UNIVERSITY OF BIRMINGHAM

Exploring *Mycobacterium Tuberculosis* Cell Envelope: Studying Proteins and Lipids Using Styrene Maleic Acid Polymer

By

Khaled Ali E. Haidar

A thesis submitted to the University of Birmingham for the degree of

DOCTOR OF PHILOSOPHY

School of Biosciences

College of Life and Environmental Sciences

University of Birmingham, UK

December 2023

Abstract

In 2022, *tuberculosis* TB caused 1.3 million deaths with nearly 10 million new cases reported, emphasizing the urgent need for new drugs targeting different pathways. The cell envelope of *Mycobacterium tuberculosis* is an excellent target for novel therapeutics due to its complex architecture, vital for pathogenicity, virulence, and drug resistance. Precise knowledge of molecular interactions on the *M. tuberculosis* envelope surface is key to identifying crucial components. Current analytical methods involve harsh organic solvents that don't selectively extract outer envelope components. Polymer nano-encapsulation techniques offer promise in studying membrane proteins. Amphipathic polymers like Styrene Maleic Acid (SMA) have been extensively used to directly solubilize proteins and lipids from the cytoplasmic membrane via SMA Lipid Particles (SMALPs).

This study presents a protocol for solubilizing and purifying three complexes of an ATP-binding protein ABC transporter from *Mycobacterium tuberculosis*, predicted to be involved in drug transport as an efflux system. Additionally, this study demonstrates that SMALPs selectively extract peripheral molecules from the cell envelope of mycobacteria. Using [^{14} C]-uniformly labelled lipids, we show that SMA can extract complex molecules like glycolipids (trehalose monomycolate, trehalose dimycolate, and glucose monomycolate), phospholipids (phosphatidyl inositol mannosides), lipoglycans (lipomannan and lipoarabinomannan), and surface-exposed proteins from the envelopes of Mycobacterium smegmatis without causing cell rupture.

Acknowledgments

I want to express my gratitude to Professor Tim Dafforn, my supervisor, for providing me with this opportunity and for his continuous guidance and passion for scientific endeavours. I am thankful to all past and present members of the Dafforn lab, especially Dr. Naomi Pollock, for her support and friendship throughout the past four years. I would like to thank my student Emily Piper for her contribution.

The 7th-floor labs at the University of Birmingham have been an incredible place, thanks to the outstanding PhD students and postdocs who have become my family. Each member has played a crucial role in making the completion of my PhD possible. There are some individuals who deserve special recognition for their support during my time in the lab. Dr. Stephanie Nestorow, I am truly grateful for your guidance, advice, and for always being welcoming and supportive. Also, for the AKTA training and willingness to address all my questions, no matter how trivial, have been immensely helpful. Dr. Amber Wilson, thank you for patiently listening to me ramble about my football team every time we chat, and for the Saturday walks, which were incredibly beneficial. (By the way, I am still waiting for the banana bread!). There are numerous others from Biosciences whom I'd like to thank: Poojha Sridhar, Dr. Peter Weatherspoon, Liam Mead, Rebecca Parr, Adam Evans, Bethan Kelly, Dr. Giedre Ratkeviciute, Dr. Ben Cooper, Dr. Simon Cualton, Dr. Gareth Hughes, Dr. David Hardy, Dr. Racheal Grime, Rachel Holyfield, Hannah Johnston, Daisy Marshall, and many more. Also, a big thanks to everyone I collaborate with on E106. You will all be greatly missed.

Immense gratitude goes out to my friends who made this journey both enjoyable and bearable. Dr. Uthman Alqopishi, thank you for the Alfursan dinner every Sunday. Ahmed Almasmuom, your volleyball games were a great stress reliever. Akram Sultan, your afternoon breaks during the entire PhD were much appreciated, and I'll see you in Bisha. Dr. Wadia Alruqayb, Fisal bin Hemaeed Dr. Mohammed Althagafi, Dr. Hassan Labni, Dr. Akram Alaufy, Mohammed Alkathami, Anas Bokari, Ahmed Algamdi, Dr. Ali Kabli, and to all the friends I made during this wonderful journey, your support has meant the world to me."

Finally, I would like to thank my family, especially my parents. I am incredibly grateful for your support during this challenging time when being away was quite difficult. I missed you tremendously. Your constant presence and encouragement have been instrumental in my success and achieving my life goals. I also want to express my gratitude to my brothers and sisters—Wejdan, Hassan, Rajeh, Wedad, Ahmed, Abdullah, and finally, Emad. You are always looking out for me, and Emad, though you're no longer with us, you will forever be remembered as my beloved younger brother."

Table of Contents

List Of Abbreviations

2D	Two- dimensional
ABC	ATB binding cassette
AG	Arabinogalactan
AUC	Analytical ultracentrifugation
CD	Circular dichroism
DAT	Diacyl-trehalose
DLS	Dynamic light scattering
E. coli	*Escherichia coli*
ECM	Extracellular matrix
EM	Electron microscopy
ETC	Electron transport chain
FT	Flow through
GFP	Green Fluorescent Protein
GlcNAc	N-acetyl- glucosamine
GMM	Glucose monomycolate
GPCRs	G-protein coupled receptors
GPL	Glycopeptidolipid
HSPs	Heat shock proteins
Ism	Insoluble material
LAM	Lipoarabinomannan
LB	Luria Bertani
LM	Lipomannan
LOSs	Lipoligosaccharides
LTBI	Latent TB infection
mAGP	Mycolyl-arabinogalactan-peptidoglycan
Man-LAM	Mannose-capped lipoarabinomannan
MAs	Mycolic acids
Mb	Membrane
MDR-TB	Multi-Drug Resistant tuberculosis
MP	Membrane protein
MSDs	Membrane spanning domain
Mtb	*Mycobacterium tuberculosis*
MurNAc	N-acetyl-muramic acid
MurNGlyc	N-glycolylmuramic acid
MWCO	Molecular weight cut-off
NBDs	Nucleotide binding domains
PAT	Pentaacyl- trehalose
PG	Peptidoglycan
PGLs	Phenolic glycolipids (PGLs)
PI	Phosphatidylinositol
PI3P	Phosphatidylinositol 3-phosphate
PIMs	phosphatidyl-myo-inositol mannose
pre-XDR-TB	Pre-Extensively-Drug resistant tuberculosis
PtkA	Protein tyrosine kinase A
RR-TB	Rifampicin-resistant TB

SAR	Scavenger receptor A
SDS-PAGE	Sodium dodecyl-sulfate polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SL	Sulfated trehalose or sulfolipids
SMA	Poly (styrene-co-maleic acid)
SMALPs	Poly (styrene-co-maleic acid) lipid particles
Smb	Solubilised membrane
TAT	Triacyl-trehalose
TB	Terrific broth
TB	Tuberculosis
TDM	Trehalose dimycolate
TLC	Thin layer chromatography
TMDs	Transmembrane domain
TMM	Trehalose monomycolate
TSB	Tryptic Soy Broth
XDR-TB	Extensively-Drug resistant tuberculosis

CHAPTER 1

GENRAL INTRODUCTION

1.1 *Mycobacterium tuberculosis*

Mycobacterium tuberculosis was first discovered and identified in 1882 via German physician and microbiologist Robert Koch. *M. tuberculosis* is an aerobic bacillus, slow-growing, non- motile, non-spore-forming, chemoorganotrophic bacteria, that multiplies every 24 hours under optimal laboratory conditions at 37°C (Gordon and Parish, 2018). Tuberculosis (TB) infection occurs via inhaling droplet nuclei containing *M. tuberculosis* from an infected to a susceptible person, droplet nuclei reach the lungs alveoli crossing the mouth or nasal cavity passages, upper respiratory tract, and bronchi. However, although TB is originally a lung infection referred to as pulmonary TB ; it may also expand to other organs and tissue (Bussi and Gutierrez, 2019). The most common symptoms of pulmonary TB are chronic cough, fever, fatigue, and weight loss (Palanivel *et al*., 2023) .Globally, about 7.5 million Tuberculosis (TB) cases were recorded annually, with 1.3 million mortality rate in 2022, according to the World Health Organization (WHO) report (World Health Organization, 2023). After discovering several drugs to treat TB, a multi-drug regimen was adopted to ensure quicker and more effective recovery., this global emergency was exacerbated by the rise in Multi-Drug Resistant tuberculosis (MDR-TB) and Extensively-Drug resistant tuberculosis (XDR-TB) strains of the bacterium (Sundararajan and Muniyan, 2021).

Moreover, TB has the ability of infecting the host without developing any symptoms and without reproduction, this called latent TB infection (LTBI), term firstly called by Austrian pediatrician Clemens von Pirquet (Palanivel *et al*., 2023, Migliori *et al*., 2021). When the immune system of the host fail to controlling the bacterium or defending itself, the *Mtb* develop an active infection, unless this occurs the *Mtb* remain in the LTBI state (Tang and Johnston, 2017). Over quarter of the world's population carries an infection of TB, and around 5 – 10 % of them will advance to develop active tuberculosis and the rest will stay in LTBI state (Palanivel *et al*., 2023).

1.2 Mechanism of infection

Infection with *Mtb* typically begins through the inhalation of droplets containing a small number of bacilli (Kaufmann, 2001). Once in the lungs, alveolar macrophages take up the bacilli through a process called phagocytosis. These macrophages usually attempt to eliminate the internalized *Mtb* by exposing them to a destructive environment within lysosomes. However, some bacilli may evade destruction during this process and manage to survive within the macrophages (Armstrong and Hart, 1975, Russell, 2001, Kaufmann, 2001). *Mtb* uses various receptors to enter macrophages, such as complement receptors, mannose receptors, the dendritic cell-specific intercellular adhesion molecule (ICAM)-3-grabbing nonintegrin (DC-SIGN), and Fc receptors (Pieters, 2008). The presence of multiple receptors employed by *Mtb* to invade macrophages likely correlates with the unique and complex structure of the *Mtb* cell surface (Brennan and Nikaido, 1995). When *Mtb* enters through the CR3 receptor, it relies on the cholesterol in the host cell membrane to prevent the fusion of phagosomes with lysosomes (Gatfield and Pieters, 2000).

Additionally, the selective internalization of *Mtb* through the binding with such specific receptors is a method to evade degradation within the macrophage and enhances the probability of its survival (Pieters, 2008). *Mtb* enhances its survival inside the macrophage by obstructing phosphatidylinositol 3-phosphate (PI3P), a crucial host membrane component essential for building phagolysosomes (Roth, 2004). The interference occurs during the delivery of the phagocytosed cargo to the lysosome, which relies on PI3P this obstruction prevents the accumulation of PI3P on the phagosomal membrane (Fratti *et al*., 2003). It is achieved by influencing the activity of the PI3 kinase hVP34, preventing the secretion of PI3P onto the phagosomal membrane(Vergne *et al*., 2003).

Additionally, *Mtb* activates SapM acid phosphatases, which hydrolyse PI3P, preventing its accumulation on the phagosomal membrane (Saleh and Belisle, 2000). Additionally, *Mtb* releases soluble molecules of PknG kinase upon its internalization within macrophage phagosomes to aid in survival. This molecule gains access to the cytosol, although the exact mechanism of this translocation remains uncharacterized or unknown as a consequence, it inhibits the formation of phagolysosomes (Walburger *et al*., 2004).

Protein analysis of mycobacterial phagosomes indicates the presence of a protein, coronin 1, which isn't typically linked to any specific cellular structure but is notably abundant in phagosomes containing viable mycobacteria. Coronin 1 functions by preventing lysosome delivery and preventing the death of mycobacteria within macrophages through the activation of the calcium-dependent phosphatase calcineurin (Jayachandran *et al*., 2007, Hasan *et al*., 1997, Gatfield *et al*., 2005, Ferrari *et al*., 1999). *Mtb* employs another strategy for survival within macrophages by hindering phagocytosis maturation, resulting in a decrease of the pH environment to around 6.2. This is achieved through alterations in the macrophage cell wall structure and specific molecules that maintain a natural pH within the macrophage. The presence of protein tyrosine kinase A (PtkA) is significant in enhancing phagolysosome acidification (Zhai *et al*., 2019; Wong *et al*., 2018).

Additionally, *Mtb* has devised a survival strategy within macrophages by initiating granuloma formation—a clustered structure of *Mtb* within macrophages surrounded by various immune cells, notably T-lymphocytes, non-active macrophages, and active ones coexisting in the same environment. Their interactions establish an ongoing immune response at the infection site, offering protection to the host from infection as long as the host's immunity remains functional, primarily via activated macrophages and effective T cells (Pieters, 2008, Karakousis *et al*., 2004).

1.3 TB worldwide

Despite a two-year disruption caused by the pandemic, there was a noticeable recovery in the number of people diagnosed and treated for TB in 2022 (World Health Organization, 2023). However, despite this global recovery, TB remains one of the primary causes of death from a single infectious agent in 2022 (World Health Organization, 2023). In 2022, Southeast Asia recorded the highest number of individuals developing TB, accounting for 46%, followed by Africa with 23%, the Western Pacific with 18%, the Eastern Mediterranean with 8.1%, the Americas with 3.1%, and finally Europe with 2.2%. Among all TB cases reported in 2022, approximately 6.3% were among individuals suffering from HIV (World Health Organization, 2023).

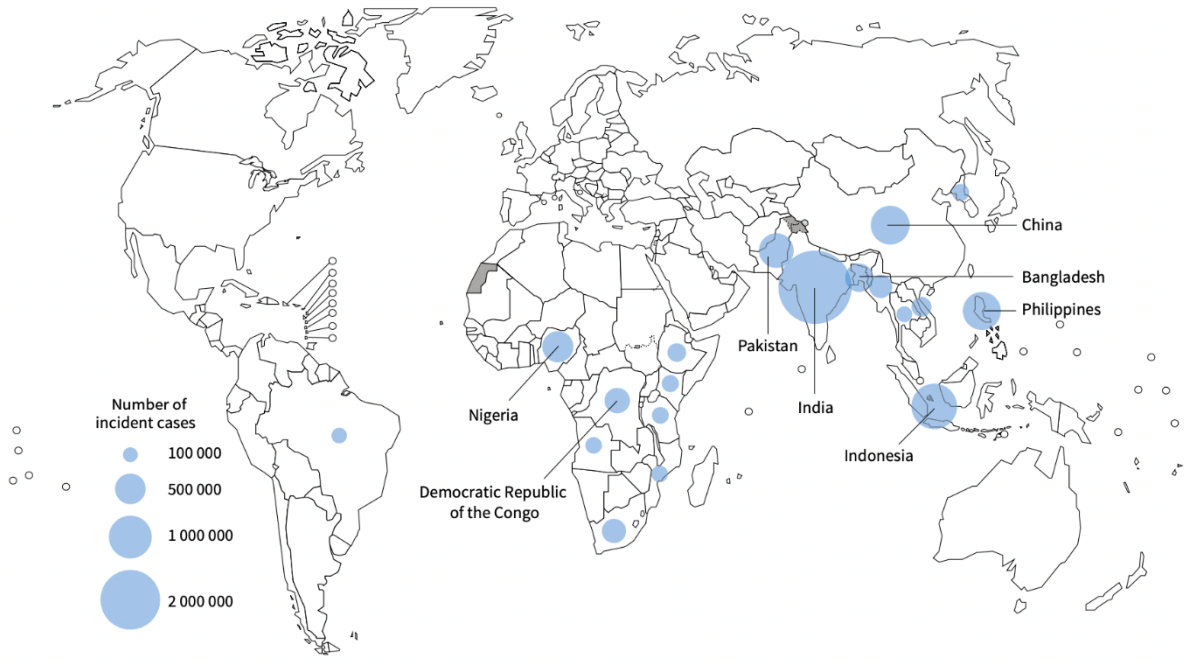


Figure 1.1 The estimated number of incident TB cases in 2022. Highlighting the top countries with at least 100,000 reported cases. Figure reproduced from World Health Organization, 2023.

