is composed of four identical transmembrane (TM) helices that together form a homotetramer. In total, each helix is comprised of 97 residues. Each helix can be divided into three domains: the extracellular N-terminal domain, TM domain, and the intracellular C-terminal domain (11). In order to study the structure and mechanism of the M2 proton channel, the truncated construct M2[18-60] has been produced which has been shown to form a stable tetramer within a lipid bilayer (13). Recent work on the M2[18-60] has shown that the protein is composed of a dimer of dimers, meaning that two pairs of helices together form the tetramer. Drug sensitivity and proton conduction have also been shown to be preserved within this region which indicates that the truncated M2[18-60] functions similarly to the wild-type

(WT) protein (9). A schematic representation of M2[18-60] and its orientation within a lipid bilayer can be seen in Figure 2.

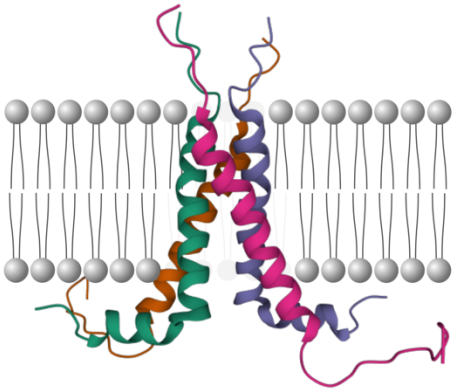

Figure 2: Schematic representation of M2[18-60] within a lipid bilayer. The hydrophilic head is represented by a sphere and hydrophobic tails by lines. The M2[18-60] is an S31N mutant, which denotes that amino acid 31 (Serine, S) is changed to a asparagine (N) (14) (PDB: 2N70).

Within the M2[18-60] construct, residues 18-23 form an unstructured N-terminus, residues 25-46 form the TM channel, residues 47-50 form a short flexible loop, and the C-terminal residues 51-59 form an amphipathic helix (10). A schematic representation of the aforementioned functional residues within the M2[18-60] construct can be seen in Figure 3.

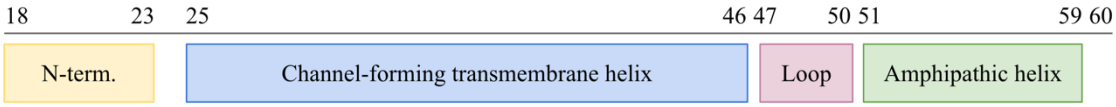

Figure 3: Schematic representation of functional residues within the M2[18-60] construct. N-term. are residues 18-23 that form an unstructured N-terminus, residues 25-46 form the channel pore, residues 47-50 form a short flexible loop, and residues 47-59 form an amphipathic helix.

2.4 Proton conduction

The exact mechanism of proton conductance through the channel has been of much debate. One theory is that protons are able to be conducted through the M2 channel via diffusion and traveling down their electrochemical gradient (15). The conductance of protons through the channel is theorized to be mediated by hydrogen bonds with water and consequently largely dependent on polar residues that facilitate channel hydration within and at the terminus of the channel. Another theory is that imidazole might play a role in relaying protons through the channel via binding and dissociation (11).

2.5 Activation

Activation (opening) of the M2 proton channel is dependent on the pH environment at the N-terminal ectodomain. The activation mechanism is dependent on His37, which acts both as a pH sensor that activates the channels proton conduction as well as being proton-selective (16). When the pH is around 5.6 (17), His37 becomes protonated which results in a rotation of the indole on Trp41. The rotation allows for protons to flow through the channel. Therefore, Trp41

acts like a gate, blocking and opening the channel as a result of the protonation of His37. At a higher pH of 7.5, the His37 is no longer protonated and therefore the pore is closed (16).

2.6 M2 clinical importance

There are several reasons why M2 would be a valid target for antiviral drugs (5). Firstly, the M2 proton channel is largely conserved compared to other mutations of influenza A proteins. This is because many of the residues found within the channel-forming TM helix are functionally important for proton conduction (5). In particular, His37 and Trp41 have been shown to be critical for channel activity (18). Secondly, there are many antiviral drugs that do not target the M2 channel which could allow for combinational therapy or a M2 inhibitor to be used in case of resistance to antiviral drugs that do not target M2. In addition, when investigating in vitro mutations that occur on M2 under antiviral conditions, mutants have been found to revert to WT configurations after removal of the antiviral drug (5). Due to its importance in the viral life cycle and largely conserved regions, the research into M2 inhibitors is certainly of clinical importance.

2.7 Mutations and drug resistance

The first antiviral drug that was shown to target the M2 proton channel was amantadine. In Figure 4, the structure of amantadine can be seen, with its hydrophobic adamantane cage and an amine group.

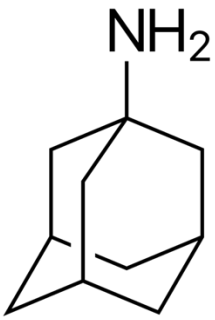

Figure 4: Structure of amantadine.

The drug was discovered during the late 20th century but due to amantadine-resistant mutants emerging in the early 21st century, the use of amantadine was discontinued. It was found that the most dominant mutation which confers drug resistance towards amantadine was that of S31N. Interestingly, it was also found that this mutation existed even before the use of amantadine, and it was concluded that the S31N mutation was not due to selective pressure, but due to spontaneous mutation that rendered a functional similarity to WT-M2. Later studies have also found that only two more mutations (V27A and L26F) would be able to function similarly to WT-M2. Another aspect that might prevent the emergence of strains resistant to M2-targeting drugs is that HA is highly dependent on the conductivity of M2. The two proteins co-evolve, which means that mutations in HA need to confer with a mutation in M2 and vice versa. The likelihood of mutations occurring of HA and M2 causing resistance, while still retaining the proteins' functionality, is highly unlikely (5).

2.8 Drug discovery process

The general process of drug discovery involves the synthesis of potential drug candidates, target identification, and screening methods *in vitro* and *in vivo*. Several screening methods of synthesized drug candidates have been performed to discover novel M2 inhibitor. Initially, viral replication and high-throughput screening assays can give an indication of whether the candidate is antiviral, however, these methods do not indicate whether the candidate is targeting M2 or other proteins in the virus. To determine if the candidate is targeting M2, liposomal flux assays, electrophysiological assays, and bacterial or yeast growth assays are required. As a means to specifically determine the binding of the potential candidate to M2, binding assays using NMR and fluorescence would need to be performed. NMR would, in particular, provide structural and functional information regarding the binding of the candidate to M2 (5). That is why the expression, purification, and study of M2 using ssNMR can lay the foundation of the drug discovery process.

3 LABORATORY THEORY

3.1 Introduction

The following section describes the theoretical background of the expression, purification and reconstitution of M2[18-60]. There are several publications that briefly mention the expression and purification methods of M2[18-60]. Since the publications don't describe the method of purification in detail, the protocol by Fu *et al.* (8) was used as a starting point and changes were made sequentially to determine the impact on M2 yield and purity.

3.2 Construct and expression

The plasmid containing M2[18-60] as a fusion to a C-terminal His_6-trpLE in a pMM-LR6 vector (8) was a generous gift from James Chou from the department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School (10). The plasmid was transformed in *E. Coli*. The complete sequence of the construct can be seen in Figure 5.

MHHHHHHHHHKAIFVLKGSLDRDLDSRIELELRTDHKELSEHLLLVDLARNDLA
RIATPGSRYVADLTKVDRYSYVLHLVSRVVGELRHDLDALHAYRAALNLGTLSG
APKVRA*KL*WMRSNDSSDPLVVAASIIGILHLILWILDRLFFKSIYRFFEHGLK

Figure 5: Sequence obtained from James Chou. The M2[18-60] sequence is highlighted in green with a fusion to a C-terminal His_6 tag, trpLE underlined, KL is the restriction site for HindIII, and M (methionine) in red is where CNBr cleavage occurs.

3.3 Inclusion Body Purification

Membrane proteins are challenging to express in bacteria due to their inherent lethality to the host cell (19). Therefore, M2[18-60] is expressed together with the fusion protein trpLE which drives the formation of inclusion bodies of overexpressed proteins. Inclusion bodies can be isolated by lysing the cells followed by medium speed centrifugation. To purify the protein further, isolated inclusion bodies are solubilized by addition of high concentration of a denaturing agent (e.g. 8M Urea or 6M Guanidinium Hydrochloride) and subjected to immobilized metal affinity chromatography (IMAC). This method utilizes the association of an affinity tag, such as His_6, to immobilized beads, such as Ni^{2+} or Co^{2+}. For this, a C-terminal His_6 tag has also been included (See Figure 6) with the trpLE-M2[18-60] construct. Low affinity proteins, such as those without a poly-histadine tag, passes through the column. Tagged proteins, such as M2, bind tightly to the immobilised Ni^{2+} beads. The bound protein can then be eluted from the chromatography column using imidazole, or acidic pH. The components of interest in the pMM-LR6 vector can be seen in Figure 6.

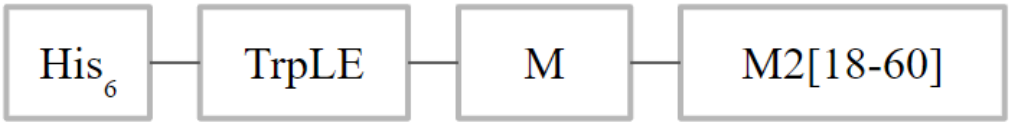

Figure 6: Proteins being expressed in the pMM-LR6 vector. A C-terminal His_6 tag is followed by trpLE with a M (methionine) connecting to M2[18-60].

As a means to cleave off His_6-trpLE and isolate M2[18-60], cyanogen bromide (CNBr) is used in the presence of formic acid (FA). The reaction with CNBr leads to a cleavage at the C-terminal of any methionines present in the protein (20). In previously published papers, the CNBr reaction has been performed in 90% FA for 1 h (8), 70% FA for 2 h (10), and 70% FA for 1-3 h (9) under a continuous nitrogen gas stream. The method performed in this thesis used 85% FA for 1 h. At a later stage, an apparatus was made in order for the nitrogen gas stream to flow through the sample. To stop the reaction, CNBr is removed through dialysis against H_2O. This also reduces the concentration of FA in the sample. The sample is then lyophilized before being further purified using reversed-phase liquid chromatography (8).

The size and grand average of hydropathicity (GRAVY) score of expressed protein and cleavage products can be seen in Table 1. Values were obtained from the ProtParam tool using the amino acid sequences of each component (21).

Table 1: Size and GRAVY score of cleavage products of the construct.

Peptide	Size (kDa)	GRAVY score
M2[18-60] monomer	5.1	0.566
M2[18-60] dimer	10.2	*
M2[18-60] tetramer	20.4	*
His_6 -trpLE	13.5	-0.353
His_6 -trpLE-M2[18-60] monomer	18.6	-0.102
His_6 -trpLE-M2[18-60] dimer	37.2	*
His_6 -trpLE-M2[18-60] tetramer	74.2	*

The GRAVY score is a measure of the hydrophobicity of proteins which depends on the hydropathy of amino acids and the length of the sequence. Positive GRAVY scores indicate that the protein is hydrophobic, while negative GRAVY scores are indicative for hydrophilic proteins (22).

**The GRAVY score of a dimer or tetramer would be the same as the monomer according to how GRAVY is calculated, but since the hydrophobic amino acids are exposed to the outside of the protein, it may appear more hydrophobic.*

3.4 Reversed-Phase Chromatography

Reversed-Phase Chromatography (RPC) can be used to separate proteins in a liquid sample. In general chromatography, the stationary phase that resides in a column interact with proteins in the mobile (liquid) phase being passed through the column. The interaction between the stationary phase and the proteins in the mobile phase causes the proteins to separate. In RPC, the stationary phase is non-polar (and therefore hydrophobic), and the mobile phase is polar (hydrophilic). A gradient elution is applied in order for the proteins of the mobile phase to be eluted (23). The separation of the proteins in the mobile phase, therefore, depends on the hydrophobicity of each protein. Hydrophobic proteins will associate more strongly to the non-polar stationary phase, while hydrophilic proteins will associate less, eluting at a lower elution volume. Elution fractions are collected at a determined volume and protein content can be

monitored through UV-Vis absorbance at 280 and 214 nm (24). These absorbances are important for protein detections since peptide bonds absorb at 214 nm and aromatic amino acid residues absorb at 280 nm (25). The results are documented on a graph containing information about the elution fraction and Absorbance Units (AU). Each peak theoretically represents an individual protein as each protein will elute at a different volume depending on their hydrophobicity. However, it is important to note that different proteins can have identical hydrophobicity, which means that they would elute at the same volume. The distance between the start of the graph and the peaks is called the retention time. Proteins with a short retention time have not been associated readily to the column and will elute quickly, while proteins with a longer retention time are associated to the column and will elute more slowly (24).

Columns

M2[18-60] samples were purified using either a Vydec TP C4 column (26) with a pore size of 300 Å, carbon load of 3%, an inner diameter of 22 mm, and a length of 250 mm or a ZORBAX C3 column (27) with a pore size of 300 Å, carbon load of 4%, an inner diameter of 9.4 mm, and a length of 250 mm. The C3 and C4 denominators indicate the number of carbons bound to the silica-based stationary phase. C4 has a slightly higher hydrophobicity while C3 has a slightly lower hydrophobicity (28).

3.5 Reconstitution

With the interest of studying M2[18-60] in as close to its native-like environment as possible, the protein needs to be reconstituted into liposomes. Liposomes are composed of a phospholipid bilayer that forms a sphere. The process of reconstituting M2[18-60] into liposomes initially requires detergent-solubilized M2[18-60]. When the detergent is removed, M2[18-60] spontaneously incorporates into the phospholipid bilayer. A liposome with an incorporated protein is then called a proteoliposome (29). In previous studies the phospholipids 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC) and 1,2-diphytanoyl-*sn*-glycero-3-phosphocholine (DPhPC) have been used, and shown to generate a stable tetramer of M2[18-60] (30).

3.6 Gel Electrophoresis

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) can be used to determine the molecular weight of proteins. The method involves the sample being placed in a polyacrylamide gel together with an electrophoresis buffer. The addition of SDS to the sample before electrophoresis mitigates the proteins' natural charges and gives them a negative charge that will correlate to the mass of the protein at the same ratio for all proteins in the sample. The molecular weight of the bands generated can be elucidated through use of a pre-stained protein markers (known as a ladder). This ladder is run concurrently with the samples. Voltage is then applied which causes the negatively charged proteins in the sample to migrate through the gel towards the positive anode. The migration speed and therefore the separation of the proteins is almost entirely dependent on the molecular weight of the protein, finally leading to a separation of the proteins depending on their molecular weight. The gel is then placed in a staining solution which allows the bands of the proteins to be seen. One band on the gel represents proteins with the same molecular weight or migration speed, and by comparing it to the protein ladder, the molecular weight of the protein can be determined

(31). Gel electrophoresis can also be performed with other gels such as Mini-PROTEAN TGX™ gel (32) together with a Tris-Tricine buffer (33) which allows for further band separation.

3.7 Mass Spectrometry

Mass spectrometry is a technique that can be used to analyze peptides, proteins, small molecules and even metabolites. The technique works by bombarding samples with electrons through the use of a high voltage. These electrons cause the sample to become negatively charged (depending on the number of electrons absorbed). The abundance of each ion is later detected, which can be shown by plotting intensity *vs* the mass/charge (m/z) ratio (34). If the m/z ratio of the peptide that is being analyzed is known, selected ion monitoring (SIM) can be performed instead. The ion is then selectively detected during a retention time which can be plotted as the intensity or relative abundance *vs* the retention time (35). Samples for mass spectrometry analysis can be acquired from SDS-PAGE by cutting out the band of interest from the gel (36). The use of in-gel enzymatic digestion using trypsin is commonly used, which cleaves at arginine and lysine amino acid residues (37).

4 NMR THEORY

4.1 NMR Spectrometry

Nuclear Magnetic Resonance (NMR) spectroscopy is a method that can be used to obtain structural and dynamic information of proteins that cannot be crystallized (38). NMR spectroscopy is dependent on the spin of nuclei and their interaction with magnetic fields (39). The following section will describe some underlying quantum mechanical processes that NMR spectroscopy is based on and what can be learned from an NMR spectrum.

Spin and quantum numbers

To understand the fundamentals of NMR spectroscopy, the concept of *spin* needs to be introduced. Spin is an intrinsic property of charged nuclei, like mass or charge, where the nucleus *appears* to 'spin' around an axis. This generates a magnetic dipole moment. Examples of nuclei that have a nuclear spin quantum number of ½ are ^{1}H, ^{13}C, and ^{15}N, which are termed NMR-active. When these nuclei are placed in an external magnetic field, B_0, they can become aligned with B_0 and absorb radiofrequency (RF) radiation (40). When nuclei are placed in B_0, they can then have different energy levels (spin states) and the energy difference between these levels is the same energy as the RF radiation.

Precession

Another concept worth mentioning is precession. This is the effect that an external magnetic field has on the nuclear magnetic moment. When B_0 is applied, the nucleus will start to process about the magnetic field at a specific frequency known as the Larmor frequency, ω_0.

Bulk magnetization

In reality, when a protein or molecule is placed in a magnetic field, not all nuclei will be aligned with or against the magnetic field. The magnetic moment of all the nuclei are oriented randomly at first, but after a short amount of time the magnetization *relaxes* and the system reaches equilibrium. This is when the net magnetization of all individual magnetic moments results in a *bulk magnetization, M_0,* aligned with B_0, as seen in Figure 7. This is very small relative to B_0 and is utilized for the detection of NMR signals (39).

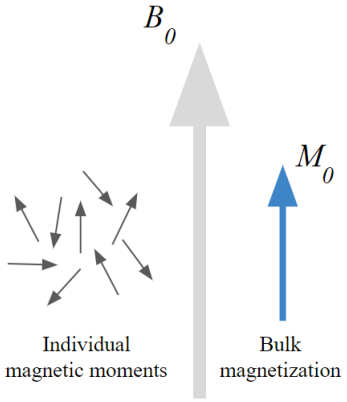

Figure 7: Illustration of the bulk magnetization M_0 aligning with the external magnetic field B_0. *Reproduced image from Keeler, 2010. (39)*

The Vector Model

It is preferable to describe what is happening with the magnetization during an NMR experiment using a vector model. In a 3D coordinate system with the x-, y-, and z-axis arranged as in Figure 8, the bulk magnetization, M_0, of the sample will be pointing along the +z direction when B_0, is applied along the +z-axis (39).

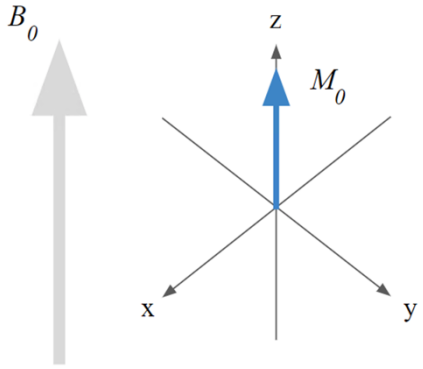

Figure 8: Illustration of a 3D coordinate system with the bulk magnetization, M_0, aligned with the external magnetic field, B_0, that is applied to the +z-axis. *Reproduced image from Keeler, 2010. (39)*

Spectral acquisition

In the simplest NMR experiment, an RF pulse that oscillates at the Larmor frequency is applied on the x-axis which causes the bulk magnetization to rotate 90° into the xy-plane. This RF pulse has the same energy as the difference between the two spin states. The exact frequency and pulse length for this 90° rotation is dependent on the nuclei of interest (41). The power levels (RF power) and duration of the pulse usually used for ssNMR are described in Table 2 (42).

Table 2: The recommended power level and duration of the 90° pulse for ^{1}H, ^{13}C, and ^{15}N.

Nuclei	**^{1}H**	**^{13}C**	**^{15}N**
Power Level (kHz)	83.3	50.0	35.7
Duration (µs)	3	5	7

As the nuclei bulk magnetization is *flipped* to the xy-plane, the nuclei will keep processing at the Larmor frequency in the xy-plane and eventually return to process along the +z-axis. This process can be seen in Figure 9.

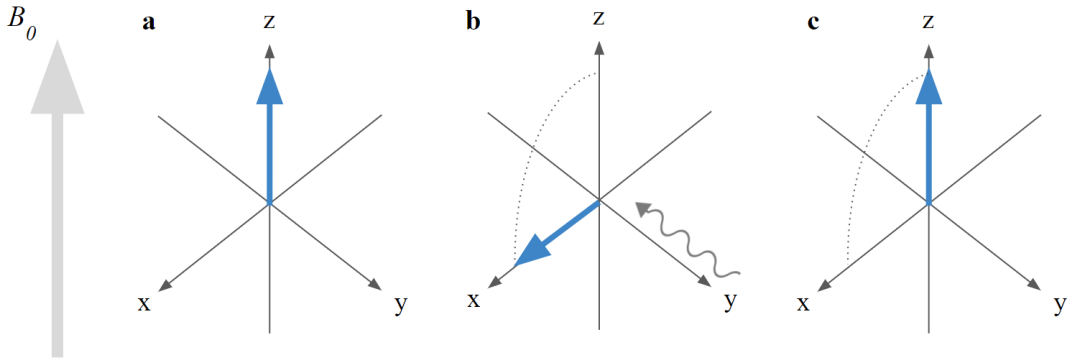

Figure 9: Illustration of the 90° pulse. The bulk magnetization of the nuclei if flipped to the xy-plane as a result of an RF pulse oscillating at the Larmor frequency applied on the y-axis. *Reproduced image from Keeler, 2010. (39)*

As the bulk magnetization returns to the +z-axis, it induces a current in a receiver coil that is wound around the sample cavity. The *free induction decay* (FID) is the signal that can be detected as the magnetization relaxes back to its equilibrium state along the +z-axis. The FID is in the time-domain and the mathematical operation called the Fourier transform needs to be applied to convert it into the frequency domain (39).

Determining the 90° pulse

To acquire maximum intensity, the power level of the 90° pulse needs to be determined. This can be done by looking for the null condition, the pulse required to flip the magnetization 180° and thereby acquiring no signal (39). By determining the 180° degree pulse (at a given power level), the 90° degree pulse can be determined by simply halving the pulse length (42).

Chemical Shifts

In a typical NMR spectrum, a frequency can be associated with a specific atom within the molecule. The height of the peaks are the intensities, and the location of the frequency on the x-axis refers to the chemical shift of the nucleus. The chemical shift, δ, is measured in parts per million (ppm) and refers to the frequency at which the nucleus resonates in relation to a reference compound, typically tetramethylsilane (TMS) or sodium salt of 2,2-dimethyl-2-silapentane-5-sulphonic acid (DDS). This enables NMR spectra from spectrometers with different magnetic field strengths to be compared (39). The x-axis of an NMR spectrum is by tradition labeled with 0 ppm (the reference compound) at the far right and increases to the left. The chemical shifts to the right are considered upfield and shielded (decrease in ppm) and the chemical shifts to the left are considered downfield and deshielded (increase in ppm). The chemical shift depends on the local electron environment that each nucleus experiences. The local electron environment will determine how much of the external magnetic field the nucleus experiences and therefore affects the possible spin states of the nucleus. The electrons surrounding a nucleus are specific for the nucleus and how it is bound to other nuclei. For example, in ^{1}H (proton) NMR, the chemical shift of the protons depends on electronegative atoms and unsaturated groups surrounding it. Protons bound to electronegative atoms will have a lower electron density around the nucleus, and they are therefore termed deshielded. By knowing the typical chemical shifts of protons in specific electron environments, the peaks of the NMR spectrum can be assigned to protons of the molecule being analyzed (43). For ^{1}H NMR the chemical shift generally ranges between 12 and 0 ppm, ^{13}C NMR ranges from 220 to

0, and ^{15}N ranges from 400 to -200 ppm (44). Typical chemical shifts of protons and carbons depending on their electron environment can be seen in Figure 10.

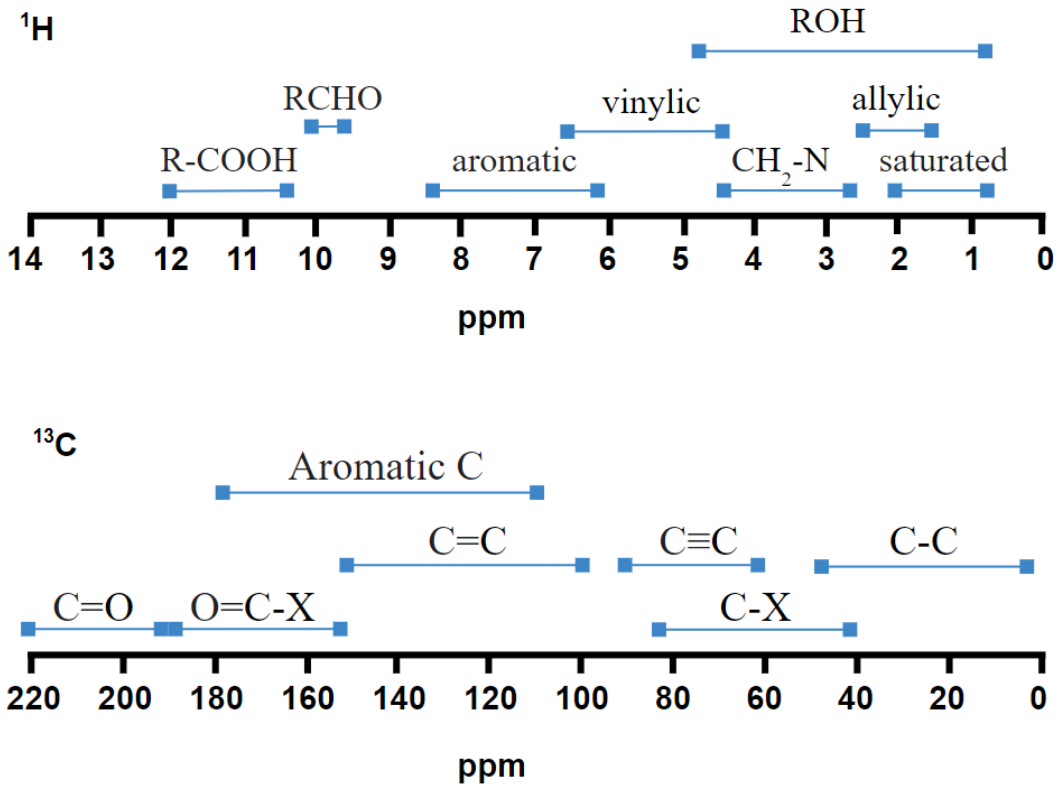

Figure 10: Typical chemical shifts of ^{1}H and ^{13}C depending on their electron environment. The ^{1}H chemical shift ranges between 14 and 0 ppm (top), and the ^{13}C chemical shift ranges between 220 and 0 ppm (bottom)

Decoupling

In ^{13}C NMR spectroscopy, decoupling is used to improve the signal-to-noise ratio (SNR). The process involves an RF pulse being applied to protons during acquisition and relaxation to remove the splitting caused by them. This method results in less splitting of the peaks of the ^{13}C NMR spectra which also makes peak assignment easier (45).

Pulse sequences

The order and timing of RF pulses in any NMR experiment can be represented in a pulse sequence diagram. The diagram shows what RF pulses are being applied to which nuclei, the power level and duration of the pulse, and where the signal is acquired. An example of a simple pulse sequence can be seen in Figure 11.

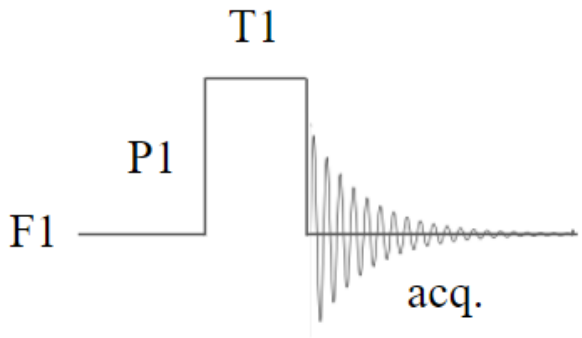

Figure 11: Example of a pulse sequence diagram. F1 is the nuclei that is receiving the RF pulse, P1 is the power level of the pulse, T1 is the duration of the pulse. Acq. stands for the acquisition of signal.

4.2 Two-dimensional NMR

Two-dimensional (2D) NMR spectrum gives new insight into how the nuclei of a molecule are coupled (connected) to each other. The 2D NMR spectrum is composed of two frequency dimensions, one on the x-axis, and one the y-axis. An example of a 2D NMR experiment is a ^{13}C-^{13}C correlation spectrum with proton-driven spin diffusion (PDSD) mixing. The result of the ^{13}C-^{13}C correlation spectrum is a 2D spectrum with ^{13}C frequencies in ppm on each axis. To acquire a ^{13}C-^{13}C correlation spectrum, a series of pulses need to be applied to the sample. An overview of a ^{13}C-^{13}C correlation pulse program can be seen in Figure 12.