Moreover, drug-resistant tuberculosis presents a significant concern in public health due to the growing resistance to primary medications and the declining efficacy of essential drugs, causing alarm in the global health community. Multidrug-resistant TB (MDR-TB) occurs when tuberculosis is resistant to both rifampicin and isoniazid, requiring treatment with alternative medications. Rifampicin-resistant TB (RR-TB) specifically denotes cases resistant solely to rifampicin, necessitating a different treatment approach. India, the Philippines, and the Russian Federation collectively account for about 42% of the global cases of individuals developing MDR/RR-TB (figure 1.2) (World Health Organization, 2023).

Figure 1.2 The estimated number of people who developed MDR/RR-TB in 2022. Highlighting the top countries with at least 100,000 reported cases. Figure reproduced from World Health Organization, 2023.

1.4 Treatment of TB

Mtb, a highly pathogenic organism, accounted for approximately 1.30 million deaths in 2022, with a rising number of individuals developing multi-drug-resistant or rifampicin-resistant TB (MDR/RR-TB) (World Health Organization, 2023). Over the past 5 years, around 34 million individuals with TB have been treated, including 2.5 million children and 825,000 people with drug-resistant strains (World Health Organization, 2023). The standard TB treatment involves a 6-month regimen comprising four drugs intensively administered in the initial 2 months—isoniazid, rifampicin, pyrazinamide, and ethambutol—followed by 4 months of isoniazid and rifampicin (Zumla *et al*., 2013). With the emergence of multi-drug-resistant TB, WHO identifies five categories: MDR-TB, RR-TB, isoniazid-resistant TB, extensively drug-resistant TB (XDR-TB), and pre-XDR-TB (World Health Organization, 2023). XDR-TB occurs when the bacteria are resistant to rifampicin, any fluoroquinolone, and either bedaquiline or linezolid. Pre-XDR-TB is resistant to rifampicin and any fluoroquinolone (World Health Organization, 2023).

The distinct structure of the mycobacterial cell wall has garnered attention as a potential drug target. Through comprehensive investigations into cell wall assembly, particularly via whole-genome sequencing, researchers aim to gain deeper insights into mycobacterial cell wall biosynthesis. This understanding could pave the way for discovering and developing chemotherapeutic agents that target the synthesis of the unique macromolecular structure in *Mtb* (Abrahams and Besra, 2018). Targeting the processes involved in constructing and maintaining the bacterial cell wall is considered a promising strategy for developing new antibacterial treatments because this structure differs from those found in mammalian cells. For many years, understanding and studying the production and maintenance of the cell wall has been a major area of focus. Drugs used in tuberculosis treatment, such as Ethambutol, Isoniazid, and Ethionamide, function by inhibiting these cell wall processes (Table 1.1) (Takeya and Hisatsune, 1963).

Table 1.1: Antimycobacterial Drugs: Providing a breakdown of the discovery year and the mechanism of action for both first and second line antimycobacterial drugs. The table has been reproduced with permission from Zumla *et al*., 2013.

Drug (year of discovery)	Target	Effect
First-line drugs		
Isoniazid (1952)	Enoyl-[acyl-carrier-protein] reductase	Inhibits mycolic acid synthesis
Rifampicin (1963)	RNA polymerase, beta subunit	Inhibits transcription
Pyrazinamide (1954)	S1 component of 30S ribosomal subunit	Inhibits translation and *trans*-translation, acidifies cytoplasm
Ethambutol (1961)	Arabinosyl transferases	Inhibits arabinogalactan biosynthesis
Second-line drugs		
***Para*-amino salicylic acid (1948)**	Dihydropteroate synthase	Inhibits folate biosynthesis
Streptomycin (1944)	S12 and 16S rRNA components of 30S ribosomal subunit	Inhibits protein synthesis
Ethionamide (1961)	Enoyl-[acyl-carrier-protein] reductase	Inhibits mycolic acid biosynthesis
Ofloxacin (1980)	DNA gyrase and DNA topoisomerase	Inhibits DNA supercoiling
Capreomycin (1963)	Interbridge B2a between 30S and 50S ribosomal subunits	Inhibits protein synthesis
Kanamycin (1957)	30S ribosomal subunit	Inhibits protein synthesis
Amikacin (1972)	30S ribosomal subunit	Inhibits protein synthesis
Cycloserine (1955)	d-alanine racemase and ligase	Inhibits peptidoglycan synthesis

1.5 Mycobacteria cell envelope

The mycobacterial cell envelope, as depicted in Figure 1.3 is a specialized and multi-layered structure(Abrahams and Besra, 2018). Bacterial cell envelopes are pivotal in shielding bacteria from harsh environments and aiding their adherence to host cells (Vinod *et al*., 2020). Additionally, the cell wall provides the bacteria with structural strength, facilitating the transport of solutes and proteins (Vinod *et al*., 2020). Unlike gram-positive and gram-negative bacteria (Figure 1.4), mycobacteria possess a distinctive cell envelope, consisting of four primary layers above the plasma membrane: peptidoglycan, arabinogalactan, mycolic acid, and a peripheral lipid layer (Figure 1.3) (Garcia-Vilanova *et al*., 2019, Alderwick *et al*., 2015).

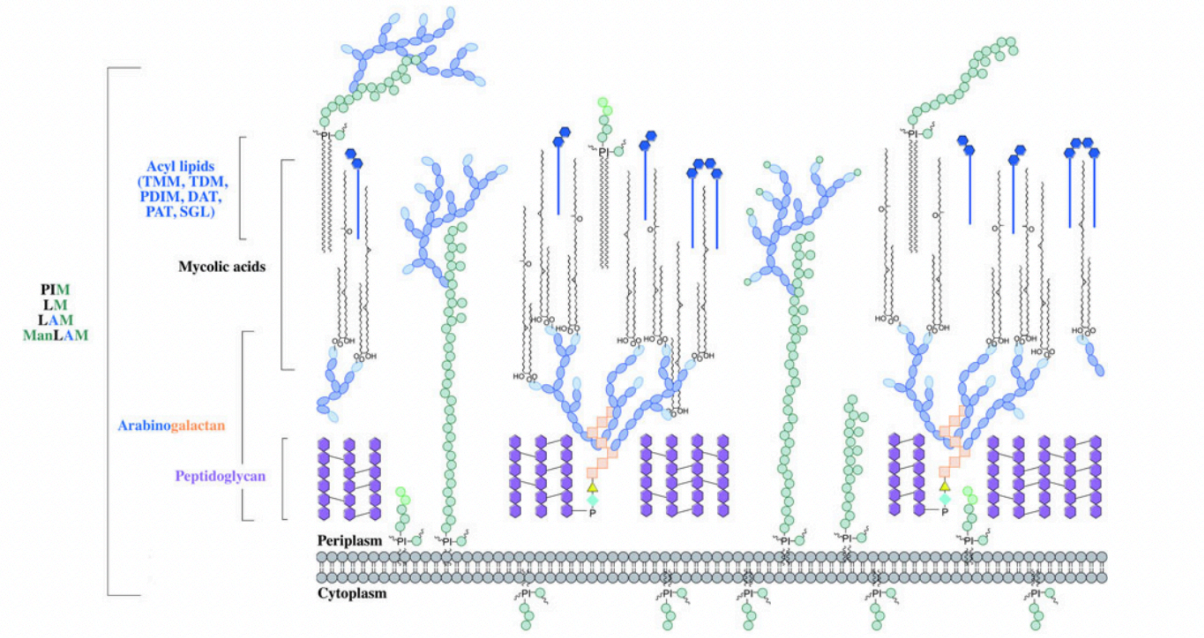


Figure 1.3: Cross-section model of mycobacterial cell wall. Highlighting: the peptidoglycan, arabinogalactan, mycolic acid, lipomannan, lipoarabinomannan, mannosylated lipoarabinomannan, and surface glycolipids. Figure reproduced with permission, from Abrahams and Besra, 2018.

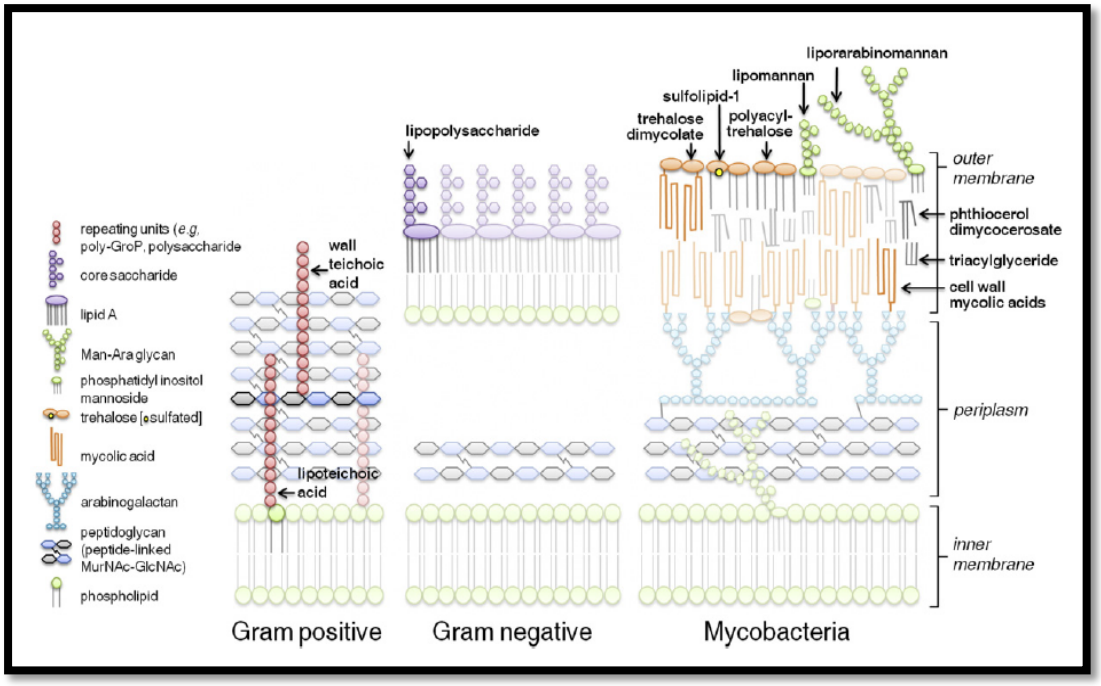


Figure 1.4: Comparing the components of bacterial cell envelopes among Gram-positive, Gram-negative, and Mycobacteria. Figure reproduce with permission from (Touchette and Seeliger, 2017)

Peptidoglycan (PG) constitutes glycan chains linked by short peptide chains, serving as the structural foundation. Below this layer lies the cytoplasmic membrane, a double layer of phospholipids enveloping the cell and serving as its core (Moynihan *et al*., 2019, Jankute *et al*., 2015). Positioned above the PG layer is the arabinogalactan (AG), a polysaccharide made of galactofuranose residues forming a galactan chain followed by a branching arabinofuranose polysaccharide (Besra *et al*., 1995, Alderwick *et al*., 2005). The inner membrane is covalently connected to the AG, composed of mycolic acid and fatty acids, while the outer part encompasses various derivatives of mycolic acid species (Touchette and Seeliger, 2017, Bansal-Mutalik and Nikaido, 2014), alongside lipoglycans: lipomannan (LM) and lipoarabinomannan (LAM), scattered within the inner membrane. LM and LAM are polysaccharides anchored by lipids, structured from a phosphatidylinositol enhanced with mannose for LM, further enriched with arabinose to produce LAM (Pitarque *et al*., 2005, Mishra *et al*., 2007, Cashmore *et al*., 2017, Touchette and Seeliger, 2017). Finally, the outermost layer is the capsule, a loosely structured assembly primarily containing

polysaccharides such as alpha-glucan, arabinomannan, and mannan, in addition to secreted proteins and lipids (Kalscheuer *et al*., 2019, Koliwer-Brandl *et al*., 2016).

1.5.1 Peptidoglycan

Peptidoglycan (PG) is present in most eubacteria, comprising a complex macromolecular structure located outside the plasma membrane (Schleifer and Kandler, 1972). Its mesh-like arrangement provides strength to the cell, allowing it to withstand osmotic pressure, maintain cellular integrity, and retain its shape (van Heijenoort, 2001). Consequently, it plays a pivotal role in bacterial growth and survival (Vollmer *et al*., 2008). In mycobacteria, PG forms part of the mycolyl-arabinogalactan-peptidoglycan (mAGP) complex. This structure involves alternating N-acetyl- glucosamine (GlcNAc) and N-acetyl-muramic acid (MurNAc) residues, connected in a β (1→4) configuration with side chains of amino acids cross-linked by a trans peptide bridge (Zumla *et al*., 2013, Lederer *et al*., 1975, Brennan and Nikaido, 1995). Moreover, mycobacterial PG distinguishes itself from typical structures due to distinct attributes, such as the presence of N-acetyl-muramic acid (MurNAc) and N-glycolylmuramic acid (MurNGlyc) residues, amidation of carboxylic acids in peptide stems (Mahapatra *et al*., 2005), and the incorporation of additional glycine or serine residues (Vollmer *et al*., 2008).

1.5.2 Arabinogalactan

The arabinogalactan (AG) sits atop the peptidoglycan (PG) layer and, as its name suggests, is a complex heteropolysaccharide consisting of arabinose and galactose, comprising roughly 35% of the cell envelope's total mass. It is covalently linked, via a phosphodiester bond, to approximately 10%-12% of the muramic acid residues within the peptidoglycan. Together, the PG and AG form a substantial macromolecular structure situated between the inner membrane and the outer mycomembrane (mycolic acid layer) (Vinod *et al*., 2020, Alderwick *et al*., 2015). Studies have unveiled the AG's primary structure, revealing it as a linear galactan composed of roughly 30 alternating β-(1-5)-linked and β-(1-6)-linked D-

galactofuranose (Galf) residues. Additionally, up to two extensively branched arabinan chains (each containing 31 D-arabinofuranose (Araf) residues) are attached to this galactan backbone, rendering AG a pivotal element in the contagiousness and pathogenicity of *M. tuberculosis* (Wu *et al*., 2017). The process of creating this structure happens in two stages: first, the formation of the galactan core occurs on the inner part of the plasma membrane, followed by the addition of arabinose to the galactan in the periplasm (Jankute *et al*., 2015, Abrahams and Besra, 2018).

1.5.3 Mycolic Acid

According to Anderson, mycolic acids (MAs) are a series of related β-hydroxy fatty acids featuring an elongated α-alkyl side chain that varies by 28 atomic mass units (Anderson, 1929). In mycobacteria, mycolic acids are essential for viability, contributing to maintaining the hydrophobic nature of the external cell surface, particularly in virulent strains where this impermeable barrier is crucial for preventing the uptake of various antibiotics (Puech *et al*., 2001, Liu *et al*., 1996). These MAs are covalently linked to the unique peptidoglycan-arabinogalactan complex of the mycobacterial cell wall (Nataraj *et al*., 2015), or can be in the outer membrane as free mycolates or they can form compounds like trehalose monomycolate (TMM) or trehalose dimycolate (TDM) by combining with various sugar-containing structures like trehalose (Butler *et al*., 1986, Bhatt *et al*., 2007, Bhatt *et al*., 2008). In the case of *M. tuberculosis*, these are highly hydrophobic fatty acids ranging from C54 to C63, with an α side chain of C22 to C24 (Takayama *et al*., 2005). Within *M. tuberculosis*, three types of MAs have been identified: α-mycolates, methoxymycolates, and ketomycolates, where alpha-mycolates are the most prevalent (>70%), followed by methoxymycolates (10%-15%) and ketomycolates (figure 1.5) (Takayama *et al*., 2005, Minnikin *et al*., 2015b). Structurally, MAs can adopt various folded configurations. Ketomycolates in *M. tuberculosis* typically adopt a consolidated W-shaped confirmation with four chains in parallel. In contrast, α-mycolates assume a completely extended U-shaped confirmation with two elongated distal chains. Methoxymycolates, however, can fold into either W or U shapes, as well as other configurations like Z and sZ (Minnikin *et al*., 2015b).