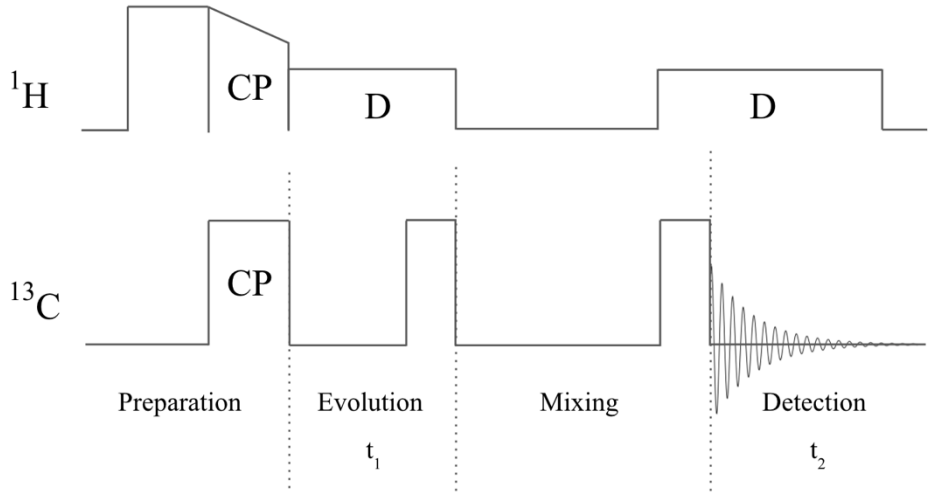

Figure 12: ^{13}C-^{13}C correlation pulse program. Cross-polarization is indicated with CP and decoupling is indicated with D. *Reproduced image from Szeverenyi et al., 1982 (46).*

During the preparation time, ^{1}H relaxes to its equilibrium state. A 90° pulse is then applied to ^{1}H before the ^{1}H magnetization is transferred to ^{13}C via cross-polarization. This locks the magnetization in the xy-plane. During the evolution time, t_1, a 90° pulse on ^{13}C is applied and the ^{13}C chemical shift information is collected in the indirect dimension. In addition, ^{1}H is decoupled during this time. During the mixing time, a magnetization transfer between ^{13}C atoms that are close in space (dipolarly coupled) occurs. Finally, detection occurs after a

second pulse is applied to ^{13}C which regenerates observable ^{13}C magnetization, this is the direct dimension. A 2D Fourier transform is applied which results in a frequency domain 2D spectrum. Peaks along the diagonal appear if no magnetization exchange between nuclei was made during the mixing time. If a magnetization transfer has occurred during the mixing time, peaks will appear off the diagonal. These peaks are called cross-peaks and can help identify carbons that are connected, meaning, that a magnetization transfer has been able to happen between them (46). In Figure 13, a simplified 2D ^{13}C-^{13}C correlation spectrum can be seen with the cross-peaks in the top left and bottom right corner.

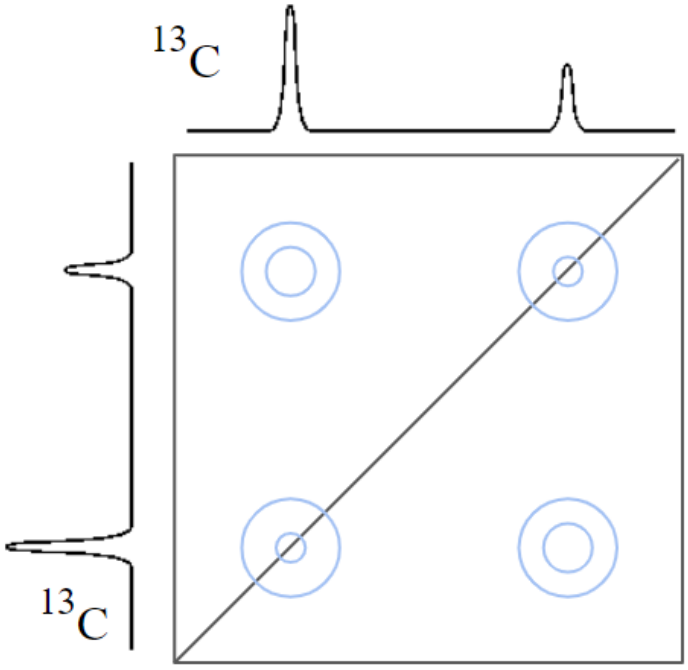

Figure 13: Illustration of a 2D ^{13}C spectra. The peaks are presented with contour plots in blue to represent intensities. The top left and bottom right peaks are cross-peaks. *Reproduced image from Horst, 1996 (47).*

Peak assignment

In order to assign each peak on a 2D spectrum to a specific atom, the average chemical shift values can be used. Each atom, for example, N, Cα, and Cβ, within an amino acid has a characteristic chemical shift value depending on the amino acids. For example, serine is an amino acid that is relatively easy to assign. The Cα chemical shifts lie within the range of 58.69 ± 2.781 standard deviations and Cβ chemical shifts lie within the range of 63.73 ± 4.918 standard deviations (48).

4.3 Solid-State NMR

Solid-state NMR (ssNMR) is a spectroscopy technique that can be used to determine the chemical structure of a wide range of substances such as proteins, molecular solids and inorganic materials (49). For proteins, it can provide structural and dynamic information that reside within a membrane and are therefore in a fixed orientation. This is also true of protein above a sufficient size and slow tumbling speed. This technique can be used for proteins with a high molecular weight or proteins positioned in a lipid membrane (50). In contrast to solution NMR, there are certain interactions that are orientation-dependent. For example, dipolar

couplings and interactions between spins and the electronic environment. Since the molecule is in a fixed position, each nucleus will experience a different magnetic field depending on their orientation in relation to the main magnetic field B_0. In order to make the chemical shift independent of the nuclei's orientation, magic angle spinning (MAS) is applied to the sample. MAS is the physical spinning of the sample at an angle of 54.7° from the external magnetic field which averages out the anisotropic interactions (49).

4.4 NMR Instrument

An NMR instrument contains a large superconducting magnet. The magnet is the most essential part of the spectrometer as it produces the magnetic field that the sample will be placed in (41). The signal-to-noise ratio of the acquired spectra will determine the quality of the data and is dependent on the magnetic field strength. A larger magnetic field strength will improve the SNR (51). The size of the produced magnetic field strength can range from 6 (41) to 23.5 T (52). The superconducting magnet is immersed in liquid helium and together, they are kept in a dewar. This creates a vacuum-insulated space where the temperature is kept close to -273.15°C (0 K). Around this dewar is a second dewar filled with liquid nitrogen which acts as a thermal buffer between the helium and surrounding room temperature (41). Inside the magnet is a cylindrical space called the bore. This is where the sample probe is located and where the sample is inserted. The RF coils and two outlets for gases that spin the rotor are located within the sample probe. The RF coils detect the sample magnetization and generates RF pulses. The sample spinner spins the sample so that the magnetic field becomes as homogenous as possible around the cross-section of the sample probe. Shim coils are located in-between the magnet and the sample probe. These coils are used to correct for any inhomogeneities in the external magnetic field (53).

5 METHOD

5.1 Introduction

The expression, purification, and reconstitution were made on both $^{13}C^{15}N$ uniformly labeled and unlabeled M2[18-60]. Expression was performed using LB and M9 minimum media at 22°C and 30°C. Several methods were used and evaluated during the purification process in order to optimize the purification. The method was initially based on the protocol by Fu *et al.* (8).

5.2 Overall method

The overall method is described in Figure 14. Expression of M2[18-60] in *E. coli* was performed by Dagmar Michl (FMP). The purification process involved sonication, centrifugation, IMAC, CNBr cleavage, lyophilization, and RPC. Reconstitution was performed using denaturing buffer, octyl-beta-glucoside (OG) detergent, DPhPC or POPC dialyzed against H_2O for 5-7 days. A 3.2mm rotor was packed with the resulting proteoliposomes and ssNMR experiments were performed using Bruker Ultrashield 700 MHz PLUS (54).

Figure 14: Overall method used for the study of M2[18-60]. Expression was performed by Dagmar Michl (FMP). Purification involved sonication, centrifugation, IMAC, CNBr cleavage, lyophilization, and RPC.

5.3 Expression

Briefly, M2[18-60] was expressed by inoculating 4 mL LB medium (with Kanamycin) with transformed in *E. Coli* BL21. The over-day culture was incubated at 37°C at 180 rpm until turbid. 2 mL of the over-day culture was inoculated with 100 mL LB or M9 minimal media and left to incubate overnight at 30°C at 180 rpm. The overnight culture was separated into two incubators in 1 L LB or M9 minimal media at 37°C at 180 rpm until the OD_{60} reached 0.6-0.8. The temperature was then reduced to 22°C or 30°C and IPTG was added to a final concentration of 0.15 mM after 20 minutes. The expression of $^{13}C^{15}N$ uniformly labeled M2[18-60] was done by using M9 minimal media with ^{13}C glucose and ^{15}N labeled ammonium chloride. The culture was centrifuged at 5000 rcf for 30 min at 4°C. Pellets were washed with 50 mL PBS buffer and stored at -80°C prior to use.

5.4 Purification according to Fu *et al.*

Purification of M2[18-60] was performed using IMAC with Ni-NTA resin under denaturing conditions using 8M Urea (55). CNBr was used to cleave off trpLE under a nitrogen gas stream

and dialyzed overnight (8) with a 3.5 kDa molecular weight cut-off (56). RPC was performed using a C3 and C4 column.

5.5 Alternative protocol for inclusion body purification

An alternative protocol for inclusion body purification was performed according to the protocol written by Dr. Sascha Lange (FMP). 20 mL Lysis Buffer (50 mM Tris, pH 8.0, 200 mM NaCl) was placed in the Falcon tube containing the pellet from M2[18-60] expression. The sample was sonicated, on ice, at 30% maximum power for 10 minutes, using intervals of 40 seconds on, 30 seconds off. The sample was centrifuged at 20133 rcf (full speed) for 20 minutes. The pellet was resuspended in 10 mL Guanidine buffer (6 M GuHCl, 50 mM Tris, pH 8.0, 200 mM NaCl, 1% (*v/v*) Triton). The sample was sonicated for 10 minutes with the same pulse-sequence as before and then centrifuged again. The supernatant was mixed with 4 mL Ni-NTA beads, pre equilibrated in Lysis buffer. The sample was left to incubate overnight at room temperature with gentle agitation. The remaining steps followed the protocol by Fu *et al*. (8).

5.6 Optimized IMAC purification and CNBr cleavage

A secondary IMAC purification procedure was implemented for one sample. After the Ni-NTA column was washed with 8M Urea, an elution buffer of 6M GuHCl, 500 mM imidazole, and 50 mM Tris (pH 7) were used to elude and collect the protein. The protein concentration was continually measured using Nanodrop (57) until it reached 0.1 g/mL. The amount of elution buffer needed was more than 5 times the column volume. The elution was lyophilized overnight. 3mL of 85% FA was added to the dried protein powder before adding CNBr. A new apparatus for aerating the sample with nitrogen gas during the CNBr reaction was used which can be seen in Figure 15.

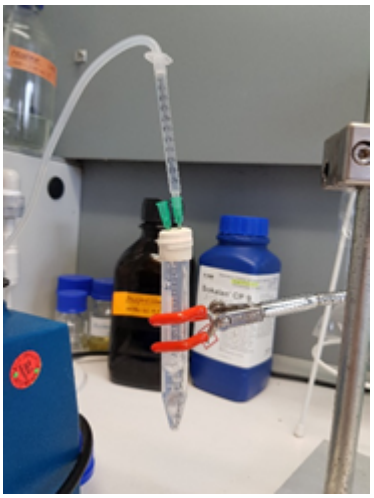

Figure 15: In-house design of apparatus for aerating sample during CNBr reaction. A plastic stopper was used to seal off the reaction from the air and two syringes were inserted through the plastic stopper to allow nitrogen gas to pass through the reaction.

The new apparatus consisted of a falcon tube with a plastic stopper on top which was used to seal off the reaction from the air. Two syringes were placed through the plastic stopper with one submerged into the sample. Nitrogen gas was passed through the submerged syringe

while the gas could escape from the other syringe. The reaction was left in a dark room with aluminum foil around the falcon tube for 1 hour at room temperature. The sample was then dialyzed with a cut-off at 3.5 kDa in H_2O for 1 hour at room temperature to remove CNBr and FA and then over-night after exchanging the H_2O. The sample was then extracted, prior to lyophilization overnight.

5.7 RPC

Dried protein powder was solubilized in 4 mL of 80% FA and loaded onto a C3 or C4 column at a flow rate of 10 mL/min. 2mL fractions were eluted. Diagrams were obtained from the ÄKTA UNICORN™ (58) software. Fractions were collected that exhibited high UV-Vis absorbance at 280 and 214 nm.

5.8 Gel Electrophoresis

Samples from RPC were diluted 1:1 in H_2O and then lyophilized. The dried protein powder was solubilized with either Laemmli or Tris-Tricine buffer. The samples were run on gel electrophoresis with a Mini-PROTEAN TGX™ gel (32) or 15% acrylamide gel. A pre-stained Protein Ladder #26616 (59) was added to the first well in each gel. The electrophoresis was run at 90 V for ~3h. Gels were placed in InstantBlue™ (60) and left on mixer Polymax 1040 (61).

5.9 Pooling of samples

Fractions from RPC that were determined to contain M2[18-60] were pooled, diluted 1:1 with H_2O, and lyophilized. The amount of the protein was determined from the UNICORN™ software.

5.10 Reconstitution

Dried protein powder was dissolved in denaturing buffer (6 M GuHCl, 40 mM phosphate, 30 mM L-Glutamin, 3 mM sodium azide, and 33 mg/mL OG detergent at pH 7.8). Lipids were dissolved in denaturing buffer and added to a protein:lipid ratio of 1:1 w/w. Dissolved protein was added to the lipid:detergent mixture, and dialyzed with a 3.5kDa molecular weight cut-off (56). The sample was dialyzed against 2 L dialysis buffer (40 mM phosphate, 30 mM glutamate, 3 mM sodium azide, pH 7.8) for 5-7 days. Unlabeled M2[18-60] was reconstituted in POPC and labeled M2[18-60] was reconstituted in DPhPC.

5.11 Packing of the rotor

After reconstitution, the sample was ultracentrifuged for 2 h. The supernatant was removed, and the pellet was packed into a 3.2mm rotor with DDS and 3 µl of supernatant. A spacer was used since the amount of sample was insufficient to fill the rotor.

5.12 ssNMR experiments

ssNMR experiments were performed on a Bruker Ultrashield 700MHz PLUS (54) and the computational setup was performed using TopSpin 4.0.7 software. The spectra that were recorded were: 1D ^{1}H and ^{13}C spectra and as well as a 2D ^{13}C-^{13}C correlation with PDSD mixing. The MAS spin rate was initially set to 5000 Hz, with a gas flow of 600 lph, and a target

temperature of 285 K. The final MAS spin rate was set to 11000 Hz with a gas flow at 800 lph, and a target temperature of 285 K. An initial ^{1}H spectrum as acquired to calibrate the spectrum to DDS and evaluate sample temperature from the water peak. Matching & tuning was performed before the first ^{1}H-^{13}C spectrum. The decoupling power of ^{1}H was optimized to 186 W by adjusting the pulse length to 6 µs. The CP power level was optimized to 126 W as well as the duration to 1 ms. A ^{13}C-^{13}C correlation spectrum with PDSD mixing was recorded with the optimized power levels and duration from before (CP conditions). The optimized power level for the 90° ^{13}C pulse was 130 W for 5µs. The spectral widths were set to 299 ppm in the direct and 220 ppm in the indirect dimension. The acquisition times in both dimensions were set to 10 ms, which amounted to 1052 points in the direct and 774 points in the indirect dimension. The dwell time for the directly detected points was 9.5 µs. The experiment was executed with 32 dummy scans and 96 scans per increment. The PDSD mixing was set to a duration of 50 ms and the ^{1}H decoupling during the evolution times was carried out using the SPINAL-64 decoupling sequence and a decoupling power of 83 kHz. Processing was performed using the Bruker TopSpin 4.0.7 software, plotting was performed using CCPNMR Analysis 2.4.2 (62). To process the 2D NMR spectrum, the 1052 acquired points in the direct dimension and the 774 points in the indirect dimension were zero-filled to give 4K points in each dimension after Fourier transform. In both dimensions, a squared sine bell window function with a phase shift of 45° was applied. The phases of the signals were manually corrected go give pure absorption line-shapes and the baseplane was corrected in both dimensions. Assigned peaks from BMRB entry 25788 (63) were overlayed on the ^{13}C-^{13}C correlation spectrum. BMRB entry 25788 entry included the assigned peaks of the two-fold symmetric structure of the M2[18-60] with the S31N mutation in DPhPC lipid bilayers originating from the article by Loren B. Andreas *et al.* (9). Assigned peaks were limited to residues 25-52. (9). This entry had the most similar conditions to the conditions used in this thesis.

6 RESULTS

6.1 Introduction

The following results are from various purifications of M2[18-60] expressed at different temperatures, in two different growth media and are labeled or unlabeled. RPC elution fractions were collected and lyophilized. Gel electrophoresis was performed in order to investigate proteins present in the factions. The ssNMR results of reconstituted M2[18-60] were obtained from the Bruker TopSpin software (64).

6.2 RPC diagrams

The results from RPC are diagrams from the UNICORN™ software. The graphs show the mAU of UV 280 nm. The amount of AU can be seen on the y-axis and the total amount of flowthrough in mL can be seen on the x-axis as well as the fraction number. The flow-through in mL can be converted into time, as the flow rate was 10mL/min. Each 10 mL, therefore, represents 1 minute. RPC was performed using purified samples expressed at 22°C and 30°C, labeled and unlabeled, and using a C3 or C4 column. Fractions were collected that exhibited significant UV 280 absorbance.

C3 column, LB media, unlabeled

In Figure 16 the results from RPC using a C3 column of unlabeled M2[18-60] expressed at 22°C in LB media can be seen.

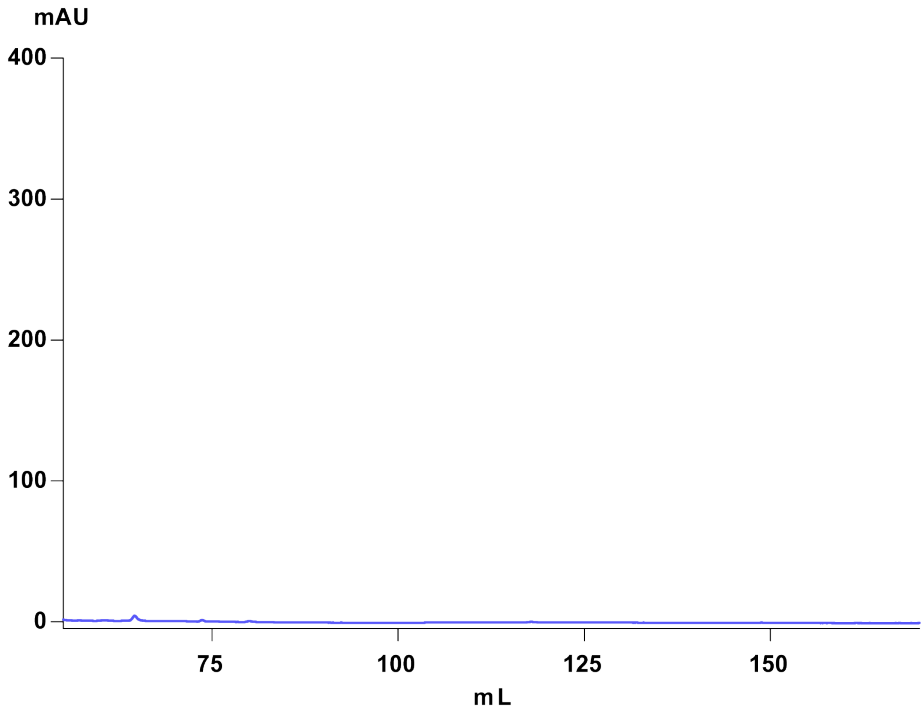

Figure 16: Results from RPC using a C3 column of unlabeled M2[18-60] expressed in LB media.

C3 column, 22°C expression, unlabeled

In Figure 17 the results from RPC using a C3 column of unlabeled M2[18-60] expressed at 22°C in M9 minimal media can be seen. Fractions 21, 22, 24, 25, 29, 30, 32, 38, 39, 40, 54, 56, 58, 64, 71, and 72 were collected.

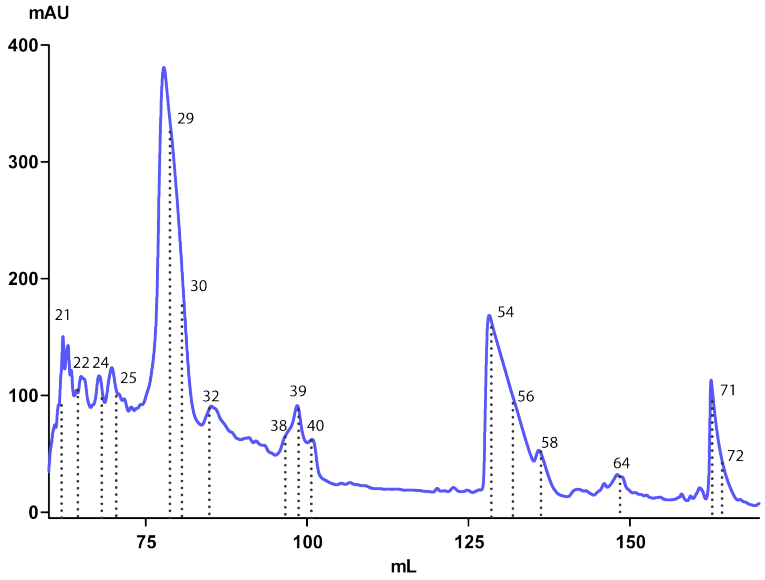

Figure 17: Results from RPC using a C3 column of unlabeled M2[18-60] expressed at 30°C in M9 minimal media. Fractions 21, 22, 24, 25, 29, 30, 32, 38, 39, 40, 54, 56, 58, 64, 71 and 72 were collected and are indicated with lines.

C3 column, 30°C expression, unlabeled

In Figure 18 the results from RPC using a C3 column of unlabeled M2[18-60] expressed at 30°C in M9 minimal media can be seen. Fractions 21, 23, 26, 27, 28, 31, 38, 39, 47, 48, 53, 55, 69, 70, and 71 were collected.

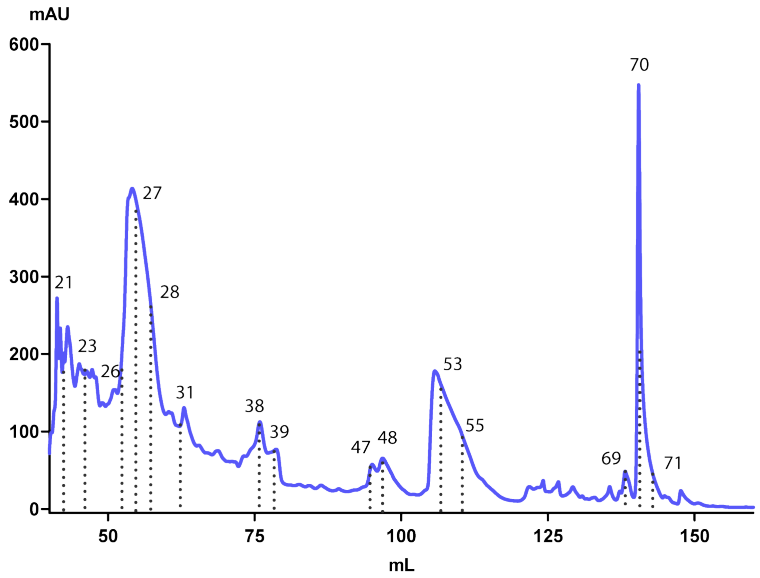

Figure 18: Results from RPC using a C3 column of unlabeled M2[18-60] expressed at 30°C in M9 minimal media. Fractions 21, 23, 26, 27, 28, 31, 38, 39, 47, 48, 53, 55, 69, 70 and 71 were collected and are indicated with lines.

C3 column, LB vs M9 minimal media, unlabeled

In Figure 19 a comparison of the RPC graphs from unlabeled M2[18-60] expressed in LB vs M9 minimal media can be seen. The blue line indicate the UV 280 AU of sample expressed in M9 minimal media and the red line indicate the UV 280 AU of sample expressed in LB media. Each graph has been adjusted so that the first elution peak of each graph was aligned.

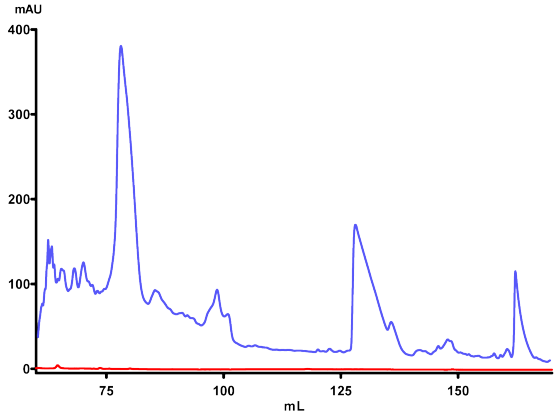

Figure 19: Comparison of M2[18-60] expressed in LB vs M9 minimal media. UV 280 AU of sample expressed in M9 minimal media are in blue and UV 280 AU of sample expressed in M9 minimal media are in red.

C3 column, 22°C vs 30°C expression, unlabeled

In Figure 20 a comparison of the RPC graphs from unlabeled M2[18-60] expressed at 22°C and 30°C can be seen. The green line is the UV 280 AU of expression at 30°C and the blue is the UV 280 AU of expression at 22°C. Circle indicates peaks that were pooled and lyophilized. Each graph has been adjusted so that the first elution peak of each graph was aligned.

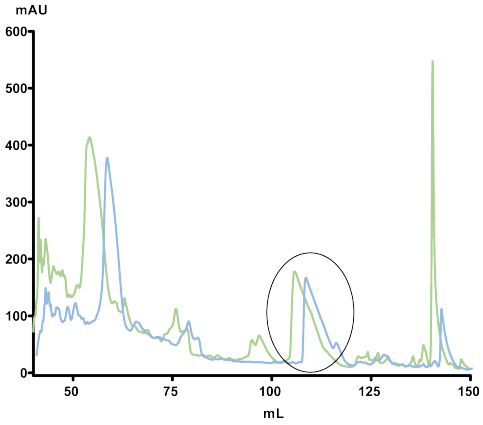

Figure 20: UV 280 of M2[18-60] expressed at 22°C (blue) and 30°C (green). Circle indicates peaks that were pooled and lyophilized.

Fractions 45-51 from RPC of unlabeled M2[18-60] expressed at 30°C and 45-50 of unlabeled M2[18-60] expressed at 22°C were pooled and lyophilized. The calculated total mass of the pooled fractions was 6.5 mg.

C4 column, 22°C expression, unlabeled

In Figure 21 the results from RPC using a C4 column of unlabeled M2[18-60] expressed at 22°C in M9 minimal media can be seen. Fractions 24, 25, 28, 32, 34, 46, 52, 56, and 63 were collected.

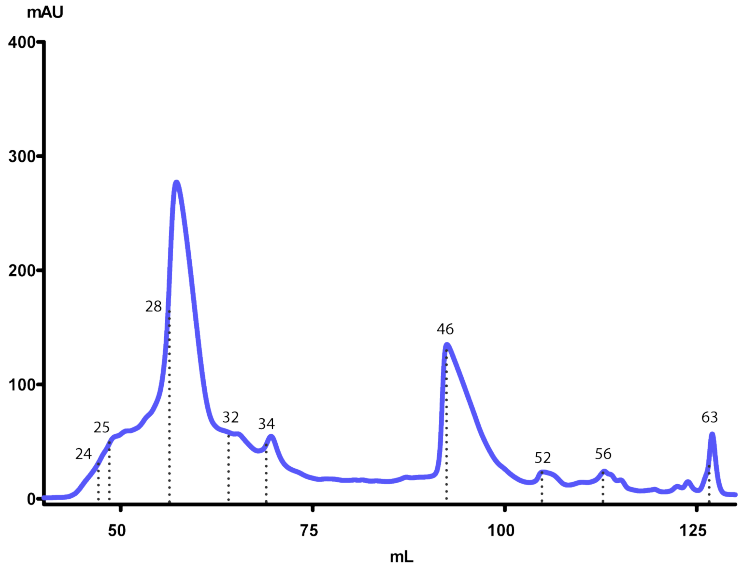

Figure 21: Results from RPC using a C4 column of unlabeled M2[18-60] expressed at 22°C in M9 minimal media. Fractions 24, 25, 28, 32, 34, 46, 52, 56 and 63 were collected and are indicated with lines.