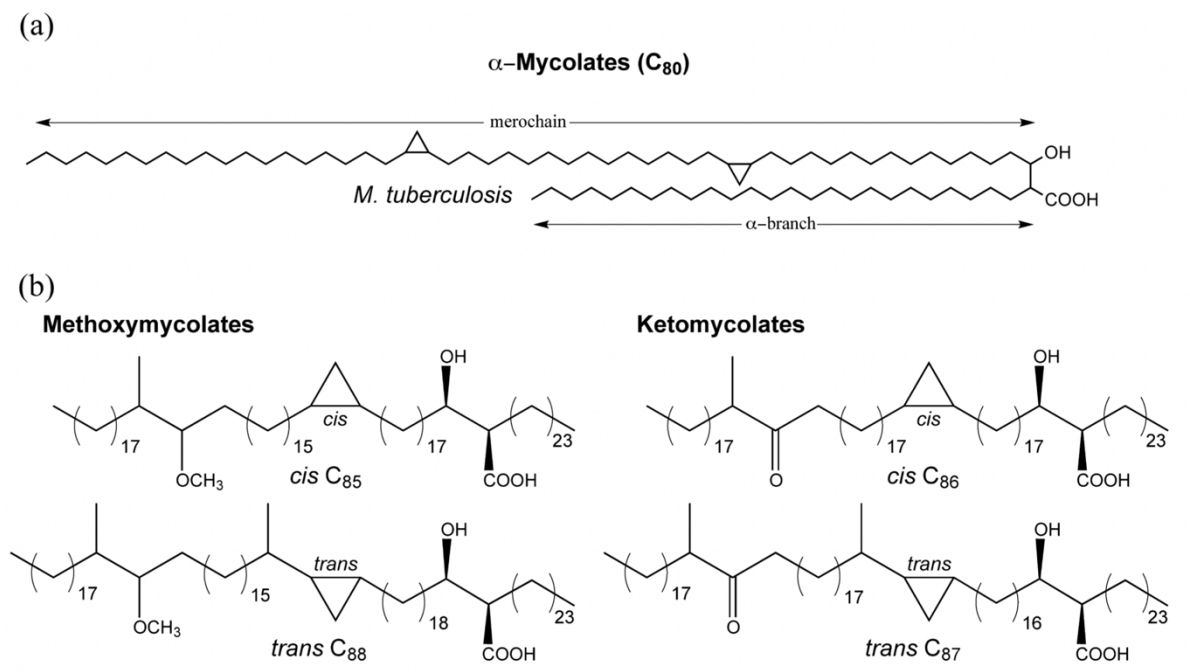


Figure 1.5 structure of mycobacterial mycolic acid. Figure reproduced with permission from Nataraj *et al* 2015.

1.5.4 Peripheral lipid layer

The exploration of tubercle bacillus began more than 90 years ago with Anderson *et al*. In 1929, they reported obtaining two fractions upon purifying the crude wax from the chloroform extract of the human type of tubercle bacillus, Strain H-37. These were a white solid substance labelled as purified wax and a soft yellowish-brown salve-like mass referred to as soft wax, possessing a faint yet pleasant perfume-like scent (Garcia-Vilanova *et al*., 2019, Anderson, 1929). Since then, glycolipids have become the most extensively studied lipids of *M. tuberculosis* due to their significance in the cell envelope's composition, known for their polarity, toxicity, and immunological properties. Among these, the acyltrehalose family has gained prominence, which includes several glucolipids like di-, tri-, and penta-acyltrehaloses (DATs/TATs/PATs), mono and di-mycolyl trehaloses (TMM/TDM), tetraacyl-trehalose sulfo(glyco)lipids (SLs), lipooligosaccharides (LOSs), phosphatidyl-myo-inositol mannose derivatives (PIMs), and phenolic glycolipids (PGLs) found in specific *M. tuberculosis* strains (Rhoades *et al*., 2011, Garcia-Vilanova *et al*., 2019, Brennan and Nikaido, 1995).

1.5.4.1 The DAT/TAT/PAT glycolipids family

Research into glycolipid families, such as diacyl- (DAT), triacyl- (TAT), and pentaacyl-trehalose (PAT), featuring polymethylated acyl groups in their chemical structure, has identified over 30 molecular species within the *M. tuberculosis* cell envelope (Gautier *et al*., 1992, Frankfater *et al*., 2019). Despite this, their specific roles within the cell envelope are still being investigated. Studies have revealed certain characteristics of these glycolipids. The DAT family, for instance, has been observed not to participate in the receptor-mediated phagocytosis of *M. tuberculosis* and does not amplify host cytotoxicity. However, members of the DAT family have been shown to induce the production of IL-10 and can restrict T-cell proliferation (Saavedra *et al*., 2001). Furthermore, the DAT family has been found to negatively regulate the host's pro-inflammatory response (Lee *et al*., 2007). Moreover, research suggests that both DATs and PATs collaborate to maintain the integrity of the *M. tuberculosis* cell envelope, enhancing its pathogenicity and aiding its survival within the host cell (Bailo *et al*., 2015). This preservation contributes to the antigenicity of *M. tuberculosis* (Papa *et al*., 1989). Studies propose that DATs might originally function as regulators of the *M. tuberculosis* cell envelope components, serving as anchors of the outer substance, thereby preventing *M. tuberculosis* recognition and phagocytosis by host cells (Rousseau *et al*., 2003, Hatzios *et al*., 2009).

1.5.4.2 The Mycolic acid ester of trehalose glycolipids family

The mycolate-containing trehalose lipids, TDM (6,6’-dimycoloyl-α-D-trehalose) and TMM (6-monomycoloyl-α-D-trehalose), are identified by their extended acyl chains (Hunter *et al*., 2006). These lipids are oriented with their mycolate groups facing the plasma membrane and their trehalose to the extracellular environment. Although they are not covalently attached to the *M. tuberculosis* cell envelope, these groups are extensively distributed, giving *M. tuberculosis* a waxy appearance (Garcia-Vilanova *et al*., 2019). Originally extracted from various *M. tuberculosis* strains using petroleum ether extraction (Bloch, 1950), TDM mycolate consists of pairs of acyl-chains—a long β-hydroxy chain containing 50 to 60 carbons and a shorter α-alkyl chain with 24 carbons. Furthermore, TDM forms an

asymmetric bilayer on the *M. tuberculosis* cell envelope by interacting with ManLAM, a component present on the surface of *M. tuberculosis*, which helps maintain the envelope's integrity (Hunter *et al*., 2006).

The significance of TDM has been observed in both in vivo and in vitro studies assessing its impact on the relationship between *M. tuberculosis* and the host during infection (Silva *et al*., 1985, Indrigo *et al*., 2002). These studies reveal that *M. tuberculosis* produces a substantial amount of TDM during infection, especially within the phagosome (Fischer *et al*., 2001). TDM plays a crucial role in host recognition of *M. tuberculosis* by interacting with Mincle (macrophage-inducible C-type lectin) on the macrophage surface. This interaction leads to the interception of mycobacterial phagosome maturation following the phagocytosis of virulent *M. tuberculosis* strains. This interception allows the bacteria to survive within the phagocyte by engaging with Mincle alongside the Fcγ-receptor transmembrane segment, inducing Th1 (e.g., TNF, IL-12), Th2 (IL-6), and inflammasome (IL-1β) pro-inflammatory mediators (Welsh *et al*., 2008, Ishikawa *et al*., 2009).

Additionally, TDM indirectly contributes to the influx of neutrophils to the infection site through its recognition via Mincle, triggers host cell apoptosis, mobilizes and vascularizes lung granulomas, stimulates host cachexia, optimally restrains antibodies, and induces non-specific host responses against bacterial and parasitic infections (Yarkoni and Bekierkunst, 1976, Parant *et al*., 1977, Bekierkunst *et al*., 1971, Bekierkunst, 1968). Conversely, TMM is recognized as the primary precursor for TDM biosynthesis and is covalently linked to the arabinogalactan-peptidoglycan (mAGP) core of the *M. tuberculosis* cell envelope (Brennan, 2003). TMM contributes to maintaining the structure and function of the *M. tuberculosis* cell envelope by interacting with TDM, arabinogalactan mycolate, or phosphatidyl-ethanolamine. The latter plays a role in translocating TMM from the inner membrane to the core of the *M. tuberculosis* cell envelope (Su *et al*., 2019).

1.5.4.3 The LOS glycolipid family

LOSs, which are trehalose-based lipids situated on the *M. tuberculosis* cell envelope, are still being investigated within the *M. tuberculosis* complex (Daffe *et al*., 1991). Insight into their biological properties has been gleaned from studies on various Mycobacterial species such as *Mycobacterium kansasii*, *Mycobacterium smegmatis*, and *Mycobacterium marinum* (Daffe *et al*., 1991, Garcia-Vilanova *et al*., 2019). Most identified LOSs consist of at least one trehalose unit connected to one or two extended poly-methylated divaricated fatty acid chains, both saturated and unsaturated, attached to the glucose residues of the trehalose polymer (Etienne *et al*., 2009). Moreover, LOSs exhibit diversity in their oligosaccharide composition, with the tetraglucose structure [D-Glcp(β1→3)-D-Glcp-(β1→4)- D-Glcp-(α1→1α)-D-Glcp] being the most prevalent (Garcia-Vilanova *et al*., 2019). Regarding their biological role, LOSs believed to play a significant part in organizing bacterial transmission, movement, and biofilm formation (Garcia-Vilanova *et al*., 2019). They possess immunogenic properties, act as phage receptors, and might be involved in establishing infections (Ren *et al*., 2007). Additionally, LOSs significantly reduces the hydrophobicity of the mycobacterial cell envelope (Minnikin *et al*., 2015a).

1.5.4.4 The sulfated trehalose glycolipid family

Sulfated trehalose or sulfolipids (SL), particularly SL-1, are present in infectious strains of *M. tuberculosis* (Garcia-Vilanova *et al*., 2019). These derivatives are uncommon in nature, and it's believed that varying levels of sulfation can immediately impact binding characteristics, altering how *M. tuberculosis* interacts with host cell biomolecules (Mougous *et al*., 2004). SLs consist of a trehalose-2-sulfate disaccharide that can undergo acylation at different sites, ranging from two sites as in the diacylated sulfoglycolipid or Ac2SGL (Kumar *et al*., 2007), up to four sites as seen in SL-1 (Converse *et al*., 2003). They possess saturated and unsaturated extremely divergent fatty acids extending up to C64 (Leigh and Bertozzi, 2008, Goren and Mor, 1990). Furthermore, SL-1 appears to play a role in establishing infection by intercepting phagosome maturation in macrophages (Goren *et al*., 1976). It inhibits

macrophage activation, phagocytosis, and the formation of the inflammasome. SL-1 also triggers the release of superoxide in activated neutrophils and monocytes and exhibits immunogenic properties, making it potentially useful for serodiagnosis (Garcia-Vilanova *et al*., 2019). Ac2SGL has been observed to activate CD1b-restricted T-cells, inducing their replication, and promotes the production of Th1 cytokines via CD8+ T cells (Gilleron *et al*., 2004).

1.5.4.5 Other trehalose glycolipids family

Finally, phthienoates (or mycolipenates) and hydroxyphthienoates (or mycolipanolates) represent trehalose-containing lipids that have received less attention in studies. While their precise role in *M. tuberculosis* pathogenesis remains unclear, their contribution to safeguarding the integrity of the cell envelope is well-defined (Minnikin *et al*., 1985).

1.5.4.6 The PGL family

The phenolic glycolipids (PGLs) are *M. tuberculosis*-specific glycolipids featuring oligosaccharides and a glycosylated structure of phthiocerol dimycocerosates (PDIM) (Arbues *et al*., 2014). While they have been extensively examined in *Mycobacterium leprae* (triglycosylated PGLI) and *Mycobacterium marinum* (monoglycosylated PGL), they are absent in most *M. tuberculosis* complex strains, except for the plentiful presence in the *M. tuberculosis* complex strain canettii (PGL-Tb) (Garcia-Vilanova *et al*., 2019, Daffé *et al*., 1987). Research conducted in a mouse model using necrotic *M. tuberculosis* has revealed that the absence of *M. tuberculosis* PGLs diminishes Th17 regulation. Production of IL-17A, crucial for controlling the development of hypoxic necrotic granulomas and reducing TB severity, is impacted when *M. tuberculosis* PGLs are lacking (Domingo-Gonzalez *et al*., 2017). Furthermore, studies have demonstrated that the sugar component of PGL can bind to the lectin domain of CR3, enhancing the infection of human macrophages by *M. tuberculosis*. This action enables *M. tuberculosis* to block Toll-like receptor 2-induced NF-κB stimulation, consequently inhibiting the pro-inflammatory response and undermining the host immune defences (Arbués *et al*., 2016).

1.5.4.7 The mannose-containing glycolipid family

This group comprises phosphatidyl-myo-inositol (PI) glycosylated derivatives known as PIMs, along with the dominant mannosyl-β-1-phpsphmycoketides (MPM). PIMs play a pivotal role in the *M. tuberculosis* cell envelope, significantly contributing to metabolic processes and playing a primary role in interactions with the host. They rank among the most abundant phospholipids in the *M. tuberculosis* cell envelope, alongside phosphatidyl-ethanolamine. PI, an acidic phospholipid, consists of phosphatidic acid linked to myo-D-inositol via a phosphate group (Brennan, 2003). Variations among PIMs arise from their differing numbers of mannose and fatty acid components. For instance, the foundational structure of PIM_1 involves a mannose linked to position C-2 of the myo-inositol of the PI anchor. By introducing an α-D-mannosyl unit or attaching di- or trimannosyl units at position C-6 of myo-inositol, PIM_2, PIM_3, and PIM_4 are formed respectively. Additionally, PIM_4 can transform into PIM_5 by substituting the terminal mannose residue at position C-2 via an α-D-mannosyl, which can further lead to PIM_6 by an additional exchange at the same position (Torrelles and Schlesinger, 2010). $AcPIM_2$ and Ac_1PIM_2 (di- and triacylated PIM_2) and $AcPIM_6$ and Ac_1PIM_6 (di- and triacylated PIM_6) are among the most prevalent PIMs present in the *M. tuberculosis* complex (Khoo *et al*., 1995). These compounds are exposed on the cell surface and actively participate in initiating infection, aiding in macrophage recognition, and facilitating the non-opsonic domain of C3 during phagocytosis processes (Villeneuve *et al*., 2005). Furthermore, studies reveal that various types of PIMs play a role in host cell recognition, phagocytosis, oxidative stress response, cytokine production, intracellular vesicular trafficking, modulation of signalling pathways, apoptosis, and antigen presentation (Gilleron *et al*., 2003, Garcia-Vilanova *et al*., 2019).

1.6 *M. tuberculosis* adhesins

1.6.1 Cell wall lipids and host interaction

M. tuberculosis contains various components like capsular glycopolymer, lipids, and lipoglycan—specifically α-D-glucan, TDM, PIMs, LM, and ManLAM. The interactions between these molecules play a crucial role in modulating the host immune response during infection (Vinod *et al*., 2020). For instance, TDM interacts with several receptors in host cells, such as C-type lectin receptor mincle, scavenger receptor A (SRA), and MARCO (Richardson and Williams, 2014, Bowdish *et al*., 2009). LM associates with TLR2, whereas α-D-glucan, higher PIMs, and Man-LAM link to DC-SIGN (Tailleux *et al*., 2003, Maeda *et al*., 2003, Driessen *et al*., 2009, Dao *et al*., 2004). Moreover, Man-LAM and PIMs associate with the macrophage mannose receptor (MR) and the T-cell surface glycoprotein CD1b (Killick *et al*., 2013, Ernst *et al*., 1998). LM and Man-LAM also interact with CD36 and the scavenger receptor, reducing TNF-α production. In experiments conducted under in vitro conditions, the application of CD36 antibodies was shown to reverse this effect (Józefowski *et al*., 2011).

1.6.2 Cell wall proteins and host interaction

According to studies conducted by Mawuenyega *et al*. and Wolfe *et al*., the *M. tuberculosis* cell wall contains approximately 528 proteins. These proteins serve various functions: 35.23% are involved in small molecule metabolism, 24.62% play roles in macromolecule synthesis and degradation, 7.58% are associated with cell operations, 19.13% are uncharacterized hypothetical proteins, and the remainder remains unidentified (Wolfe *et al*., 2010, Mawuenyega *et al*., 2005). To better comprehend Mycobacterial cell wall adhesion proteins, they have been categorized into three main groups: i) surface proteins containing signal peptide sequences, ii) surface proteins lacking signal peptide sequences, and iii) transmembrane proteins (Vinod *et al*., 2020).