C3 vs C4 column, 22°C expression, unlabeled

In Figure 22 a comparison of the RPC graphs from unlabeled M2[18-60] expressed at 22°C run in a C3 and C4 column can be seen. The blue line indicates the UV 280 AU when using a C4 column and green is the UV 280 AU when using a C3 column. Each graph has been adjusted so that the first elution peak of each graph was aligned.

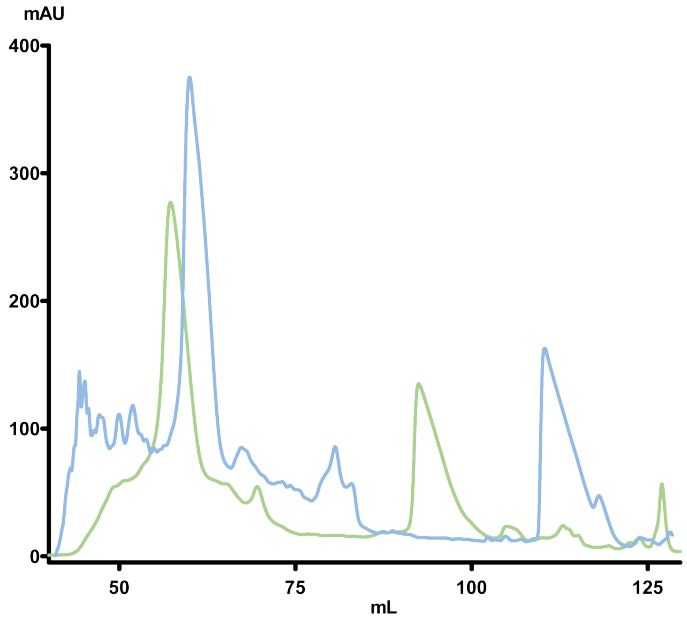

Figure 22: UV 280 mAU of RPC performed using a C4 (blue) and C3 column (green).

Peaks of interest

In Figure 23 the peaks of interest from RPC can be seen of unlabeled M2[18-60] expressed at 22°C and 30°C using a C4 and C3 column.

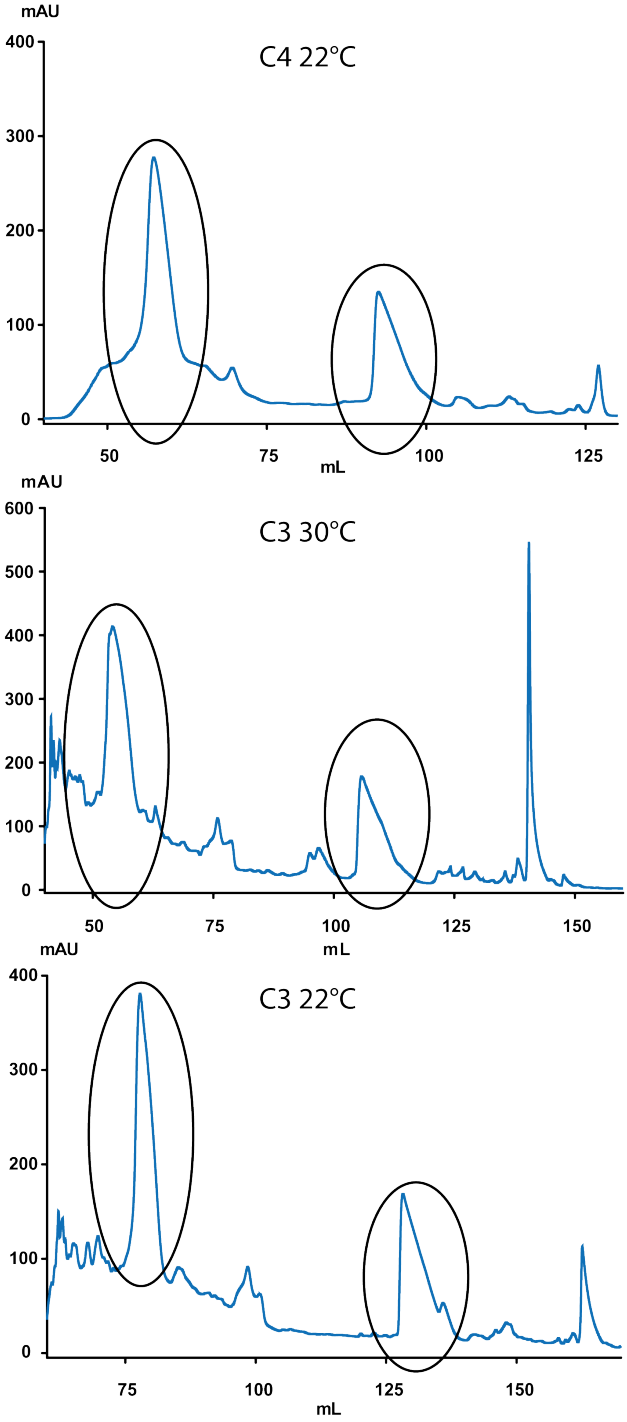

Figure 23: Peaks of interest from RPC of unlabeled M2[18-60] expressed at 22°C using a C4 column (top), 22°C using a C3 column (middle), and 22°C using a C3 column (bottom).

C4 column, 22°C expression, labeled

In Figure 24 the results from RPC using a C4 column of uniformly $^{13}C^{15}N$ labeled M2[18-60] expressed at 22°C in M9 minimal media can be seen. Fractions 30, 33, 38, 43, 46, 47, 48, 49 and 50 were collected.

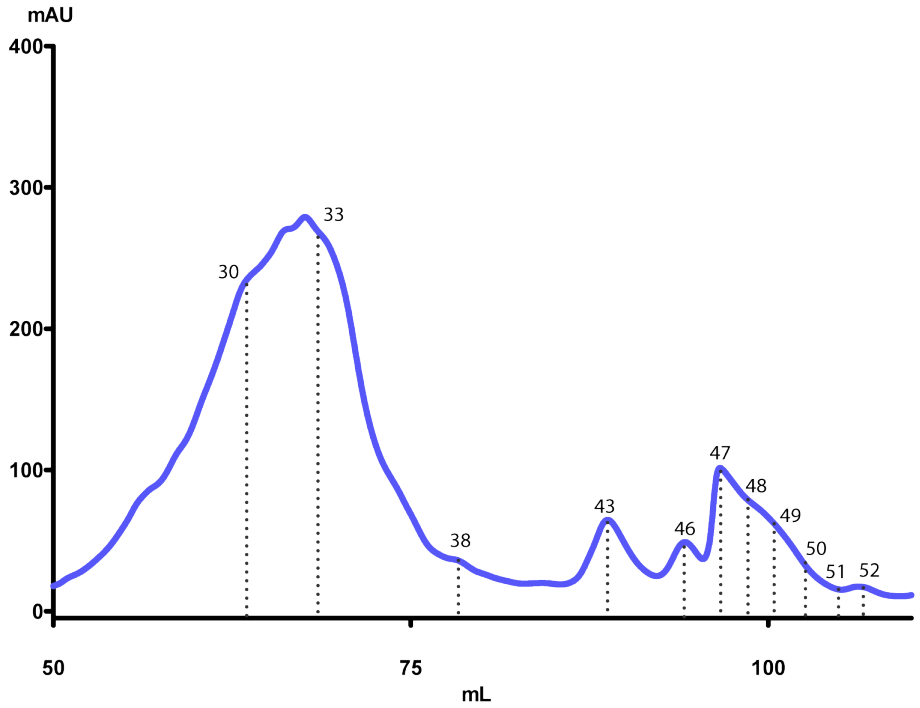

Figure 24: Results from RPC using a C4 column of labeled M2[18-60] expressed at 22°C in M9 minimal media. Fractions 30, 33, 38, 43, 46, 47, 48, 49, 50, 51 and 52 were collected and are indicated with lines.

Fractions 46 to 52 were pooled and lyophilized. The calculated total mass of the fractions that were pooled was 4 mg.

6.3 Gel electrophoresis

The following results include the gels that were run on elution fractions from RPC. The first well in each gel contains Protein Ladder #26616 and bands are labeled with known sizes.

C3 column, 22°C vs 30°C expression, unlabeled

In Figure 25 the gel image can be seen of fractions from RPC using a C3 column of unlabeled M2[18-60] expressed at 22°C and 30°C in M9 minimal media. A sample of the M9 minimal media before (BF) induction with IPTG is labeled BF in the middle column. Gel electrophoresis was run on a 15% gel.

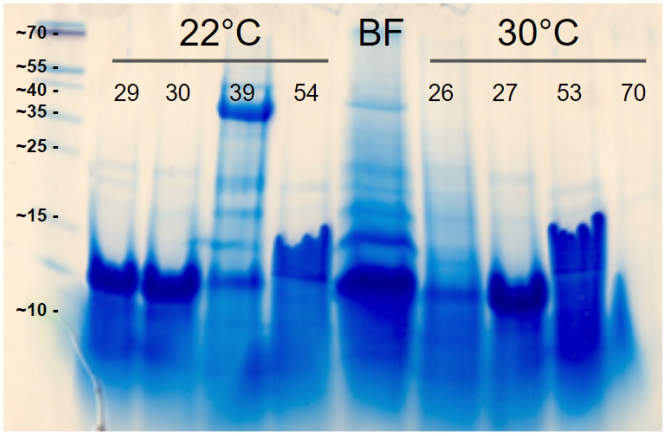

Figure 25: Gel image of collected fractions from RPC using a C3 column of unlabeled M2[18-60] expressed at 22°C and 30°C in M9 minimal media. Column labeled BF is the sample taken of the M9 minimal media before induction with IPTG.

C4 column, 22°C expression, unlabeled

In Figure 26 the gel image can be seen of fractions from RPC using a C4 column of unlabeled M2[18-60] expressed at 22°C in M9 minimal media. Marked boxes were cut out and sent for MS analysis. Box labeled 1 has the size of ~12 kDa, 2 is ~40kDa, 3 is ~20kDa, 4 is ~13 kDa and 5 is ~15 kDa. Gel electrophoresis was run on a Mini-PROTEAN TGX™ gel.

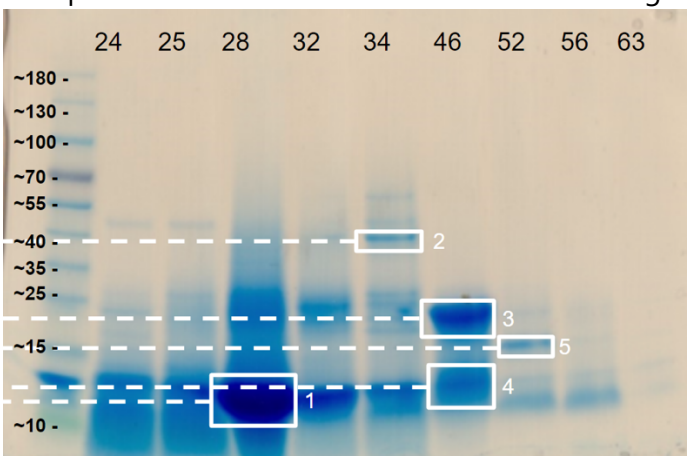

Figure 26: Gel image of fractions using C4 column of unlabeled M2[18-60] expressed at 22°C in M9 minimal media. The first lane contains the Protein Ladder #26616 with marked sizes in kDa. Subsequent lanes are marked with fraction number from RPC of purified M2[18-60] expressed at 22°C in M9 minimal media. Marked boxes were collected for mass spectrometry analysis. Box labeled 1 has the size of ~12 kDa, 2 is ~40kDa, 3 is ~20kDa, 4 is ~13 kDa, and 5 is ~15 kDa.

C4 column, 22°C expression vs C3 column 30°C and 22°C expression, unlabeled

In Figure 27 the gel image can be seen of fractions collected of the two peaks of interest from RPC using a C4 and C3 column of unlabeled M2[18-60] expressed at 22°C and 30°C in M9 minimal media. The topmost marked bands have the size of ~20 kDa and the bottom-most marked bands have the size of ~13 kDa. Gel electrophoresis was run on a Mini-PROTEAN TGX™ gel.

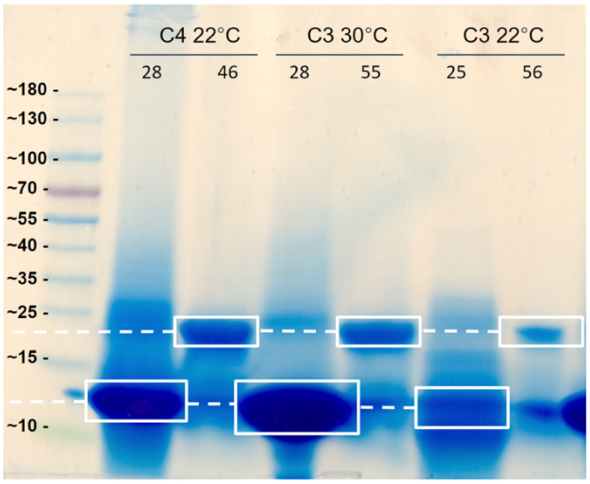

Figure 27: Gel image of fractions collected of the two peaks of interest from RPC using a C4 and C3 column of unlabeled M2[18-60] expressed at 22°C and 30°C in M9 minimal media. The first lane contains the Protein Ladder #26616 with marked sizes in kDa. The topmost marked bands have the size of ~20 kDa and the bottom-most marked bands have the size of ~13 kDa.

C4 column, 22°C expression, labeled

In Figure 28 the gel image can be seen of collected fractions from RPC using a C4 column of uniformly $^{13}C^{15}N$ labeled M2[18-60] expressed at 22°C in M9 minimal media. The prominent bands on columns 43 to 50 have the size of ~13 kDa and the other prominent bands on columns 47 to 49 have the size of ~20 kDa. Gel electrophoresis was run on a Mini-PROTEAN TGX™ gel.