1.6.2.1 Surface proteins with the signal peptide

Out of the 528 proteins in the cell wall, 87 carry secretion signal sequences, yet not all possess adhesive properties. These proteins encompass various types like pili, the antigen 85 complex protein, alanine-and proline-rich antigenic glycoproteins, 19kDa lipoprotein, and Proline-Glutamic acid-containing protein (Vinod *et al*., 2020). Pili consists mainly of a recurring subunit known as pilin, featuring a hydrophobic proteinaceous structure protruding from the bacterium surface (Telford *et al*., 2006). In *M. tuberculosis*, there are two kinds of pili: curli pili (*MTB*; RV3312a) and type IV pili. Curli pili, around 2-3 nm in diameter, can attach to the host's extracellular matrix protein, laminin (Alteri *et al*., 2007). Research indicates that wild-type mtp strains exhibit increased binding to THP-1 macrophages compared to mtp mutants deficient in curli pili (Naidoo *et al*., 2014). The antigen 85 complex proteins contribute to TDM synthesis in Mycobacteria and can bind to extracellular matrix constituents such as fibronectin (FN), elastin, and tropoelastin (Vinod *et al*., 2020). The main antigen 85 complex proteins, Ag85A or FbpA, Ag85B or FbpB, and Ag85C or FbpC, have been identified (Wilson *et al*., 2004, Kremer *et al*., 2002).

Studies have indicated specific binding sites of Ag85 to motifs in the heparin-binding domain of FN and tropoelastin (Vinod *et al*., 2020). The Alanine- and proline-rich antigenic glycoprotein (Apa) (Rv1860; 45−47 kDa antigenic complex) is a mannosylated 325 AA protein rich in alanine and proline residues. O-mannosylation occurs on specific threonine residues like Thr49, Thr57, Thr76, and Thr316 (Dobos *et al*., 1996). These sites are crucial for Apa's interaction with host cell receptors, including SP-A (pulmonary surfactant protein-A) and C-type lectin receptors like DC-SIGN and macrophage mannose receptors (Ragas *et al*., 2007). Another surface adhesion molecule in *M. tuberculosis* is the binding transporter protein PstS-1 (Rv0934, 38-kDa), post-translationally glycosylated with α-D-mannose, aiding its attachment to the macrophage mannose receptor and promoting phagocytosis (Esparza *et al*., 2015). PstS-1 also stimulates the pro-inflammatory response through interactions with TLR2 and TLR4, activating ERK 1/2 and p38 MAP kinase pathways (Jung *et al*., 2006).

1.6.2.2 Surface proteins without signal peptides

Numerous cytoplasmic bacterial proteins have been detected on the bacterial surface, despite lacking a signal peptide (Henderson and Martin, 2011). This external localization of cytoplasmic proteins extends across bacteria, fungi, and parasites (Vinod *et al*., 2020). Among these cytoplasmic proteins found in bacterial cell walls are glyceraldehyde-3-phosphate dehydrogenase (GAPDH), α-enolase, phosphoglycerate mutase (PGM), triose-phosphate isomerase (TPI), and phosphoglycerate kinase (PGK) (Henderson and Martin, 2011). In *M. tuberculosis*, cytoplasmic proteins like malate synthase, chaperones, and GAPDH have been identified on the surface (Govender *et al*., 2014). While these proteins have distinct functions within the cytoplasm, their roles on the surface exhibit varied interactions with host components. Malate synthase (MS; Rv1837c), an enzyme in the glyoxylate pathway, localizes on the *M. tuberculosis* surface. Its role involves catalysing the condensation of acetyl-CoA with glyoxylate and water to form L-malate along with coenzyme A. Additionally, the C-terminal of malate synthase has been identified as an adhesion site for laminin and FN found on host cells (Kinhikar *et al*., 2006).

Protein chaperones, including heat shock proteins (HSPs), are also present in the cell wall. These ATP-dependent proteins facilitate intracellular folding of native proteins (Fink, 1999). In *M. tuberculosis*, two types of chaperones have been identified: Cpn60.1 (Rv3417c) and Cpn60.2 (Rv0440) (Kong *et al*., 1993). Cpn60.1 interacts with TLR4 and potentially engages with the CD14 receptor, while Cpn60.2 binds to TLR2 and TLR4 but demonstrates weaker affinity for CD43 receptors on macrophages (Lewthwaite *et al*., 2001). GAPDH, a cytoplasmic housekeeping gene that converts glyceraldehyde-3-phosphate to 1,3-bis-phosphoglycerate (Dumke *et al*., 2011, Bergmann *et al*., 2004), serves as a ligand in *M. bovis* BCG, attaching to the host DC-SIGN (Vinod *et al*., 2020). In both *M. tuberculosis* and *M. avium*, surface bound GAPDH interacts with human epidermal growth factor, thereby stimulating the growth of these pathogens (Carroll *et al*., 2010). Early Secreted Antigen-6 (ESTA-6; Rv3875) is an *M. tuberculosis* surface protein functioning as an adhesion protein and inducing host cell lysis. Studies suggest that during *M. tuberculosis* doubling within pneumocytes type I and II, ESTA-

6 is overexpressed to anchor to ECM motif proteins, inducing cell lysis and subsequent release and spread of the pathogen (Kinhikar *et al*., 2010).

1.6.2.3 Transmembrane proteins and host interaction

In a study investigating the protein profile of the *M. tuberculosis* cell membrane by Gu *et al*., they confirmed 79 out of a total of 739 proteins as membrane proteins. One such protein, Heparin-binding hemagglutinin (HBHA; Rv0475), plays a role in forming extracytosolic lipid inclusions in the cytoplasm and interacts with host proteins when on the bacterial surface. HBHA, a 198-amino acid, 28 kDa protein, comprises a leucine-rich transmembrane domain near the N-terminal (residues 5-18), a coiled-coil domain (residues 24-109), and a C-terminal cationic lysine-rich domain, C-LRD (residues 160-198) (Gu *et al*., 2003). C-LRD significantly aids in *M. tuberculosis* adherence to non-phagocytic host cells and binds to host heparan sulfate glycosaminoglycans like heparin in the ECM and on the host cell surface (Pethe *et al*., 2001). Another membrane protein found in the *M. tuberculosis* cell wall is protein kinase D (PknD), a transmembrane protein featuring an extracellular domain and an intracellular domain with kinase activity. PknD has been observed to interact with brain endothelium and laminin, facilitating the invasion of the pathogen across the blood-brain barrier and aiding in the onset of central nervous system infection (Be *et al*., 2008).

1.7 ABC-transporters

The superfamily of ATB binding cassette (ABC) transporter is an important integral membrane protein that is found all living organisms (Cassio Barreto de Oliveira and Balan, 2020, Braibant *et al*., 2000). ABC transporters are categorized into importers and exporters based on the direction in which their substrates are translocated (Saurin *et al*., 1999). Importers, exclusive to prokaryotes, are responsible for importing various types of substrates like amino acids, vitamins, sugars, oligopeptides, and metals (Cuthbertson *et al*., 2010, Cassio Barreto de Oliveira and Balan, 2020, Braibant *et al*., 2000). Due to their ability to handle various substances, ABC importers play crucial roles in several biological functions such as transportation, nourishment, virulence, and disease development. ABC transporters consist of a minimum of four domains, including two regions membrane spanning domain (MSDs)/ transmembrane domain (TMDs) usually consisting of four to eight transmembrane helices each, and two nucleotide binding domains (NBDs) (figure 1.6)(Cuthbertson *et al*., 2010, Braibant *et al*., 2000). These domains could exist independently, forming separate units necessary to create a tetramer for function. Alternatively, they could combine into a multi-domain chain that includes all four components or form either a homo- or hetero-dimer to operate (Locher, 2016, Braibant *et al*., 2000).

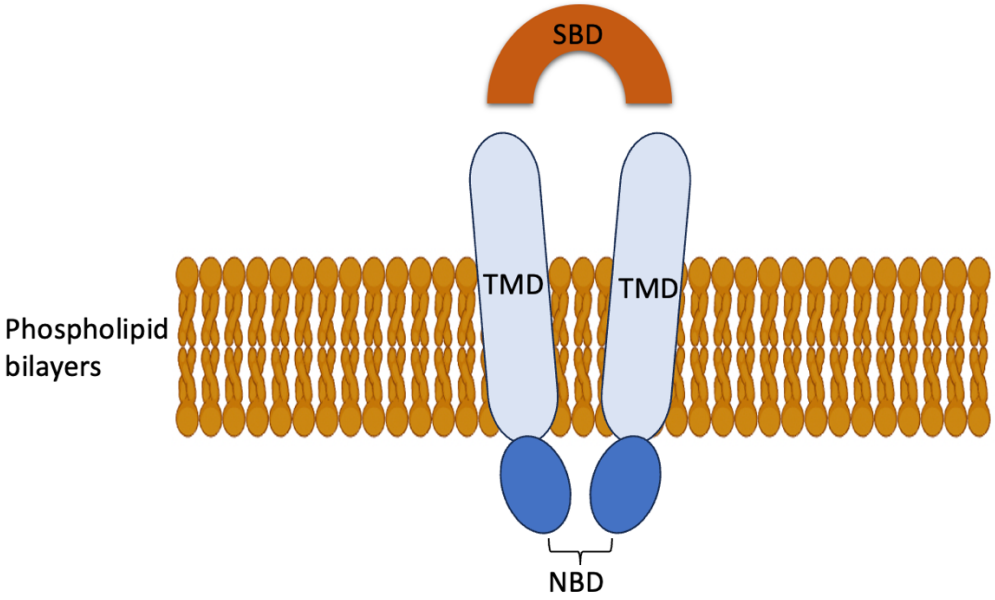


Figure 1.6: Schematic figure outlining the key domains of ABC transporters.

The NBDs contain protected stretches of amino acids, the ABC transporter signature, and a six-amino-acid motif. They hydrolyse and release ATP, providing the energy required for molecule transport, earning them the name "the engine of the transporter." (Braibant *et al*., 2000, Beis, 2015). NBDs consist of two sub-domains catalytic core domain and an α-helical domain (Beis, 2015). The catalytic core is the conserved domain, holds the necessary motifs for binding and breaking down ATP. These include the Walker A (also known as the P-loop), Walker B, along with the Q-loop, A-loop, and the H-motif (figure 1.7) (Beis, 2015, Ambudkar *et al*., 2006). The Walker A motif connects with the phosphate groups of the nucleotide. Meanwhile, the Walker B motif contains a glutamate residue that serves as a general base, triggering a water molecule to initiate a nucleophilic attack at the γ-phosphate of ATP (Beis, 2015). Additionally, the α-helices contain the ABC signature motif (C-loop), which includes an LSGGQ peptide sequence that plays a role in attaching the nucleotide (Locher, 2016, Braibant *et al*., 2000, Beis, 2015). To form the full ATP binding site, the two NBDs must come together and form a dimer, which occurs in a head-to-tail arrangement. This is crucial because both proteins contain motifs essential for creating a functional active site at each ATP binding site (Beis, 2015).

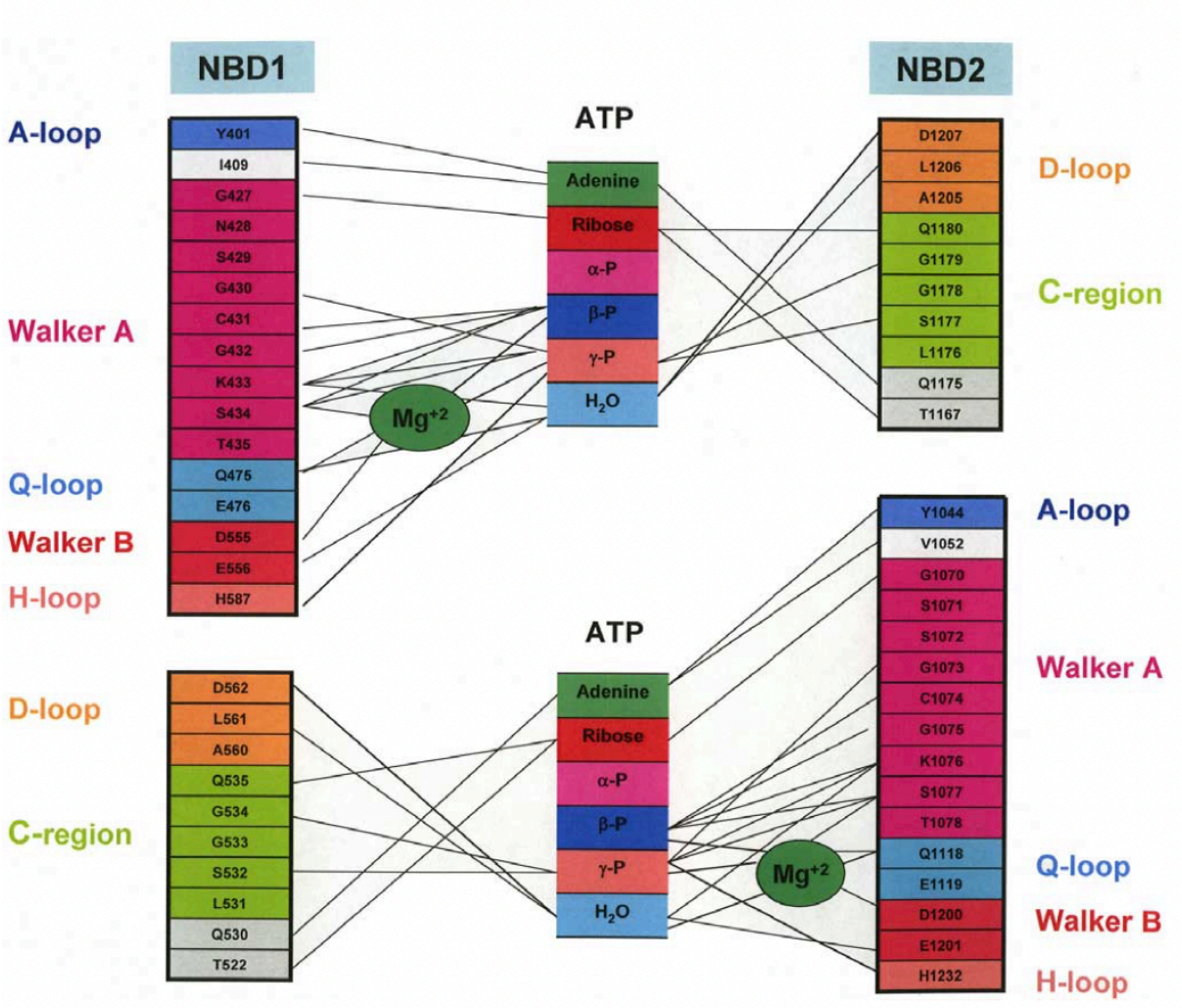


Figure 1.7 This describes how the nucleotide binding domain of a human ABC transporter interacts with ATP. It highlights the specific residues of the conserved ABC transporter motifs and their locations where they interact with ATP in the human protein. Figure produced with permission from Ambedkar *et al* 2006.

The A-loop, located near the Walker A motif, is a region commonly found in many ABC-transporters. It interacts with the adenosine part of ATP using tyrosine (Ambudkar *et al*., 2006). However, not all ABC-transporters contain this A-loop (Ambudkar *et al*., 2006). The Walker A and B motifs both interact with a bound Mg^{2+} ion, crucial for ATP binding through its connections with the α and b phosphates (Beis, 2015, Ambudkar *et al*., 2006). The Walker A motif shares a standard sequence in nucleotide binding domains, shown as (A/G)XXXXGK(S/T). Here, "X" represents any amino acid, and the letters in parentheses indicate multiple amino acids that can form this motif (Braibant *et al*., 2000). The Walker B motif comprises DEØØØØD, where "Ø" stands for a hydrophobic amino acid (Rees *et al*., 2009).

The Q-loop, a flexible section within the nucleotide binding domain, plays a crucial role in interacting with the membrane spanning domains. It spans about nine residues and typically starts with a conserved glutamine, followed by seven or eight other amino acids. This region flexibility allows it to interact with the active site, primarily through the conserved glutamine. These interactions are believed to regulate the protein conformational changes, enabling either the transportation or release of the substrate (Jones and George, 2002, Beis, 2015). The H-motif, situated after the Walker B motif, consists of a brief sequence of hydrophobic amino acids followed by a histidine. This section interacts with the gamma-phosphate of ATP and has been identified as crucial for the proper functioning of several ABC-transporters (Ambudkar *et al*., 2006).

The nucleotide binding domains have higher conservation compared to the membrane-spanning domains (MSDs). Some motifs within MSDs are specific to particular sub-families. Typically, an MSD comprises four to eight transmembrane helices usually six. When these domains are expressed as separate proteins, they are encoded within the same operon as the respective nucleotide binding domain they pair with (Braibant *et al*., 2000). The general arrangement of transmembrane domains displays considerable diversity, categorized into three types: importer type I, importer type II, and exporter (Rees *et al*., 2009). The connection between the energy from ATP hydrolysis and moving substances happens through a structurally similar but differently arranged alpha helix found in the loops within the cell, linking the membrane sections (Beis, 2015). Because there aren't many structures available for membrane proteins, including MSDs, it's hard to reach a clear agreement on their structural likenesses. Figuring out a structural consensus is made tougher by the wide range of transported substrates, from small molecules like lipids to larger polysaccharides. These variations could significantly impact how the complex is structured (Beis, 2015).

In addition to the two primary domains, the MSDs and NBDs, there is an extra domain responsible for importing. This domain is called the substrate binding domain (SBD) and operates with an additional protein. It is found in the periplasmic membrane, attached

either by an anchoring lipid tail or linked directly to the transporter in Gram-positive bacteria. These domains grab onto the transporter's cargo and deliver it to the ABC transporter for import (Braibant *et al*., 2000, Berntsson *et al*., 2010).

1.7.1 ABC transporter function

ABC transporters utilize energy from ATP hydrolysis to transport a wide range of substrates across cellular membranes, often moving them against concentration gradients (Rees et al., 2009). In prokaryotes, ABC importers facilitate the uptake of essential nutrients such as amino acids, sugars, and ions, while ABC exporters, present in both prokaryotes and eukaryotes, expel toxins, drugs, lipids, and other molecules from cells (Rees et al., 2009, Cassio Barreto de Oliveira and Balan, 2020). These transporters handle a diverse array of substances, including ions, peptides, sterols, and drugs (Thomas and Tampé, 2020, Rees et al., 2009, Cassio Barreto de Oliveira and Balan, 2020). Notably, some ABC transporters, like P-glycoprotein (ABCB1), contribute to multidrug resistance by pumping chemotherapeutic agents out of cancer cells (Borst and Elferink, 2002). ABC transporters also play key physiological roles, such as maintaining lipid and cholesterol homeostasis, facilitating antigen presentation in immune responses, supporting mitochondrial function and heme synthesis, and regulating the blood-brain barrier and nutrient transport (Robey et al., 2018, Lee et al., 2014, Borst and Elferink, 2002). In addition to transport functions, some ABC proteins are involved in processes like DNA repair, RNA translation, and chromosome maintenance (Borst and Elferink, 2002). In bacteria, ABC transporters are vital for cell viability, virulence, and pathogenicity, and they also play a crucial role in detoxifying cells by removing harmful substances (Rees et al., 2009, Borst and Elferink, 2002).

1.7.2 ABC transporter structure

ABC transporters can be classified into seven types based on the sequence and structure of their transmembrane domains (TMDs) (figure 1.8) (Davidson et al., 2008). These types are present across bacteria, which have the highest proportion of ABC systems (Davidson et al., 2008). Types I–III primarily function as importers (ter Beek et al., 2014, Rempel et al., 2019),

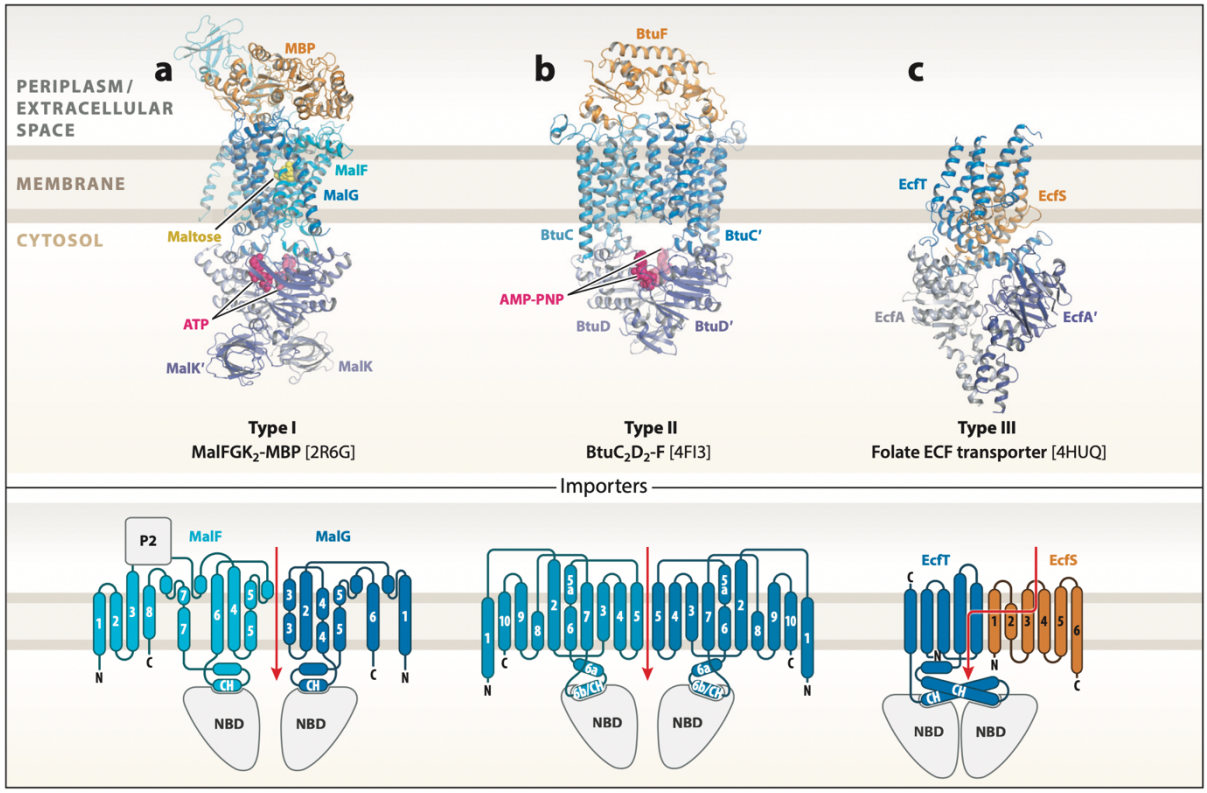
while types IV and V mainly act as exporters, though there are exceptions (Thomas and Tampé, 2020). Type VI systems work as extractors, and type VII transporters are involved in mechanotransduction and other functions like lipoprotein extraction or cell division. While type IV was originally considered an exporter, many type IV transporters have other roles (Thomas and Tampé, 2020). Type I importers (figure 1.8 a) are primarily responsible for transporting larger amounts of essential substances like sugars, amino acids, peptides, and osmoprotectants (ter Beek et al., 2014). In contrast, type II importers (figure 1.8 b) specialize in absorbing smaller nutrients, such as vitamin B12, metal chelates, and heme (ter Beek et al., 2014). The transmembrane domains (TMDs) of type I importers contain at least five helices (Oldham et al., 2007), while those of type II importers have ten (Locher et al., 2002, Korkhov et al., 2012). Both types rely on coupling helices to facilitate communication between their nucleotide-binding domains (NBDs) and transmembrane domains (TMDs), like other ABC transporters (Thomas and Tampé, 2020). Type I and II importers work together with substrate-binding proteins (SBPs), which bring the substrate to the transporter's transmembrane domains (TMDs) (Thomas and Tampé, 2020). SBPs are either attached directly to the transporter or anchored to the membrane by a helix or lipid in as in gram-positive bacteria or float freely in the periplasm as in gram-negative bacteria (Thomas and Tampé, 2020). Type III importers (figure 1.8 c), energy-coupling factor (ECF) transporters are mainly responsible for absorbing micronutrients, such as vitamins, nickel (Ni2+), cobalt (Co2+), and tryptophan like type II importer (Rempel et al., 2019). However, ECF transporters are different than the rest of ABC systems because their two transmembrane domains (TMDs) are structurally and functionally different from each other (Thomas and Tampé, 2020).

Type IV transporters act as multidrug efflux systems (figure 1.8 d), removing toxic compounds like antibiotics in bacteria (Wong et al., 2014). The structure of these transporters was initially identified through the crystal structure of the homodimeric Sav1866 from *Staphylococcus aureus* (Dawson and Locher, 2006) and later confirmed by examining other ABC exporters, including the lipopolysaccharide (LPS) transporter MsbA (Wong et al., 2014, Ward et al., 2007). Each half of a type IV ABC transporter has an elbow helix at its N terminus, a short structure that sits on the membrane surface facing the

nucleotide-binding domain (NBD) and likely helps stabilize the transporter (Thomas and Tampé, 2020). Following this is a core of six transmembrane helices (TM helices) that form the transmembrane domain (TMD). Two of the TM helices can swap between both TMD which give type IV transporters a special feature. Additionally, the NBDs are positioned farther from the membrane compared to other ABC transporters that related to a part of the TM helices extends into the cytosol (Thomas and Tampé, 2020). Type IV transporters have two coupling helices coupling helix 1 (CH1) and coupling helix 2(CH2). CH1 interacts with either the NBD on the same side of the transporter or both NBDs, depending on the stage of the transport cycle. On the other hand, CH2 only interacts with the NBD on the opposite side of the transporter (Thomas and Tampé, 2020). Type V transporters structure (figure 1.8 e) was first discovered in the human sterol transporter ABCG5/8 (Lin et al., 2015). Unlike type IV systems, where transmembrane domains (TMDs) swap between halves, each half of ABCG5/8 contains six tightly packed TMD helices that remain within their respective halves and do not cross over (Thomas and Tampé, 2020).

Type VI system (figure 1.8 f) is multicomponent membrane extractors, ABC transporter LptB2FG and four additional components, LptA, LptC, LptD, and LptE are essential to extract mature LPS from the periplasmic leaflet of the inner membrane and delivered to the outer membrane. The LptB2FG complex is an example of a new class of ABC transporters, featuring nucleotide-binding domains (NBDs) from the LptB subunits and transmembrane domains (TMDs) formed by LptF and LptG. The distinct structural configuration of LptF and LptG, there is no domain alternating between the subunits, sets the LptB2FG complex apart from other ABC transporters. Moreover, both LptF and LptG contain periplasmic domains that resemble β-jellyroll structures (Luo et al., 2017, Dong et al., 2017). In the absence of nucleotides, LptF and LptG create a large, outward-facing V-shaped hydrophobic cavity, which likely plays an important functional role (Owens et al., 2019, Li et al., 2019). Finally, type VII systems (figure 1.8 g) are involved in transperiplasmic mechanotransduction. A prime example is the MacAB-TolC complex, which acts as a tripartite efflux pump. Gram-negative bacteria use this system to expel toxic substances, including antibiotics and virulence factors (Thomas and Tampé, 2020, Du et al., 2018). The pump extends across the full width of the periplasm, connecting the inner and outer membranes. It is constructed

with the MacB homodimer, an ABC transporter embedded in the inner membrane, the hexameric MacA situated in the periplasm, and the TolC trimer positioned in the outer membrane. Together, they form a seamless channel to transport substances out of the cell (Thomas and Tampé, 2020).



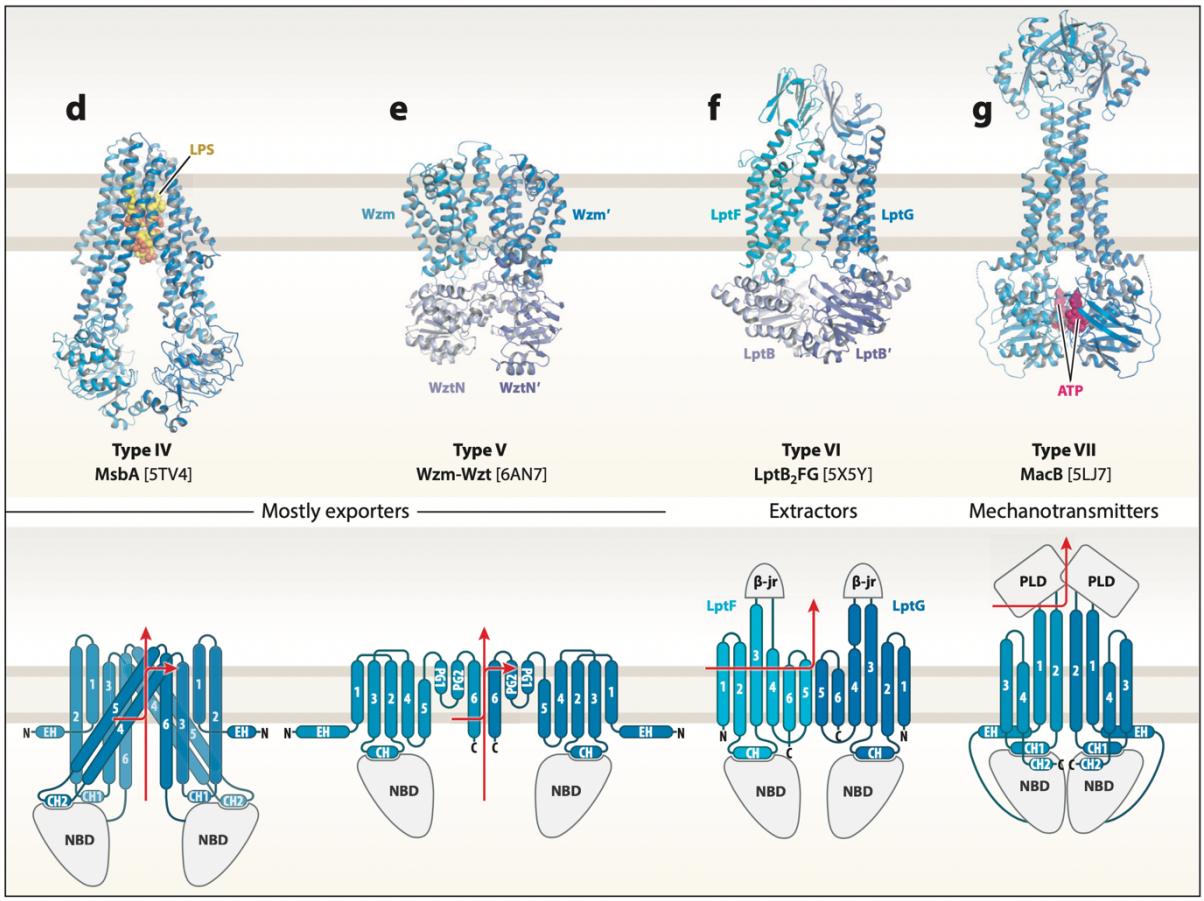


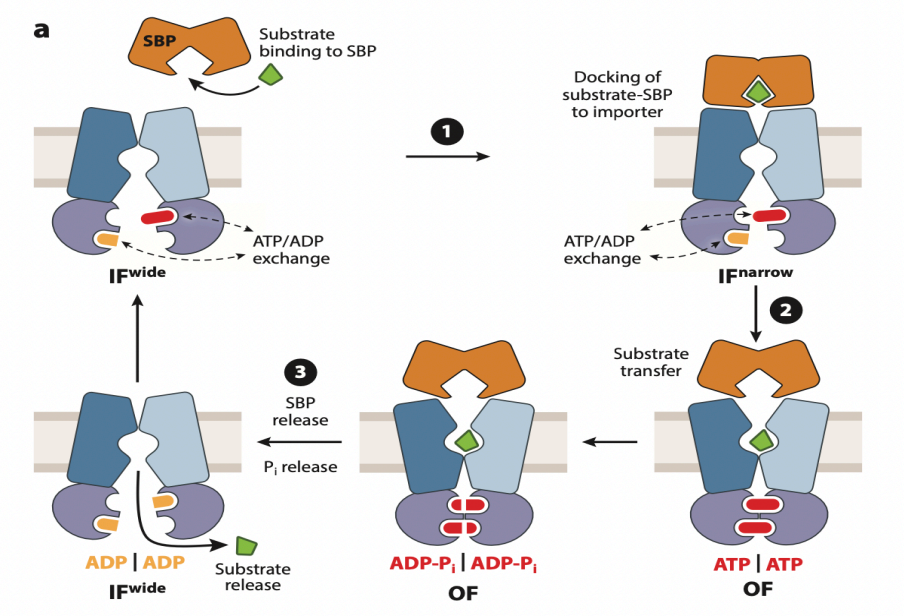
Figure 1.8: A summary of the various ABC transport systems, illustrated with representative experimental structures and matching TMD topology diagrams. Figures reproduce with permission from (Thomas and Tampé, 2020)

1.7.3 ABC transporter mechanism

ABC transporters undergo various conformational changes while transporting substances into or out of the cell. Importer ABC transporters bring substances into the cell, while exporter ABC transporters release them. For example, in the case of type I importers like the maltose transporter MalFGK2 (figure 1.9 a), the process begins when the substrate-binding protein (SBP) encounters a substrate and captures it by closing its two lobes, much like a Venus flytrap (Thomas and Tampé, 2020). The SBP, now holding the substrate, docks onto the transmembrane domains (TMDs) of the importer in its inward-facing (IF) position. This docking triggers structural changes in both the TMDs and nucleotide-binding domains (NBDs), drawing the NBDs closer together. When ATP binds, the NBDs fully dimerize, causing the TMDs to switch to an outward-facing (OF) position (Thomas and Tampé, 2020).