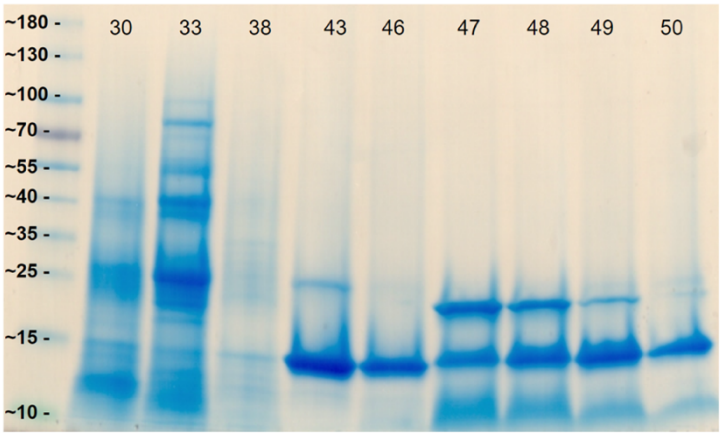

Figure 28: Gel image from gel electrophoresis. The first lane contains the Protein Ladder #26616 with marked sizes in kDa. Subsequent lanes are marked with fraction number from RPC using a C4 column of purified uniformly $^{13}C^{15}N$ labeled M2[18-60] expressed at 22°C in M9 minimal media.

Supernatant and pellet of reconstituted labeled vs unlabeled M2[18-60]
In Figure 29 the gel image can be seen of samples from the pellet (P) and supernatant (S) from the reconstitution of labeled and unlabeled M2[18-60]. Gel electrophoresis was run on a Mini-PROTEAN TGX™ gel.

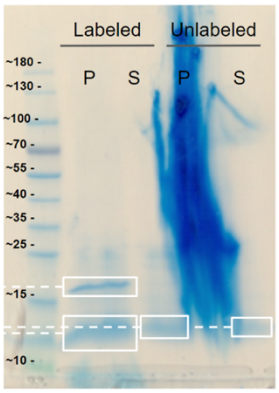

Figure 29: Gel image of the pellet (P) and supernatant (S) taken from the reconstitution of labeled and unlabeled M2[18-60]. The first lane contains the Protein Ladder #26616 with marked sizes in kDa. Marked boxes have the size of ~17 kDa (top) and ~13 kDa bottom.

6.4 Mass spectrometry

In Figure 30 the chromatogram from mass spectrometry analysis of the five samples taken from the gel in Figure 26 can be seen. The topmost graph is gel sample number 1 and so on. The use of SIM allowed for M2[18-60] to be identified in all samples. M2[18-60] was most abundant in samples 3 and 4 as seen by the increase in relative abundance in graphs 3 and 4.

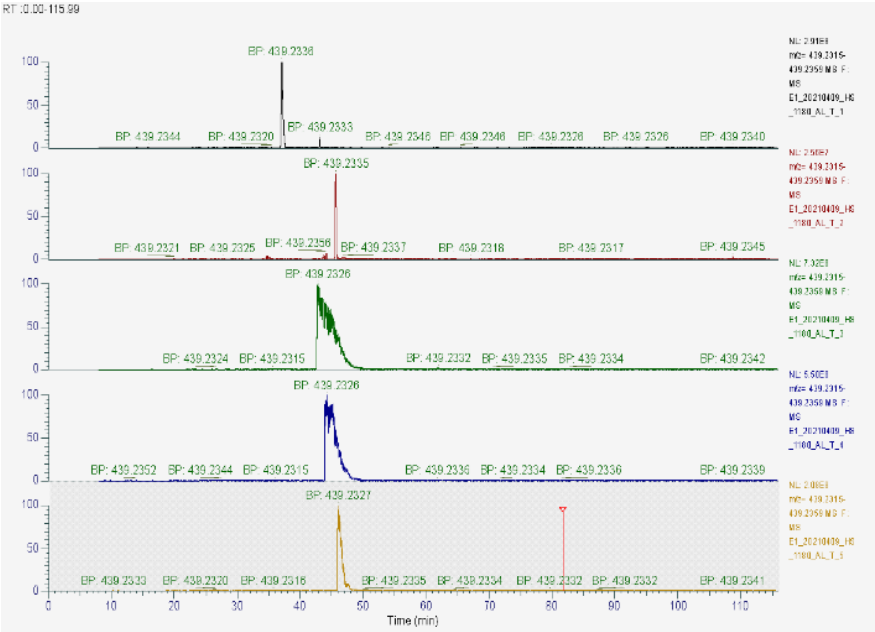

Figure 30: Chromatogram of the five SDS-PAGE gel samples. The topmost graph is the first gel sample and so on.

6.5 Reconstitution

Figure 31 and Figure 32 show the samples of reconstitution of uniformly labeled M2[18-60] after reconstitution and dialysis for 7 days. The solution had white particles floating around which were lighter than water. Even after centrifugation, white particles could be seen in the supernatant.

Figure 31: Dried protein powder in Eppendorf tube dissolved in denaturing buffer. Small white particles can be seen floating in the buffer.

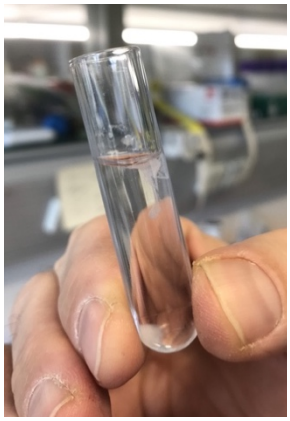

Figure 32: Supernatant after centrifugation of dissolved protein. Small white particles can be seen floating in the buffer.

6.6 SsNMR

In Figure 33 the ^{13}C-^{13}C correlation spectrum of M2[18-60] reconstituted into proteoliposomes using DPhPC can be seen. The spectrum is limited to the aliphatic region where four serines were able to be identified as well as one peak that might be indicative of a fifth serine.

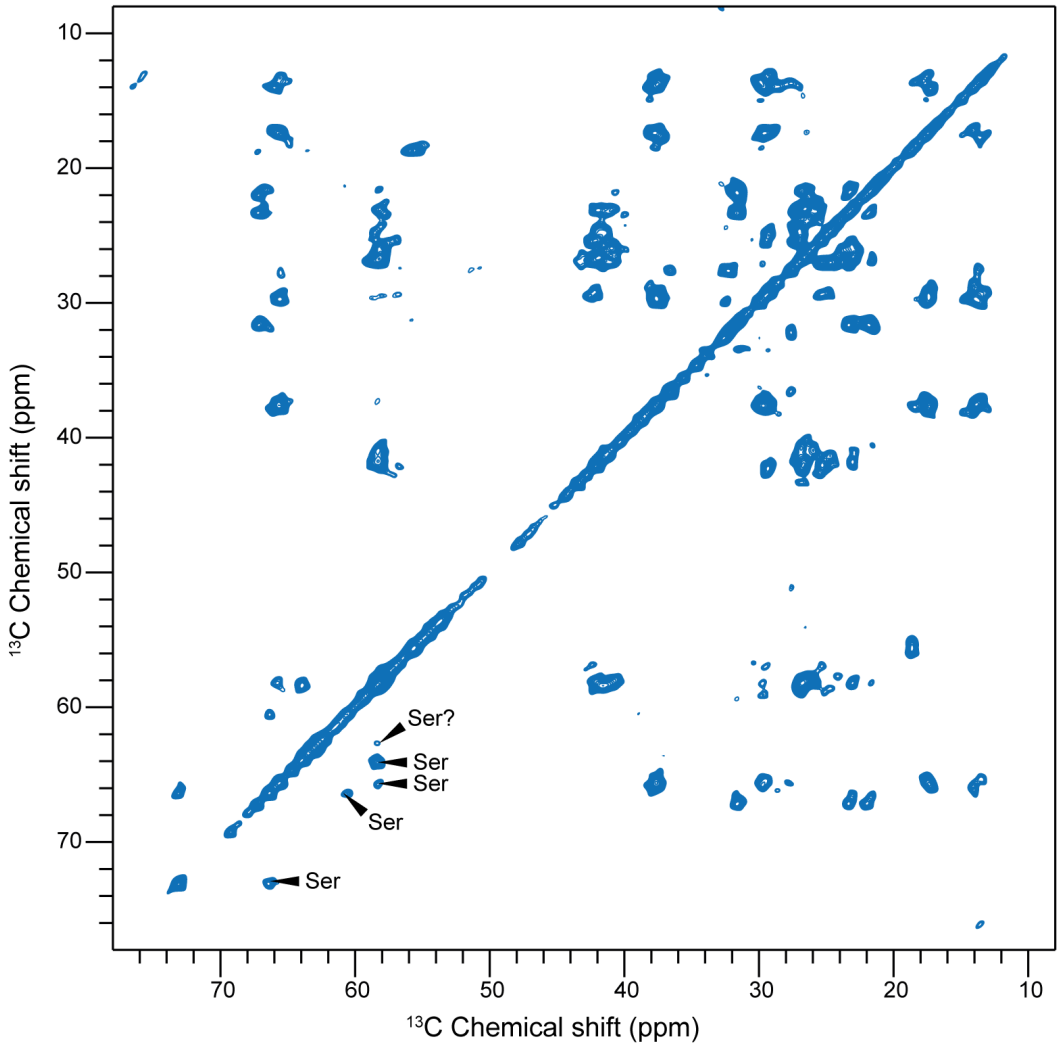

Figure 33: ^{13}C-^{13}C correlation spectrum of the aliphatic region of M2[18-60] reconstituted into proteoliposomes using DPhPC. Four serines and possibly a fifth are indicated.

In Figure 34 the ^{13}C-^{13}C correlation spectrum of M2[18-60] overlayed with assigned peaks from BMRB entry 25788 can be seen. The M2[18-60] used in the overlayed investigation was an S31N mutant. 31N and 32I have been highlighted in red.

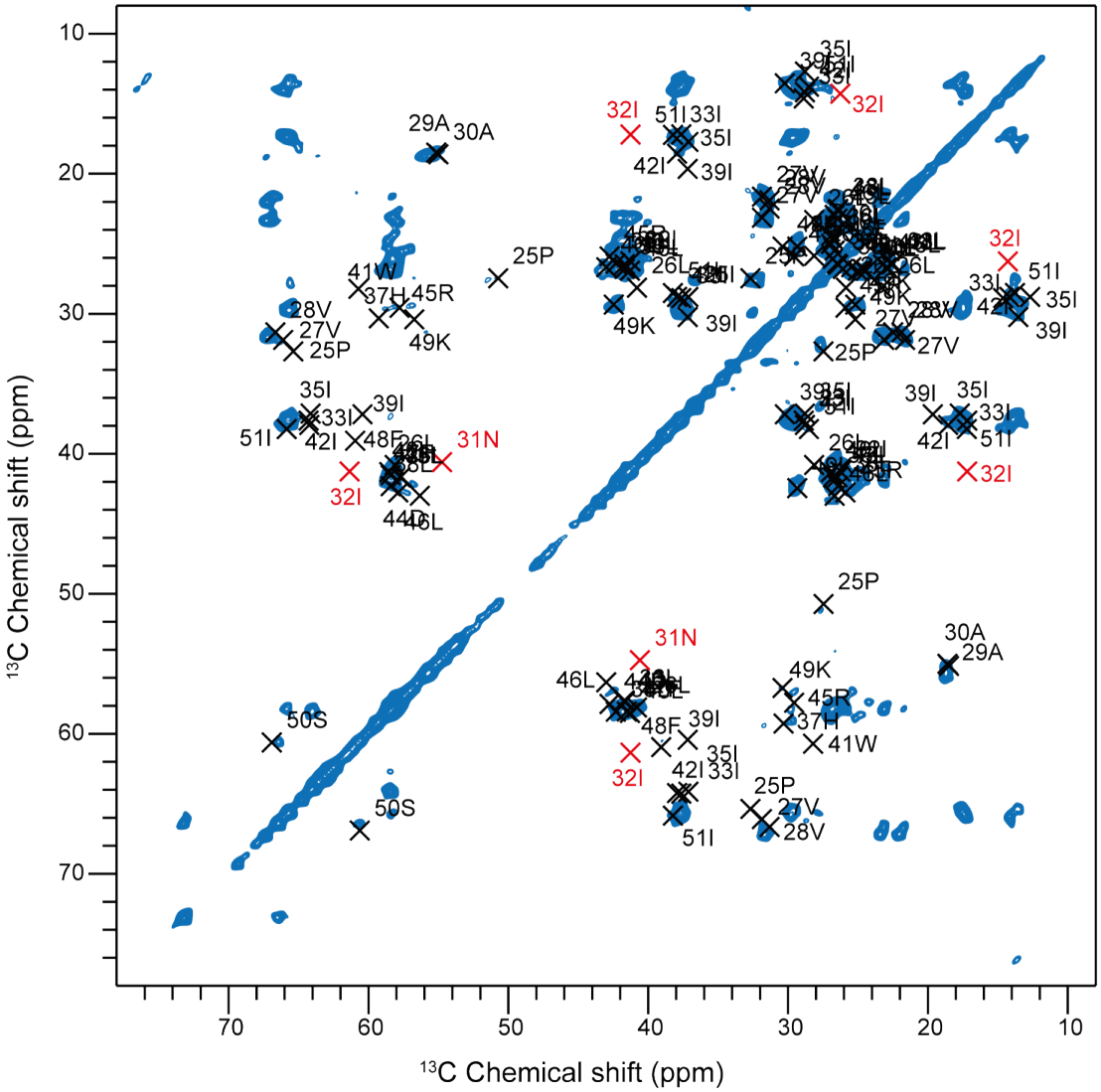

Figure 34: ^{13}C-^{13}C correlation spectrum of M2[18-60] overlayed with assigned peaks from BMRB entry 25788. 31N and 32I have been highlighted in red.

7 DISCUSSION

7.1 Summary of results

The following sections will contain discussions of the results that have been presented in the previous section as well as a discussion concerning the methodologies used. Despite the difficulties in optimisation, a robust growth and purification protocol has now been established. The presence of folded M2[18-60] similar to that in previously published works has been confirmed by ssNMR.

7.2 RPC results

Peaks of interest
The peaks of interest highlighted in Figure 23 were selected because they had the most consistent replication within the samples and had the largest peaks at UV 280 nm. The two peaks theoretically represent two different proteins being eluted since they have different retention times.

LB vs M9 minimal media
One of the most prominent results that can be seen is that there is a large difference in protein yield between the expression of M2[18-60] in LB and M9 minimal media. This can be seen in Figure 19. The graphs show a remarkably higher absorbance at 280 nm, indicating a higher concentration of protein generated in M9 media than LB. This result is expected since the expression in M9 minimal media drives the formation of inclusion bodies.