This change closes off the pathway from the inside of the cell and opens it to the outside, allowing the substrate to enter the translocation channel and bind to a specific site on the TMD. After ATP is hydrolysed, the release of phosphate (Pi) causes the transporter to return to its IF state, allowing the substrate to move into the cell (Thomas and Tampé, 2020).

Type II importers operate differently due to their high ATPase activity, even in the absence of substrate binding (figure 1.9 b). Additionally, the SBP undergoes only minor changes upon substrate binding. The transport pathway remains open to the outside when no SBP is attached. Once the loaded SBP docks onto a preformed site on the transmembrane domains (TMDs) of the ATP-bound importer, it releases the substrate into the pathway (Thomas and Tampé, 2020). Unlike type I importers, type II importers lack a specific binding site for the substrate along the pathway, which acts more like a smooth, non-stick Teflon surface. After ATP is hydrolysed and phosphate (Pi) is released, the cytosolic gate opens, pushing the substrate out (Thomas and Tampé, 2020). The TMDs then close, releasing the SBP and temporarily collapsing the importer to prevent leaks. Due to the high ATP concentration in the cell, this state quickly transitions back to the outward-facing (OF) position, releasing the SBP and preparing the transporter for the next cycle (Thomas and Tampé, 2020).



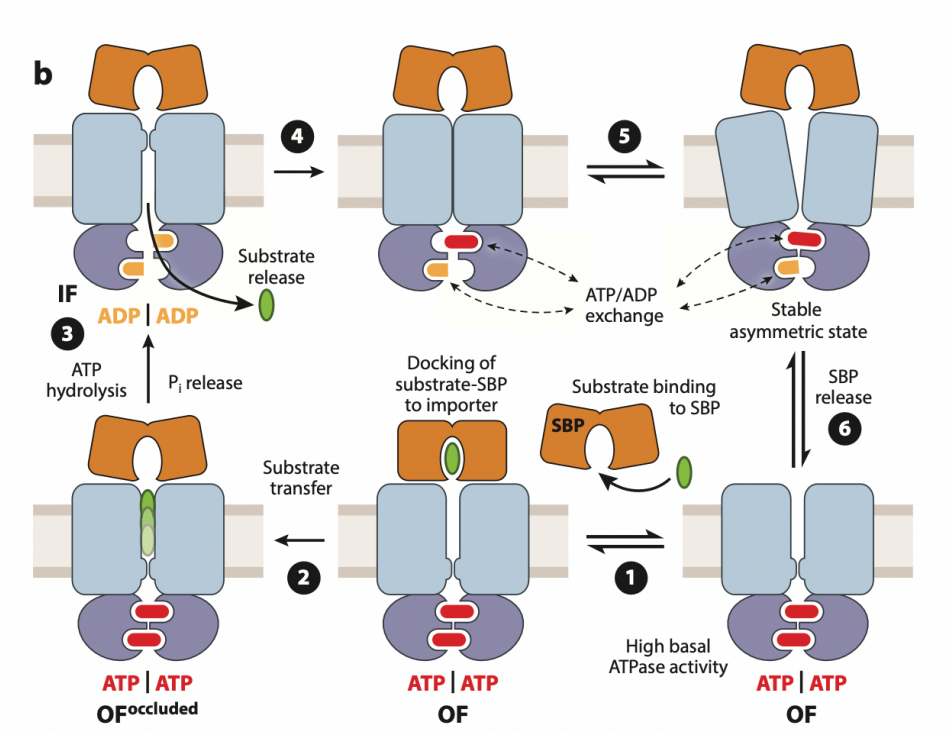


Figure 1.9: Models of the transport cycle of type I and type II ABC importers. (a) Substrate transport cycle of the maltose importer MalFGK2. (b) Substrate translocation cycle of type II importers, based on structures of the cobalamin importer BtuC2D2-F and the molybdate importer MolB2C2. Figures reproduce with permission from (Thomas and Tampé, 2020)

As an example of an ABC exporter transporter cycle (figure 1.10), TmrAB, a transporter in *Thermus thermophilus* similar to human TAP1/2, has been fully mapped for the first time, revealing its complete functional cycle and conformational changes in a lipid environment (Thomas and Tampé, 2020). Under normal conditions, the transporter exhibits two inward-facing conformations: IF^{narrow} and IF^{wide}. The IFwide state features a wider intracellular gate and larger cavity due to the movement of a segment called TM6, which acts as a "gatekeeper" for substrate binding. After ATP is hydrolysed, the transporter enters an "unlocked-return" state (UR^{asym} and UR^{asym*}), where the nucleotide-binding domains (NBDs) remain paired, but the intracellular gate begins to open, resetting the transporter to its inward-facing shape (Thomas and Tampé, 2020).

The outward-facing (OF) state, where the transporter is open to the outside, is brief and only seen when ATP hydrolysis is slowed. Locking the transporter in a pre-hydrolysis state

showed that the ATP-bound and hydrolysis states are very similar, with both open (OF^{open}) and closed ($OF^{occluded}$) extracellular gates, indicating that ATP binding alone can cause the shift from inward-facing to outward-facing (Thomas and Tampé, 2020). After ATP is broken down, the transporter slowly returns to its inward-facing state, triggered by the release of inorganic phosphate (Pi) (Thomas and Tampé, 2020). During this reset, the NBDs separate, and the intracellular gate opens due to movements of TM3 and TM4. This shows that the transporter's return to the inward-facing state is controlled by Pi release, not ATP hydrolysis (Thomas and Tampé, 2020).

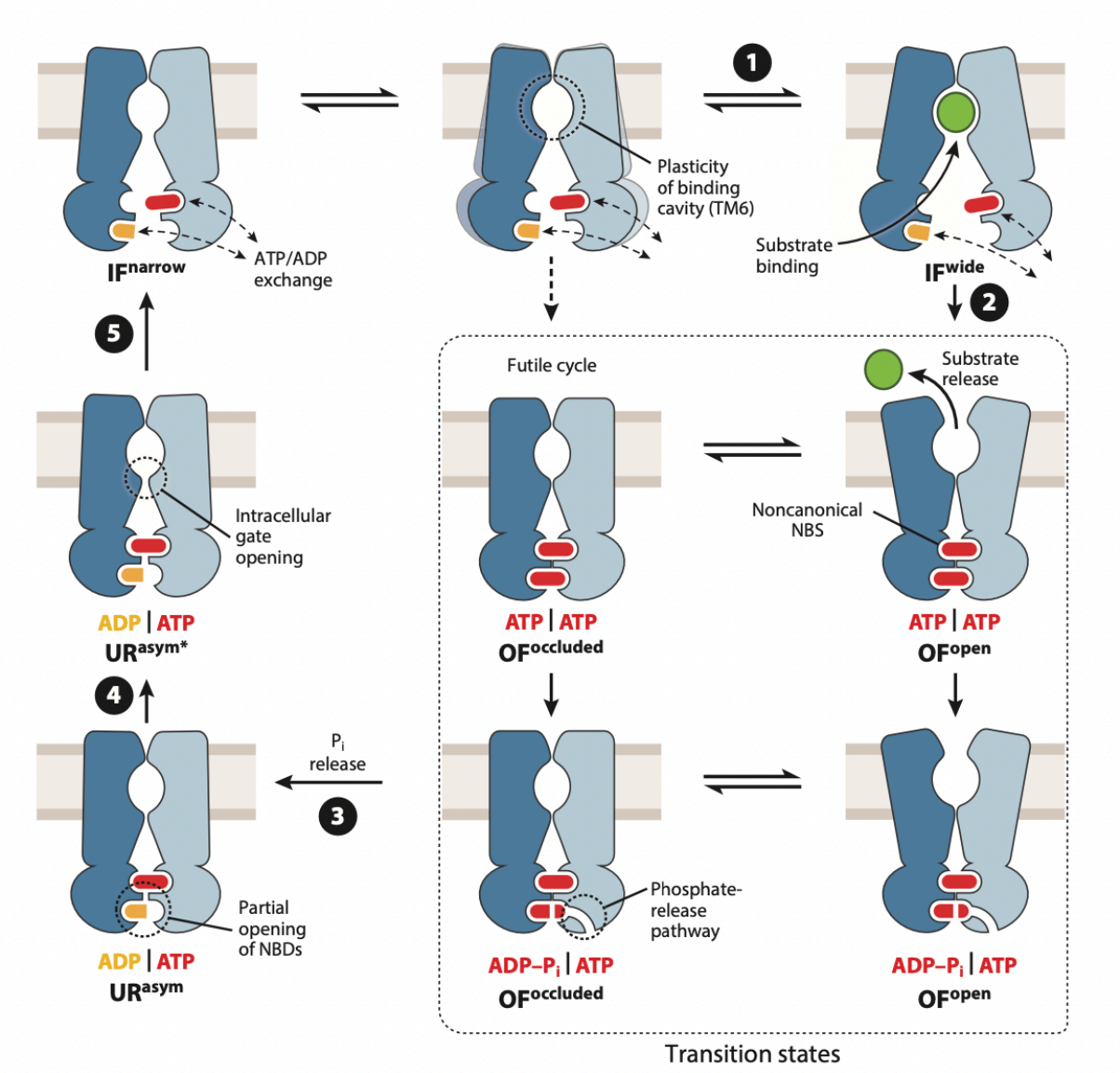


Figure 1.10: Translocation cycle of asymmetric type IV ABC transporters. Figures reproduce with permission from (Thomas and Tampé, 2020)

1.7.4 Mycobacterium ABC transporter

Most ABC importer systems identified in *M. tuberculosis* belong to group I, which includes transporters responsible for importing sugars, amino acids, peptides, and anions. However, four systems involved in iron import have been categorized under group II, while a specialized cobalt transporter has been classified into group III, known as the Energy Coupling Factor (ECF).In *M. tuberculosis*, these importers are accountable for the bacterium's adjustments within the human host (Cassio Barreto de Oliveira and Balan, 2020). *M. tuberculosis* ABC exporter systems are classified into groups IV, V, and VII (figure 1.11). Interestingly, experimental analysis has revealed that some components exhibit secondary or tertiary structural characteristics that do not align with their expected functional groups. On the other hand, exporters are present in both prokaryotes and eukaryotes, there are around 28 identified genes of ABC transporters exporters in *M. tuberculosis* (Cuthbertson *et al*., 2010, Cassio Barreto de Oliveira and Balan, 2020). Among those genes, approximately 26 possess 14 fully functional systems, divided into four distinct categories, recycling/transport of membrane components, electron transport chain (ETC), virulence and adaptation, and drug efflux (Cassio Barreto de Oliveira and Balan, 2020). or form either a homo- or hetero-dimer to operate (Locher, 2016, Braibant *et al*., 2000). Braibant *et al*. classified the superfamily of *M. tuberculosis* ABC transporters into twelve distinct subfamilies, it is predicted that *Mtb* expresses members from nine of these subfamilies (figure 1.12) (Braibant *et al*., 2000). Rv1273c and Rv1272c are membrane proteins in *Mycobacterium tuberculosis* that belong to the ATP-binding cassette (ABC) transporter family (Gupta et al., 2020a). These proteins function as exporter ABC transporters, playing a role in the efflux of macrolides and other drugs, alongside other transporters like DrrABC, Rv1686c/87c, Rv1456c/57c/58c, and Rv2686c/87c/88c (Cassio Barreto de Oliveira and Balan, 2020). Rv1273c, specifically, is an ATP-binding protein of the MsbA subfamily within the ABC transporter family. Rv3781c and Rv3783c are membrane proteins classified as ABC transporter exporters, possibly involved in the recycling and transport of membrane components and lipopolysaccharides (Cassio Barreto de Oliveira and Balan, 2020). Rv3781c, referred to as RfbE, contains nucleotide-binding domains (NBDs), while Rv3783c, known as RfbD, has membrane-spanning domains (MSDs) similar to those associated with O-antigen and lipopolysaccharides in various species (Braibant et al., 2000).

Rv1218c and Rv1217c are components of an ABC transporter system in *M. tuberculosis* that acts as an exporter, playing a role in drug efflux and contributing to antibiotic resistance (Cassio Barreto de Oliveira and Balan, 2020). These three complexes are believed to play a key role in actively transporting drugs across the membrane, contributing to multidrug resistance through an export mechanism.

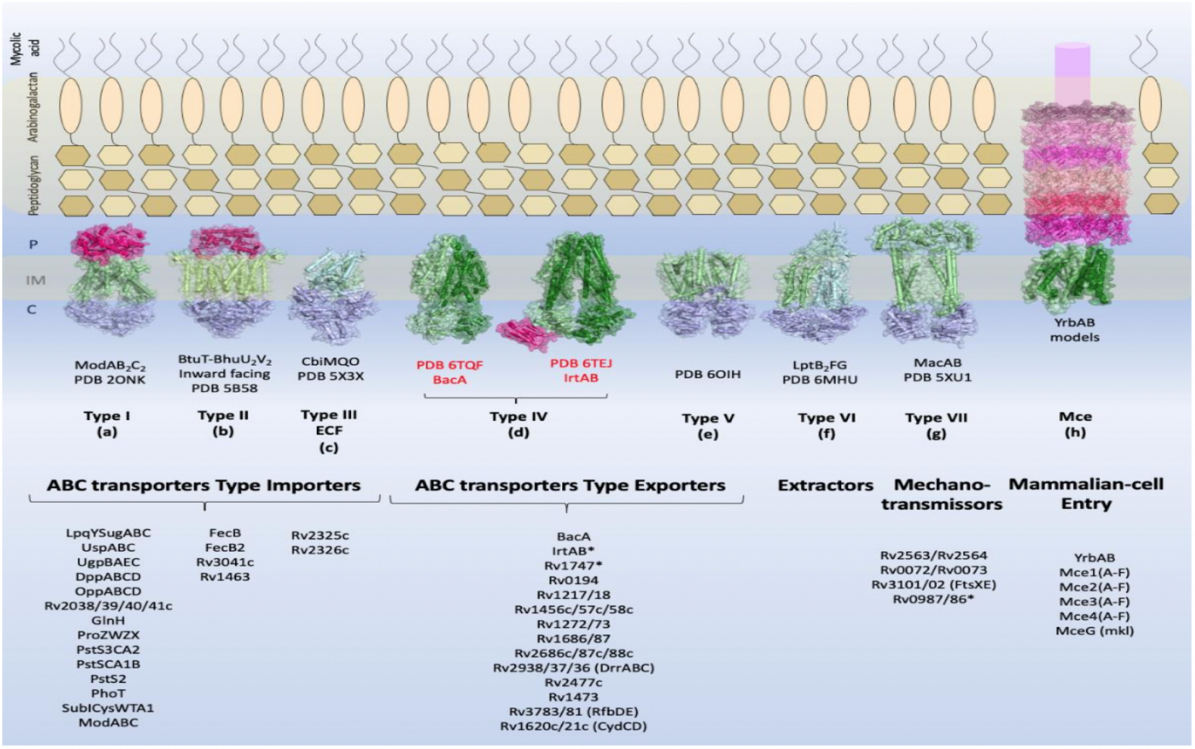


Figure 1.11: Overview of the types of ABC transporters and the components identified in *M. tuberculosis*.
Figure reproduce with permission from (Cassio Barreto de Oliveira and Balan, 2020)

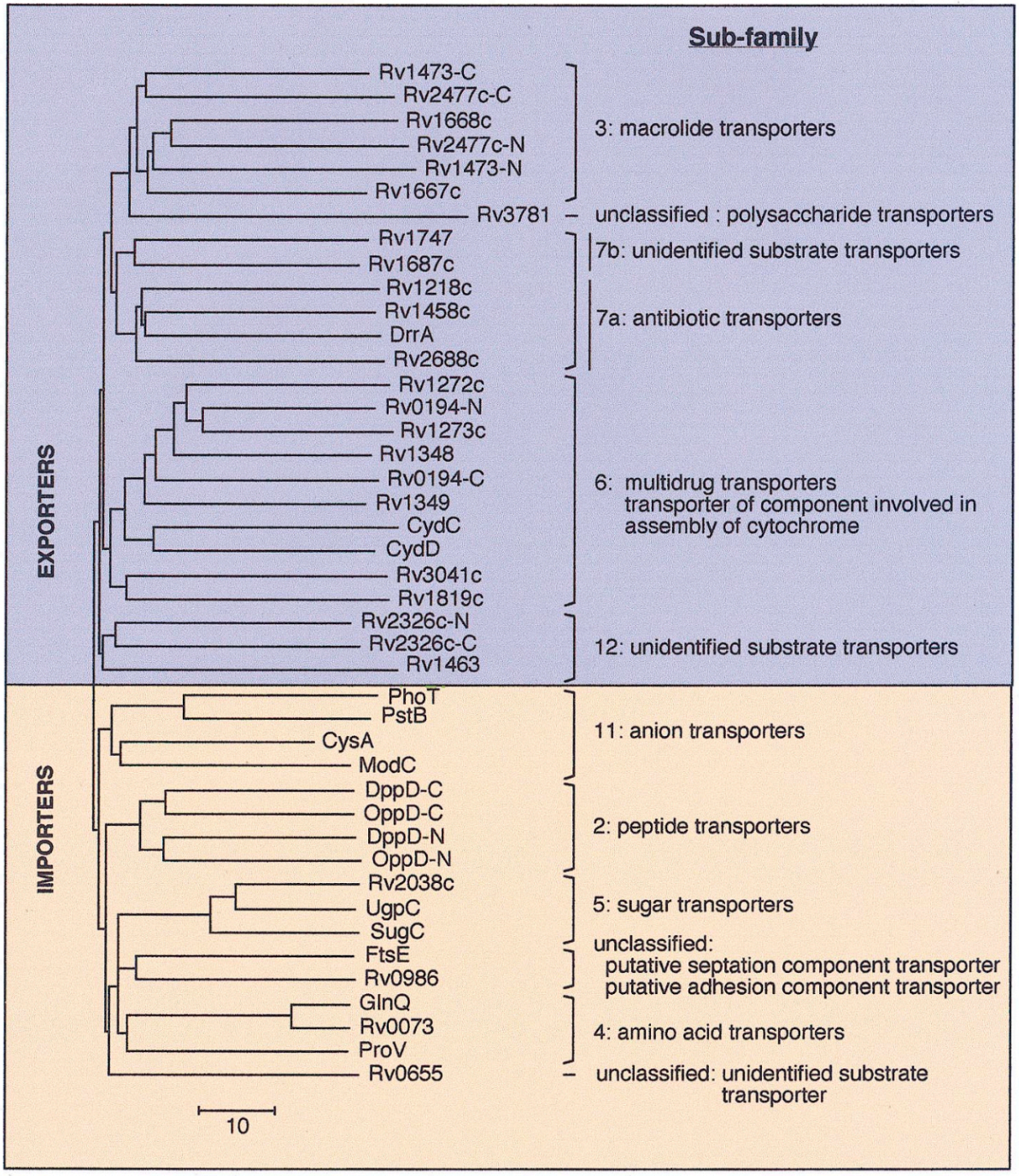


Figure 1.12: Classification of the ABC- transporter in *Mtb.* Highlighting the predicted importers and exporters along with the predicted sub-family. Figure produced with permission from Braibant *et al* 2000.

1.8 Styrene Maleic Acid

The application of poly (styrene-co-maleic acid) (SMA) has gained popularity as a biochemical method to extract a range of membrane proteins within their natural lipid environment without any detergent. SMA enables the extraction of these proteins while frequently preserving both their structural and functional integrity (Stroud *et al*., 2018). Styrene maleic acid (SMA) is a synthetic co-polymer that is amphipathic, made up of maleic acid (hydrophilic) and styrene (hydrophobic), It is known to interact with lipid bilayers, triggering the creation of disc-like assemblies known as styrene maleic acid lipid particles SMALPs (Stroud *et al*., 2018). These SMALPs are stable, self-assembling particles containing around 140 lipids, encased by SMA. When added to a lipid suspension at pH 6.5 or higher, SMA forms nano discs about 10 nm in diameter (Lee and Pollock, 2016).

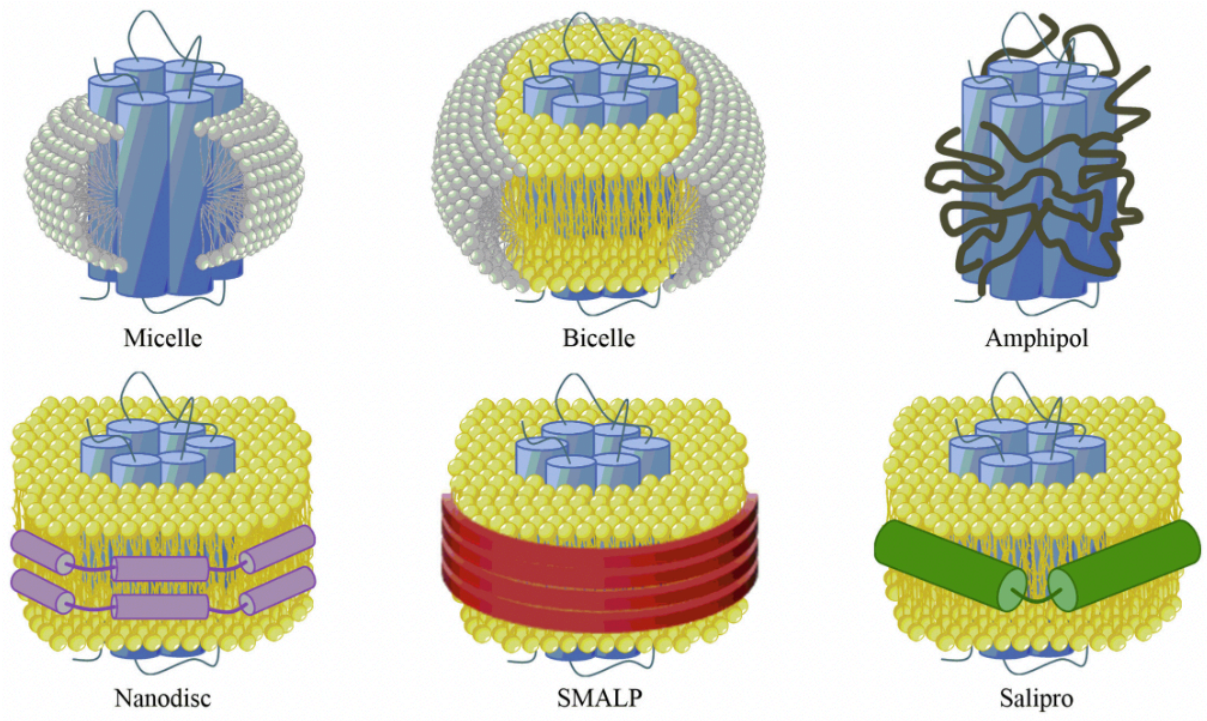


Figure 1.13: Schematic figure illustrates various amphiphilic systems used in membrane protein structural studies, including detergent micelles, bicelles, amphipols (Apols), nanodiscs (NDs), SMALPs, and Salipro. The membrane protein is shown as blue cylinders within these systems. Lipids and detergents are represented by yellow and gray, respectively, while amphipol is depicted in black. Nanodiscs feature membrane scaffold proteins (MSP) in purple, SMALPs use SMA polymers shown in red, and Salipro incorporates saposin A, represented by green cylinders figure reproduce with permission from (Lee et al., 2022)

Ever since the first discovery of SMA ability to create SMALPs (Styrene Maleic Acid Lipid Particles) and solubilize various biological membranes (Knowles *et al*., 2009), this method has been employed to directly extract diverse targets from different biological membranes. These targets include several ABC transporters (P-glycoprotein) and G protein-coupled receptors GPCRs (Adenosine A2a receptor)(Wheatley et al., 2016). Additionally, the homodimer ABC transporter Sav1866 easily solubilised in SMA and was isolated in substantial, highly pure quantities within Sav1866–SMALPs. These entities yielded favourable outcomes in circular dichroism and SAXS analysis. However, they, like other ABC transporters within SMALPs, did not exhibit ATPase activity. Nonetheless, they provided satisfying results regarding the secondary structure of Sav1866–SMALPs (Pollock et al., 2022). This technique disrupts the lipid bilayer, producing lipid patches that house membrane proteins (MPS). SMALPs are highly efficient for purifying various membrane proteins like G-protein coupled receptors (GPCRs), transporters, and ion channels (Stroud *et al*., 2018). Targets isolated through SMA extraction have been expressed in mammalian, insect, yeast, and bacterial systems, and this approach has been utilized to study membrane proteins and explore their potential as therapeutic targets (Gulati *et al*., 2014, Grime *et al*., 2021, Dilworth *et al*., 2018). The stability provided by SMALPs allows for further structural analysis using methods like electron microscopy (EM), analytical ultracentrifugation (AUC), and dynamic light scattering (DLS)(Lee *et al*., 2016).

They have distinct advantages over other purification methods as they do not need traditional detergents that can interfere with lipid-protein or protein-protein connections. Unlike other systems, SMALPs do not require additional protein components to maintain their structure (e.g., membrane scaffold proteins like MSP)(Stroud *et al*., 2018, Lee and Pollock, 2016, Lee *et al*., 2016). Plus, they are cost-effective since the starting material, styrene-maleic anhydride, is relatively inexpensive and can easily convert to styrene-maleic acid (Lee and Pollock, 2016). Additionally, SMALPs allow conducting experiments at room temperature without affecting the structural or functional characterisation of the solubilized membrane proteins. This is distinct from the behaviour observed in membrane proteins solubilized using traditional detergents (Pollock et al., 2022).

However, SMALPs have limitations, excessive polymer amounts, similar to detergents, can affect biochemical assays, studies suggest that the SMA polymer might interact with the carbon grids used in cryo-EM studies, causing reduced resolution, akin to the way high detergent concentrations complicate sample preparation (Postis *et al*., 2015). SMA is sensitive to pH, functioning optimally at pH 7.5; below 6.5, it can precipitate, causing SMALP disassembly and protein precipitation. Moreover, high concentrations of divalent cations decrease the solubility of the SMA co-polymer, separating it from the lipid and reducing particle stability (Lee *et al*., 2016).

1.9 Aims

The aim of this project is to create a robust protocol for solubilizing and purifying mycobacterial membrane proteins using SMALPs. Aiming to explore the potential of the SMALP method in extracting intact complexes of ABC transporters from *Mycobacterium tuberculosis*. Specifically, the project will evaluate the effectiveness of SMA in isolating and purifying three key mycobacterial ABC transporter complexes—Rv1273c/72c, Rv3781c/83c, and Rv1218c/17c.

In addition, the project will carry out an initial assessment into the use of SMALPs to selectively extract peripheral molecules from mycobacterial cell envelopes. The aim being to explore SMA's capacity to extract various labelled lipids ([14C]-uniformly labelled lipids), including complex glycolipids (such as trehalose monomycolate, trehalose dimycolate, and glucose monomycolate), phospholipids (specifically phosphatidyl inositol mannosides), lipoglycans (including lipomannan and lipoarabinomannan), and surface-exposed proteins from Mycobacterium smegmatis envelopes without causing cell rupture. The future aim being the use SMA to directly explore directly the lipidome and membrane proteome of TB.

CHAPTER 2

MATERIALS AND METHODS

2.1 Liquid growth media

Luria Bertani (LB) broth powder (Miller's), and terrific broth (TB) powder (Avantor) were made following the manufacture's protocol in deionised water and autoclaved for 20 minutes at 121C, 1.4 bar. Media was left at room temperature to cool down before addition of antibiotics and inoculation.

2.2 Solid growth media

LB agar was made by dissolving LB broth as mentioned above with the addition of 12 g/L of agar granules (Lennox). LB agar was autoclaved for 20 minutes at 121^{o}C, 1.4 bar, when media cooled to 50^{o}C appropriate antibiotics was added. LB agar plates were poured and left to set on the bench under sterilised conditions and stored at 4^{o}C.

2.3 Antibiotic and supplements

All antibiotics and supplements were prepared in deionised water. Solutions were stored in -20^{o}C after filtration via 0.2um syringe filter.

Antibiotic	Stock Concentration	Working Concentration
Streptomycin	50 mg/mL	50 μg/mL
IPTG	1 M	200 μM

Table 2.1. Antibiotic and supplements used in this study.

2.4 Preparation of Styrene Maleic Acid

Styrene maleic acid was prepared following the protocol described in Lee *at el.* (2016). In summary, 1 M of NaOH was added to styrene maleic anhydride co-polymer and autoclaved at 121 °C, 1.4 bar for 20 minutes, and left to cool. Concentrated HCl was added to precipitate the polymer, until the pH of the solution was <5. The material was then centrifuged at 11,000 x g at room temperature for 15 minutes to wash the precipitate and the supernatant was discarded. The pelleted polymer was washed in distilled water and again centrifuged at 11,000 x g at room temperature for 15 minutes three times. Then 0.6 M NaOH was added to the polymer and incubated in orbital shaker at 37 °C overnight to

dissolve. Once the polymer dissolved, the polymer was precipitated again using concentrated HCl, then washed in distilled water and centrifuged 3 times as stated above and dissolved in 0.6 M NaOH. HCl and NaOH were used to adjust the pH of the solution to 8.0, then the solution was frozen at -20°C. Where required the polymer was freeze dried to a fine powder and stored at room temperature.