22°C vs 30°C
For expression at 22°C vs 30°C, the RPC results in Figure 20 show a couple of notable differences. Firstly, the graphs are slightly shifted, with the graph of RPC of M2[18-60] expressed at 22°C being shifted to the right, indicating a longer retention time. However, the shift is minimal and and is likely to be a matter of misalignment. This discrepancy could be adjusted by matching the mL scale on the x-axis to another elution peak in each graph. Secondly, there is one small peak between 90-100 mL that is present for M2[18-60] expressed at 30°C but not for 22°C. This could be a result of other proteins being expressed at 30°C due to stress, incomplete cleavage of the construct, or artifacts that have become stuck in the column and later being washed out during this time. Thirdly, the largest peak seen on the graph might be an outlier as it is uncertain what this might be.

C3 vs C4 column
There does not seem to be a large difference in performing RPC with a C3 vs a C4 column. Since the RPC stage is only for the purification of M2 from the inclusion bodies, the slight difference in peak retention time between C3 and C4 seen in Figure 22 is insignificant. There is a slight difference in the height of the peaks as well a shift to the right when using the C4 column, indicating that it has a longer retention time, which was expected. In addition, when using the C3 column, there are a few more peaks present, which might indicate that proteins are able to be separated more precisely using a C3 column. However, since the peaks of interest are still distinguishable on each graph, the column of choice does not seem to have a large significance in purification of M2[18-60].

Additional comments

Since RPC will separate proteins depending on the hydrophobicity, it is important to consider the GRAVY scores of each protein that is being eluted. Since the His6-trpLE domain has a low GRAVY score (-0.135), it is hydrophilic and does not associate readily to the hydrophobic stationary phase. The His6-trpLE domain will therefore elute early. However, M2[18-60] has a high GRAVY score (0.566) and will therefore associate to the hydrophobic stationary phase and elute later. When considering the peaks of interest, this would mean that the leftmost peaks is the His6-trpLE domain (being eluted early) and the rightmost peaks are M2[18-60]. Some other things to consider is whether M2[18-60] is running as a monomer, dimer or tetramer. If the M2[18-60] is in its tetrameric form, hydrophobic amino acids will be exposed on the. The result of this is that the hydrophobic residues would associate to the hydrophobic stationary phase, leading to a later elution than expected from the GRAVY score of the monomer. Another aspect to consider is how an un-cleaved product would be eluted from the column. An un-cleaved product has a GRAVY score of which would be eluted in between the His6-trpLE domain and M2[18-60].

7.3 Gel electrophoresis results

Initially it was thought that M2[18-60] would run on the gel as a monomer, producing a band at ~5 kDa. Since the gels were not able to separate proteins at this size, it was hard to determine whether there was a band below the 10 kDa band of the protein ladder. After considering how the M2[18-60] would be affected by the buffers, and reading previous publications, it was determined that it ran through the gel as a tetramer since it is able to form a stable tetramer and appear as a band at ~20 kDa.

22°C vs 30°C

The use of SDS-PAGE was discontinued after the results seen in Figure 25. As the figure shows, the gel gave very poor results and bands were not separated well. Despite the poor quality of the gel, there are several bands of interests that merit discussion. Firstly, for fraction 39, the topmost band at ~35 kDa is relatively large and corresponds to a small peak from the RPC results. Judging from the size of the protein, this might correspond to the un-cleaved His_6-trpLE-M2[18-60] dimer with a size of 37.2 kDa. Theoretically, this dimer could be possible if the cleavage of His6-trpLE domain was unsuccessful and only a single dimer is running through the gel. To support this, the theorized dimer would have a GRAVY score of -0.102, causing it to elude after the His6-trpLE domain but before M2[18-60].

C4 column, 22°C expression, unlabeled

The most prominent bands on the gel in Figure 26 are from fractions 28 and 46. These elution fractions also correspond to the peaks of interest found in Figure 23. The band labeled 1 has a size of ~13 kDa and the band labeled 3 has a size of ~20 kDa. Theoretically, these band corresponds to the His6-trpLE domain and M2[18-60] tetramer receptively. Interestingly, there are several bands at ~13 kDa in all the lanes which could be spill-over from the first four wells or His_6-trpLE that has not been separated by RPC. Two other bands of interest are the bands labeled 2 and 5. Band 2 appears at ~40 kDa in fraction 34. The construct resulting in a band this size might be the dimer discussed above. Band 5 appears at ~15 kDa in fraction 52. The construct resulting in a band this size might be a His_6-trpLE-M2[18-60] monomer as a result of incomplete cleavage. However, it would be unlikely that this construct elutes so late since it quite hydrophilic with a GRAVY score of -0.102.

C4 column, 22°C expression vs C3 column 30°C and 22°C expression, unlabeled

Figure 27 seems to confirm that the peaks of interest correspond to His_6-trpLE domain and M2[18-60] tetramer. Each elution fraction taken from the first peak of interest in each RPC indicate a band at ~13 kDa and each elution fraction taken from the second peak of interest indicate a band at ~20 kDa. The bands are very smeared which means that the gels are overloaded with protein. In addition, bands indicating a His_6-trpLE domain in fractions where only M2[18-60] tetramer was expected, can be seen. This could be due to spill-over or constructs smaller than M2[18-60], however, the small mass of these bands makes it difficult to determine the mass of the protein. For BF (before induction), many bands are smeared which makes the interpretation difficult. One large band can be seen at ~13 kDa, which might indicate the overexpression of other protein in the host cell.

C4 column, 22°C expression, labeled

In Figure 28 there are several bands that are well defined. When looking at fractions 47, 48 and 49, there are two separate bands. The top three bands for these fractions correspond to a mass of ~20 kDa, indicating the presence of M2[18-60] tetramer. When looking at fraction 43, two separate bands can be seen, one at ~23 kDa and the other at ~13 kDa. The band at ~23 kDa is relatively faint and does not match any of the protein sizes that are expected. In addition, fractions 43-50 all have a clear band at ~13 kDa which indicates the presence of a His_6-trpLE domain in all of the samples. Fraction 30 does not have well defined bands, indicating overloading or non-separated proteins. The same can be said about fraction 33, however, strong bands can be seen at ~25 kDa, ~40 kDa, ~55 kDa and ~85 kDa. The identity of these proteins is uncertain since they do not match any of the expected cleavage products.

Supernatant and pellet of reconstituted labeled vs unlabeled M2[18-60]

In Figure 29 a clear overloading of the unlabeled pellet of reconstituted M2[18-60] is seen in the third column of unlabeled pellet sample. The large smear is due to overloading of the gel with sample. However, the other samples show promising results. For the pellet of labeled reconstituted M2[18-60], two faint bands appear. One with the size of above 15 kDa and one ~12 kDa. It is unclear whether the band above 15 kDa is M2[18-60]. In both samples of the supernatant, only one band appears at ~12 kDa. This indicates that there is no M2[18-60] in the supernatant, and that M2[18-60] has been successfully incorporated into liposomes. Despite the presence of the band at ~12 kDa in the pellet of labeled M2[18-60], proposed to be a His_6-trpLE domain, the ssNMR spectra would have detected any significant impurity.

Additional comments

It is still unclear whether M2[18-60] would run through the gel as a monomer, dimer or tetramer as all configurations are possible. In addition, larger sized proteins that were seen can be an un-cleaved product and could potentially also run through the gel as a monomer, dimer, tetramer or other impurities. As a band of ~20 kDa was seen in fractions collected where M2[18-60] was believed to be eluted, it is more likely that the theory that M2[18-60] is stable enough to form a tetramer even in the presence of the denaturing conditions of buffers used for gel electrophoresis. In future investigations, it would be beneficial to re-run gels that appeared to be overloaded in order to get a more precise indication of the band size and avoid smears.

7.4 Mass spectrometry

The spectra from mass spectrometry seen in Figure 30 indicates that M2[18-60] is present in all samples. The results indicate that either M2[18-60] has not been purified and cleaved off properly, or that RPC is not separating M2[18-60] from other proteins in the sample. In addition, the cutting of the bands from the gel may cause impurities in the samples as even small impurities can appear to be substantial. One notable thing is that samples 3 and 4 have a larger abundance of M2[18-60] compared to the other samples, which indicates that M2[18-60] is present in larger concentrations in fraction 46. This also means that the rightmost peak of interest contains more M2[18-60] than the other peaks.

7.5 Different methodologies used

The different operation parameters used for sonication and centrifugation showed no significant difference. Changes that were made from the original protocol by Fu *et al.* (8), for example the change in sonication pulse frequency and time were made due to differences in machine manufacture.

7.6 IMAC purification

Purification of M2[18-60] using IMAC greatly affects M2[18-60] purity. Since the Ni-NTA beads have an affinity for any His sequence, it is possible that other proteins in the sample associate to the beads. In addition, depending on the buffers used to elute the protein, more or less of the protein will be eluted. As was seen when monitoring the concentration of protein in the elute during optimized IMAC, the required amount of elution buffer was more than 5 times the column volume. For future purification with IMAC, it would be interesting to monitor the protein concentration during elution when using FA or other elution buffers. In addition, it would be preferable monitor the amount of elution buffer required to elute all of the protein as well as analyzing samples taken at different intervals using gel electrophoresis.

7.7 CNBr cleavage

Initial results from RPC and gel electrophoresis do not indicate a well purified protein. Since bands that had a larger size than expected cleavage products appeared on the gel, it was thought that the CNBr cleavage was unsuccessful. To optimize the reaction, the in-house apparatus seen in Figure 15 was constructed so the nitrogen gas would continuously flow through the sample, minimizing unwanted reactions with oxygen in the surrounding air. The CNBr cleavage would be a valid variable to investigate since previously published papers mention significantly different durations of the CNBr cleavage and FA concentrations. Shorter reaction times may result in insufficient cleavage, while longer reaction times might result in unspecific cleavage. To optimize the reaction time, samples could be collected at regular intervals so that cleavage products could be analyzed using gel electrophoresis.

7.8 Reconstitution

The white particles seen in the supernatant of the labeled, reconstituted M2[18-60] in Figure 31 and Figure 32 are the same density as the denaturing buffer since they do not become pelleted after ultracentrifugation. It is unclear what these particles might be, perhaps some impurities from RPC, detergent, lipids or other proteins. These particles were not present in the reconstitution of unlabeled M2[18-60] (not pictured). Since POPC was used for

reconstitution of unlabeled M2[18-60] and DPhPC was used for reconstitution of labeled M2[18-60], further work could investigate whether white particles appear when using DPhPC for reconstitution of unlabeled M2[18-60] and POPC for reconstitution of labeled M2[18-60]. In addition, further gel analysis or mass spectrometry of the samples taken from the supernatant might indicate what the particles might be.

7.9 ssNMR

The ssNMR spectra show good resolution and spectral quality. This means that M2[18-60] is able to form a stable tetramer within liposomes since ^{13}C-^{13}C correlation spectrum in Figure 33 indicate five serine cross-peaks as expected due to the five serine residues present in the amino acid sequence. To obtain better signal, more sample would need to be produced with the aim to avoid the use of a spacer, which reduces the SNR. The overlayed assignments seen in Figure 34 show that the majority of the peak clusters have corresponding assignment clusters. The M2[18-60] that was analyzed in the overlayed peak assignments was a S31N mutant and is therefore not expected to have a corresponding peak on the ^{13}C-^{13}C correlation spectrum at this residue. Neighboring residues would also be expected to have a different chemical shift due to this mutation. This was confirmed by the lack of assignable peak for residues 31N and 32I. In addition, the overlayed peak assignments were limited to residues 25-52. This explains why only one serine was assigned (S50), as other serienes in the sequence are outside of this scope. For future investigations, it would be beneficial to assign all peaks of the acquired spectra.

7.10 Possible application of protocol to purify Mg and MgM

Since the focus of at the FMP is to study other viroporins, it would be interesting to investigate whether the methods used in this thesis could be applied to the Mg and MgM viroporins of DV and WNV, respectively. Since the purification of the Mg protein of DV was previously unsuccessful, this method could potentially be used instead. However, the method is limited by the use of CNBr since it will cleave at any methionines.

7.11 Future considerations and pharmacological applications

For future studies of M2 and the potential pharmacological applications, the purification of M2[18-60] Need to be optimised further. The ssNMR results indicate that it is possible to produce a pure enough sample to be analyzed, however, a more pure sample would need to be produced for the purpose of assay development. Once a potential drug candidate has been selected, an optimized protocol based on the methods described in this thesis could be used to purify and study the interactions between the drug candidate and M2[18-60] using ssNMR.

8 CONCLUSION

The purification of M2[18-60] was not as straightforward as initially expected. Although several papers have previously reported well-defined spectra, they they lack a detailed purification methodology, which hindered the study of M2[18-60] at the FMP. Despite the purification taking longer than expected, a sufficient sample was produced in order to confirm the presence and purity of M2[18-60] using ssNMR. The overlayed assignment clusters corresponded to peaks found on the ^{13}C-^{13}C correlation spectrum, showing that M2[18-60] is well folded and stable. A full assignment of the peaks would have to be made in order to compare the sequence of M2[18-60] produced at FMP with the peak assignments of the S31N mutant.

The discrepancies in the results from RPC, gel electrophoresis and mass spectrometry, show that the purification methods used are not fully optimized for isolating M2[18-60]. In addition, there does not appear to be a large difference between the expression of M2[18-60] at different temperatures, nor the use of a C3 or C4 column during RPC. Aspects that need to be considered for future investigations include analysis of the elution profile of M2[18-60] using IMAC, the optimal conditions for cleavage of the His_6-trpLE domain, and which peaks from RPC should be selected for pooling and reconstitution of M2[18-60]. The latter aspect could potentially be investigated using western blotting, where an antibody targeting M2[18-60] or His_6 could be used in order to determine the presence of each cleavage product in the fraction samples.

This initial investigation into the purification methods for M2[18-60] shows that there is a long way to go before an optimized protocol can be established. The investigations made in this thesis pave the way for future purification of M2[18-60] which can ultimately lead to investigations being made into antiviral drug design and assay development. The use of ssNMR would then be an invaluable tool to study the binding site of the antiviral drug and can be used to further investigate potential antiviral drug development.

9 REFERENCES

1. Laws DD, Bitter H-ML, Jerschow A. Solid-State NMR Spectroscopic Methods in Chemistry. Angewandte Chemie International Edition. 2002;41(17):3096–129.

2. FMP Berlin: Research [Internet]. [cited 2021 Apr 19]. Available from: https://www.leibniz-fmp.de/lange

3. Tomar PPS, Oren R, Krugliak M, Arkin IT. Potential Viroporin Candidates From Pathogenic Viruses Using Bacteria-Based Bioassays. Viruses [Internet]. 2019 Jul 9 [cited 2021 May 20];11(7). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6669592/

4. Moorthy NSHN, Poongavanam V, Pratheepa V. Viral M2 ion channel protein: a promising target for anti-influenza drug discovery. Mini Rev Med Chem. 2014;14(10):819–30.

5. Wang J, Li F, Ma C. Recent progress in designing inhibitors that target the drug-resistant M2 proton channels from the influenza A viruses. Peptide Science. 2015;104(4):291–309.

6. Gan S-W, Tan E, Lin X, Yu D, Wang J, Tan GM-Y, et al. The Small Hydrophobic Protein of the Human Respiratory Syncytial Virus Forms Pentameric Ion Channels *. Journal of Biological Chemistry. 2012 Jul 1;287(29):24671–89.

7. TEV Protease | NEB [Internet]. [cited 2021 Jun 7]. Available from: https://international.neb.com/products/p8112-tev-protease#Product%20Information

8. Fu Q, Piai A, Chen W, Xia K, Chou JJ. Structure determination protocol for transmembrane domain oligomers. Nat Protoc. 2019 Aug;14(8):2483–520.

9. Andreas LB, Reese M, Eddy MT, Gelev V, Ni QZ, Miller EA, et al. Structure and Mechanism of the Influenza A M218–60 Dimer of Dimers. J Am Chem Soc. 2015 Dec 2;137(47):14877–86.

10. Schnell JR, Chou JJ. Structure and Mechanism of the M2 Proton Channel of Influenza A Virus. Nature. 2008 Jan 31;451(7178):591–5.

11. Pielak RM, Chou JJ. Influenza M2 proton channels. Biochimica et Biophysica Acta (BBA) - Biomembranes. 2011 Feb 1;1808(2):522–9.

12. Potter CW. A history of influenza. Journal of Applied Microbiology. 2001;91(4):572–9.

13. Andreas LB, Eddy MT, Pielak RM, Chou J, Griffin RG. Magic Angle Spinning NMR Investigation of Influenza A M2(18−60): Support for an Allosteric Mechanism of Inhibition. J Am Chem Soc. 2010 Aug 18;132(32):10958–60.

14. Bank RPD. RCSB PDB - 2N70. Two-fold symmetric structure of the 18-60 construct of S31N M2 from Influenza A in lipid bilayers [Internet]. [cited 2021 Apr 21];

Available from: https://www.rcsb.org/3d-view/2N70/0

15. Hu F, Luo W, Hong M. Mechanisms of Proton Conduction and Gating in Influenza M2 Proton Channels from Solid-State NMR. Science. 2010 Oct 22;330(6003):505–8.

16. Tang Y, Zaitseva F, Lamb RA, Pinto LH. The Gate of the Influenza Virus M2 Proton Channel Is Formed by a Single Tryptophan Residue. Journal of Biological Chemistry. 2002 Oct;277(42):39880–6.

17. Kass I, Arkin IT. How pH Opens a H+ Channel: The Gating Mechanism of Influenza A M2. Structure. 2005 Dec 1;13(12):1789–98.

18. Pinto LH, Lamb RA. The M2 Proton Channels of Influenza A and B Viruses *. Journal of Biological Chemistry. 2006 Apr 7;281(14):8997–9000.

19. Cook GA, Stefer S, Opella SJ. Expression and purification of the membrane protein p7 from hepatitis C virus. Biopolymers. 2011;96(1):32–40.

20. Hwang PM, Pan JS, Sykes BD. Targeted expression, purification, and cleavage of fusion proteins from inclusion bodies in *Escherichia coli*. FEBS Letters. 2014 Jan 21;588(2):247–52.

21. ExPASy - ProtParam tool [Internet]. [cited 2021 Apr 29]. Available from: https://web.expasy.org/protparam/

22. Chang KY, Yang J-R. Analysis and Prediction of Highly Effective Antiviral Peptides Based on Random Forests. PLoS One [Internet]. 2013 Aug 5 [cited 2021 Apr 29];8(8). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3734225/

23. Amersham Pharmacia Biotech. Reversed Phase Chromatography: Principles and Methods. Amersham Biosciences; 1999.

24. Cytiva. Instrument management handbook, ÄKTA Laboratory-scale Chromatography Systems [Internet]. Cytiva; 2020. Available from: https://cdn.cytivalifesciences.com/dmm3bwsv3/AssetStream.aspx?mediaformatid=10061&destinationid=10016&assetid=16189

25. Bond MD, Panek ME, Zhang Z, Wang D, Mehndiratta P, Zhao H, et al. Evaluation of a dual-wavelength size exclusion HPLC method with improved sensitivity to detect protein aggregates and its use to better characterize degradation pathways of an IgG1 monoclonal antibody. Journal of Pharmaceutical Sciences. 2010;99(6):2582–97.

26. Vydac® 214TP C4 HPLC Columns, Avantor [Internet]. VWR. [cited 2021 May 13]. Available from: https://us.vwr.com/store/product/22799535/vydac-214tp-c4-hplc-columns-avantor

27. nth6x. Property And Principle Of Reversed Phase C4 HPLC Column - Hawach [Internet]. hplccolumn. 2020 [cited 2021 Apr 29]. Available from: https://www.hawachhplccolumn.com/property-and-principle-of-reversed-phase-c4-hplc-column/

28. Difference between C8 and C18 Columns Used in HPLC System : Pharmaceutical Guidelines [Internet]. [cited 2021 May 20]. Available from: https://www.pharmaguideline.com/2018/05/difference-between-c8-and-c18-columns.html

29. Shen H-H, Lithgow T, Martin LL. Reconstitution of Membrane Proteins into Model Membranes: Seeking Better Ways to Retain Protein Activities. Int J Mol Sci. 2013 Jan 14;14(1):1589–607.

30. Andreas LB, Eddy MT, Chou JJ, Griffin RG. Magic-Angle-Spinning NMR of the Drug Resistant S31N M2 Proton Transporter from Influenza A. J Am Chem Soc. 2012 May 2;134(17):7215–8.

31. The principle and method of polyacrylamide gel electrophoresis (SDS-PAGE) | MBL Life Science -JAPAN- [Internet]. [cited 2021 Apr 29]. Available from: https://ruo.mbl.co.jp/bio/e/support/method/sds-page.html

32. Mini-PROTEAN® TGX™ Precast Gels | Life Science Research | Bio-Rad [Internet]. [cited 2021 May 7]. Available from: https://www.bio-rad.com/de-de/product/mini-protean-tgx-precast-gels?ID=N3GRW04VY

33. Tricine Sample Buffer for Protein Gels, 30 ml #1610739 | Life Science Research | Bio-Rad [Internet]. [cited 2021 May 26]. Available from: https://www.bio-rad.com/de-de/sku/1610739-tricine-sample-buffer-for-protein-gels-30-ml?ID=1610739

34. Overview of Mass Spectrometry - DE [Internet]. [cited 2021 Apr 29]. Available from: //www.thermofisher.com/de/de/home/life-science/protein-biology/protein-biology-learning-center/protein-biology-resource-library/pierce-protein-methods/overview-mass-spectrometry.html

35. Kitson FG, Larsen BS, McEwen CN. Chapter 1 - What Is GC/MS? In: Kitson FG, Larsen BS, McEwen CN, editors. Gas Chromatography and Mass Spectrometry [Internet]. San Diego: Academic Press; 1996 [cited 2021 May 13]. p. 3–23. Available from: https://www.sciencedirect.com/science/article/pii/B9780124833852500026

36. Cohen SL, Chait BT. Mass Spectrometry of Whole Proteins Eluted from Sodium Dodecyl Sulfate–Polyacrylamide Gel Electrophoresis Gels. Analytical Biochemistry. 1997 May;247(2):257–67.

37. Paulo JA. Sample preparation for proteomic analysis using a GeLC-MS/MS strategy. J Biol Methods [Internet]. 2016 Jul 12 [cited 2021 Apr 29];3(3). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4959781/

38. Krishnan V, Rupp B. Macromolecular Structure Determination: Comparison of X-ray Crystallography and NMR Spectroscopy. In 2012.

39. Keeler J. Understanding NMR spectroscopy. 2nd ed. Chichester,

U.K: John Wiley and Sons; 2010. 511 p.

40. Patching S. NMR-Active Nuclei for Biological and Biomedical Applications. J Diagn Imaging Ther. 2016 Jun 18;3(1):7–48.

41. Hornak JP. The Basics of NMR [Internet]. 1997. Available from: http://www.cis.rit.edu/htbooks/nmr/bnmr.htm

42. Hoffmann J, Ruta J, Shi C, Hendriks K, Chevelkov V, Franks WT, et al. Protein resonance assignment by BSH-CP-based 3D solid-state NMR experiments: A practical guide. Magnetic Resonance in Chemistry. 2020;58(5):445–65.

43. 13.4: Chemical Shifts in ^{1}H NMR Spectroscopy [Internet]. Chemistry LibreTexts. 2015 [cited 2021 Feb 17]. Available from: https://chem.libretexts.org/Bookshelves/Organic_Chemistry/Map%3A_Organic_Chemistry_(McMurry)/13%3A_Structure_Determination_-_Nuclear_Magnetic_Resonance_Spectroscopy/13.04%3A_Chemical_Shifts_in_H_NMR__Spectroscopy

44. Levitt MH. Spin dynamics: basics of nuclear magnetic resonance. 2nd ed. Chichester, England ; Hoboken, NJ: John Wiley & Sons; 2008. 714 p.

45. Jardetzky O, Roberts GCK. NMR in Molecular Biology. New York: Academic Press; 1981. 681 p. (Molecular biology).

46. Szeverenyi NM, Sullivan MJ, Maciel GE. Observation of spin exchange by two-dimensional fourier transform 13C cross polarization-magic-angle spinning. Journal of Magnetic Resonance (1969). 1982 May;47(3):462–75.

47. Horst J. Two Dimensional NMR Spectroscopy [Internet]. 1996 [cited 2021 Jun 11]. Available from: http://www.cryst.bbk.ac.uk/PPS2/projects/schirra/html/2dnmr.htm

48. BMRB - Biological Magnetic Resonance Bank [Internet]. [cited 2021 May 17]. Available from: https://bmrb.io/ref_info/csstats.php

49. Reif B, Ashbrook SE, Emsley L, Hong M. Solid-state NMR spectroscopy. Nature Reviews Methods Primers. 2021 Jan 14;1(1):1–23.

50. Ladizhansky V. Applications of solid-state NMR to membrane proteins. Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics. 2017 Nov 1;1865(11, Part B):1577–86.

51. Webb A. Increasing the Sensitivity of Magnetic Resonance Spectroscopy and Imaging. Anal Chem. 2012 Jan 3;84(1):9–16.

52. NMR spectrometer of superlatives [Internet]. Max-Planck-Gesellschaft. 2020 [cited 2021 Feb 17]. Available from: https://www.mpg.de/15191397/0721-bich-056839-eines-der-drei-staerksten-hochaufloesenden-1-2-ghz-nmr-spektrometer-weltweit-steht-nun-in-goettingen

53. Nuclear Magnetic Resonance Instrumentation - Bai - - Major Reference Works - Wiley Online Library [Internet]. [cited 2021 May 17]. Available from: https://onlinelibrary.wiley.com/doi/

abs/10.1002/9780470027318.a6108.pub3

54. Niles S. Site Planning for AVANCE Systems 300-700 MHz. Bruker Biospin GmbH. 2008;106.

55. His-tagged Proteins–Production and Purification - DE [Internet]. [cited 2021 Apr 29]. Available from: //www.thermofisher.com/de/de/home/life-science/protein-biology/protein-biology-learning-center/protein-biology-resource-library/pierce-protein-methods/his-tagged-proteins-production-purification.html

56. Spectra/Por 3 Dialysis Tubing, 3.5 kD MWCO [Internet]. Repligen. [cited 2021 May 12]. Available from: https://store.repligen.com/products/spectra-por3dialysistubing-3-5kdmwco-18mmflat-width-15meters-roll50ft

57. Nanodrop micro-volume spectrophotometers - Isogen Lifescience [Internet]. [cited 2021 May 17]. Available from: https://isogen-lifescience.com/products/instruments/nanodrop-micro-volume-spectrophometers?gclid=CjwKCAjwhMmEBhBwEiwAXwFoEab5A0q2JV939psUUUb-PSK7m6ixt2MSQY3YXoYmi9hC-fzarVT9UhoCnDEQAvD_BwE

58. UNICORN™ start software | VWR [Internet]. [cited 2021 May 7]. Available from: https://us.vwr.com/cms/life_science_ge_unicorn_start

59. PageRuler ™ Prestained Protein Ladder, 10 to 180 kDa [Internet]. [cited 2021 May 7]. Available from: https://www.thermofisher.com/order/catalog/product/26616

60. Expedeon InstantBlue™ Stain Protein Stain Vol.: 1L Molecular Weight Markers | Fisher Scientific [Internet]. [cited 2021 May 7]. Available from: https://www.fishersci.co.uk/shop/products/expedeon-instantblue-stain-protein-marker/10616474

61. Heidolph Instruments: Polymax 1040 (10° tilt angle) [Internet]. [cited 2021 May 7]. Available from: https://heidolph-instruments.com/en/products/Shakers-Mixers/Polymax-1040-10deg-tilt-angle~p1209

62. Welcome to CCPN — Collaborative Computational Project for NMR [Internet]. [cited 2021 May 17]. Available from: https://www.ccpn.ac.uk/

63. Andreas L, Reese M, Eddy M, Gelev V, Ni QZ, Miller E, et al. Two-fold symmetric structure of the 18-60 construct of S31N M2 from Influenza A in lipid bilayers [Internet]. Biological Magnetic Resonance Bank; 2016 [cited 2021 May 24]. Available from: https://bmrb.io/data_library/summary/?bmrbId=25788

64. TopSpin | NMR Data Analysis | Bruker [Internet]. [cited 2021 May 17]. Available from: https://www.bruker.com/en/products-and-solutions/mr/nmr-software/topspin.html?gclid=CjwKCAjwqIiFBhAHEiwANg9szrdgdVA_OV65DJBkcjl8W73CPOvtcuRt8_StIRL0R1AKET0g4WS5ahoCkUwQAvD_BwE