2.5 Plasmid

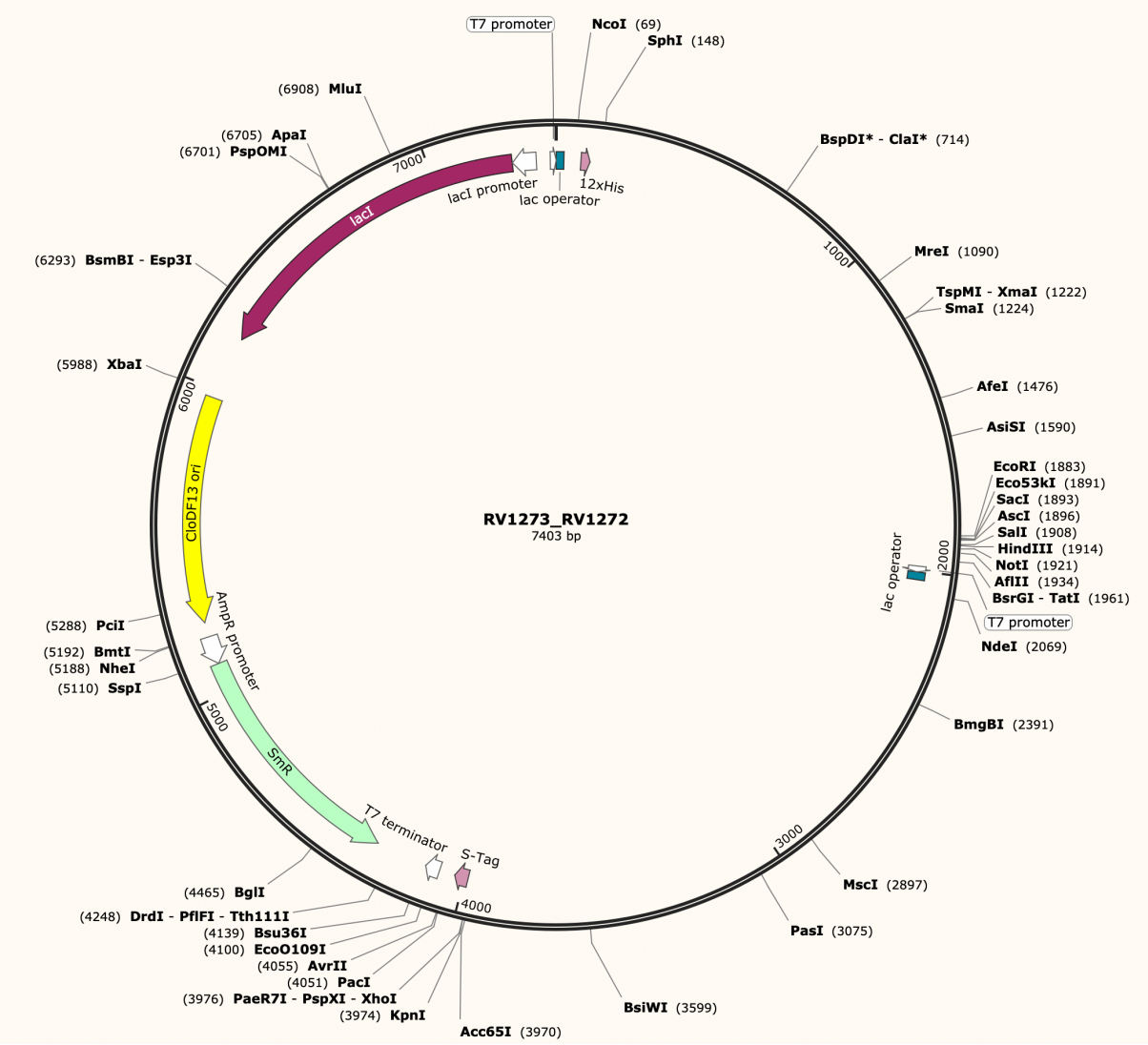


Figure 2.1: pCDF_Rv1273c-Rv1272c Plasmid.

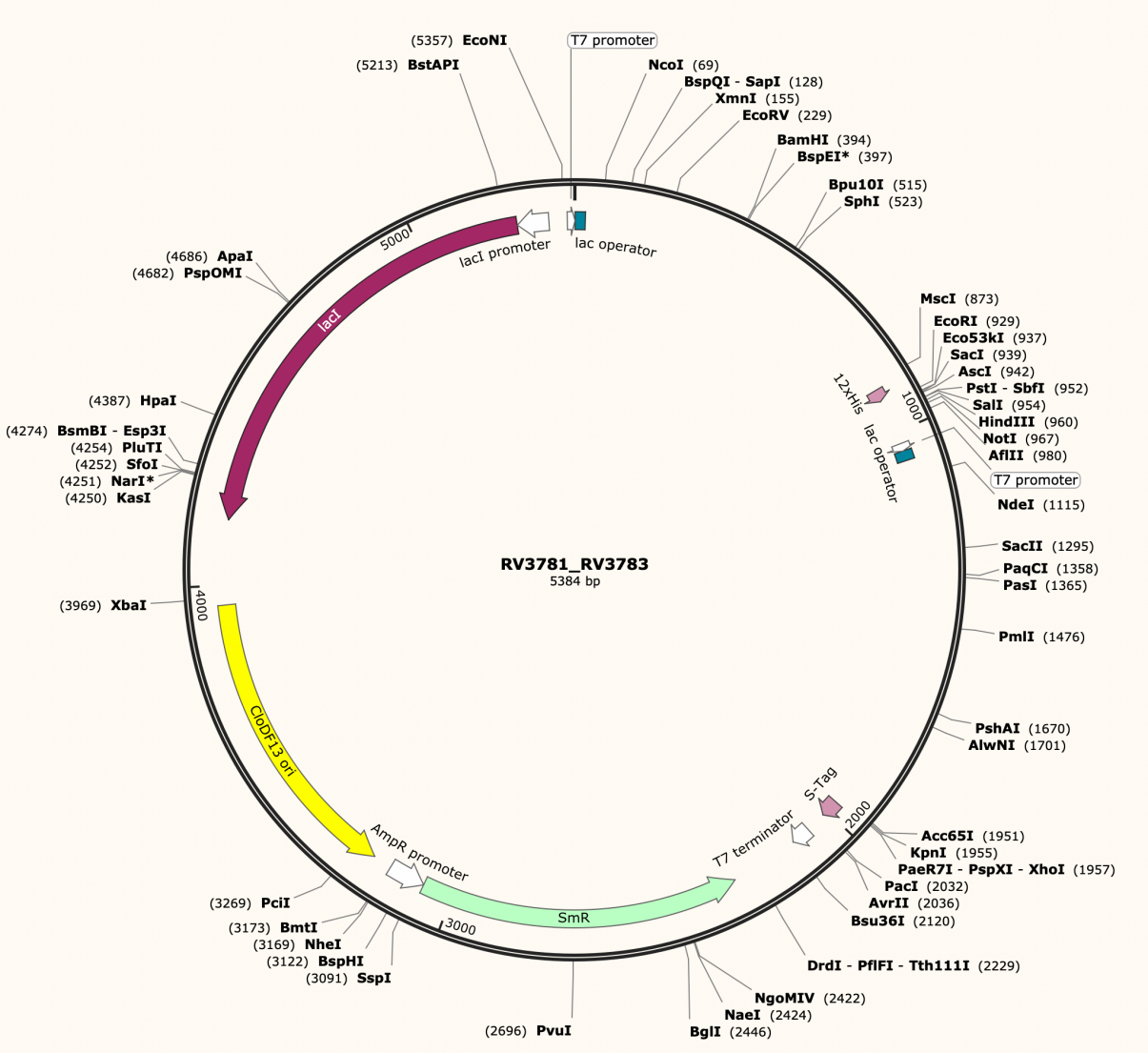


Figure 2.2: pCDF_Rv3781c-Rv3783c Plasmid

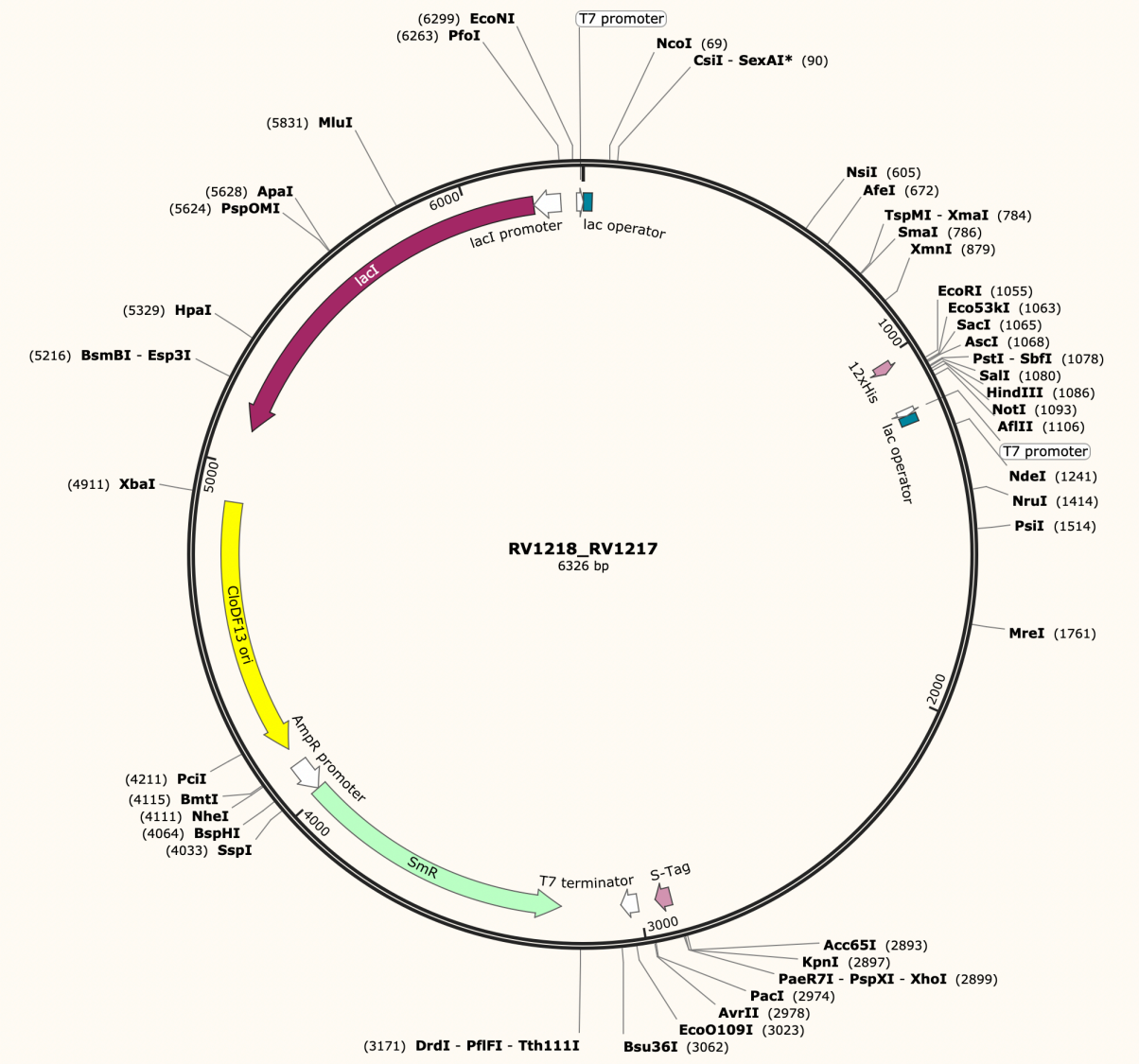


Figure 2.3: pCDF_Rv1218c-Rv1217c Plasmid

2.6 Growing BL21(DE3 huA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS λ DE3 = λ sBamHIo ΔEcoRI-B int: :(lacI: PlacUV5:T7 gene1) i21 Δnin5) competent cells

LB agar plate was streaked with *E. coli* BL21(DE3) (New England Biolabs) and incubated overnight at 37 °C. A single colony was used to inoculate in 5 mL of LB broth and incubated overnight at 37 °C. Around 500 µL of the overnight culture was added to a fresh 50 mL of LB and incubated at 37 °C with shaking at 180 rpm and monitored at OD_{600}. When OD_{600} of 0.4 was reached the culture was chilled on ice and cells were collected by centrifugation using sterile centrifuge tube at 4,000 rpm in an Eppendorf Multi-purpose centrifuge for 5 minutes

at 4 °C and the supernatant discarded. The cells were resuspended in 20 mL of ice-cold 50 mM $CaCl_2$ and incubated on ice for 20 minutes. Cells were collected by centrifugation using sterile centrifuge tube at 4,000 rpm in an Eppendorf Multi-purpose centrifuge for 5 minutes at 4 °C and the supernatants discarded. Cells were resuspended in 2.5 mL of ice-cold 50 mM $CaCl_2$ supplemented with 10 % glycerol for long term storage at -80 ^{o}C.

2.7 Transformation of BL21(DE3) competent cells

For high-yield protein expression, 100 µL of BL21(DE3) competent cells were used to transform the plasmids containing the desired protein gene. Vials of plasmids were thawed on ice. Subsequently, 1 µL of plasmid DNA was added to the competent cells and the mixture was incubated on ice for 30 minutes. The cells were then heat-shocked for 45 seconds at 42 °C using a heating block, followed by immediate incubation on ice for an additional 5 minutes. Afterward, 900 µL of fresh LB media was added to the cells, which were then allowed to recover at 37°C with shaking at 180 rpm for 2 hours. The transformed cells (100 µL) were then plated onto LB agar plates supplemented with the appropriate antibiotic. Sterile glass beads were used to evenly spread the cells on the plate, which was then incubated overnight at 37° C.

2.8 Preparation of Glycerol Stock

For long term storage of transformed cells, one colony from LB agar plate was used to inoculate 10 ml of LB broth supplemented with streptomycin 50 µg/mL in 50 ml falcon tube and incubated at 37°C overnight with shaking at 180 rpm. 500 µL of the culture was added to 500 µL of sterile 50% glycerol in a 2 mL cryovial and frozen at -80°C.

2.9 Large Scale Protein Expression

A single colony from a LB agar plate was used to inoculate 20 ml of LB media supplemented with streptomycin 50 µg/mL in 50 ml falcon tube and incubated at 37°C overnight with shaking at 180 rpm. The 20 ml of the overnight culture was later used to inoculate 2 L of terrific broth (TB) in 5 L conical flask and incubated at 37°C with shaking at 180 rpm. Cell growth was mentored via measuring the OD_{600} every hour. When cultures reach an OD_{600} between 0.6 – 0.8, cultures were induced with IPTG at a final concentration of 200 µM IPTG and incubated overnight at 18°C with shaking at 180 rpm. Cells were harvested by centrifugation at 4,000 x g for 20 minutes at 4°C. The pelleted cells were transferred to a 50 mL falcon tube and stored at -80°C.

2.10 Membrane Protein Solubilisation and Purification

Disruption buffer (Buffer A)	50 mM TrisHCl pH 7.5 (r.t) 5% (v/v) glycerol 2mM EDTA Before use, dissolve one PierceTM EDTA-free Protease Inhibitor Tablet (Thermo Scientific) per 50 mL buffer
Solubilisation buffer (Buffer B)	50 mM TrisHCl pH 8.0 (r.t) 10% (v/v) glycerol 500 mM NaCl
SMA solubilisation buffer	Solubilisation buffer + 5% (w/v) SMA
Ni-NTA wash buffer.	1- Solubilisation buffer + 10 mM imidazole 2- Solubilisation buffer + 20 mM imidazole 3- Solubilisation buffer + 50 mM imidazole 4- Solubilisation buffer **+** 100 mM imidazole
Ni-NTA elution buffer	Solubilisation buffer + 500 mM imidazole

Table 2.2. Buffers used in membrane protein solubilisation and purification.

2.10.1 Rv1273c/Rv1272c, Rv3781c/Rv378c, and Rv1218c/Rv1217c Membranes Preparation

All steps of membrane preparation were done on ice to maintain protein integrity. If frozen pellets were to be used for membrane preparation the cells were defrosted completely on ice. Cell pellets were resuspended in disruption buffer (buffer A) at least 2 mL per gram of pellet by hand homogenisation using a borosilicate glass homogeniser. Protein suspensions were passed through an Emulsiflex C3 Homogeniser (Avestin) not less than 5 times to disrupt the cells under pressure of around 17,000 psi. Cells were then centrifuged at 12000 x g for 30 minutes at 4°C to remove cell debris. Supernatants were transferred into pre-weighed Beckman Ultra tubes (26.3 mL, 25 X 89 mm) making sure they were 2/3 full and centrifuged at 120,000 x g for 1 hour in the XL 90 (Beckman) to sediment the membranes. The collected membrane fractions were weighed and resuspended in solubilisation buffer (buffer B) by hand homogenisation to a final concentration of 60 mg/mL. The resuspended membrane solution was stored in -80°C for long term and thawed on ice before use.

2.10.2 SMA Solubilisation

The resuspended membranes were added to SMA solubilisation buffer in 1:1 ratio and solubilised in the room temperature for 2 hours with light agitation. The resulting solution was transferred into pre-weighed Backman Ultra tubes (26.3 mL, 25 x 89 mm) making sure they were 2/3 full and centrifuged at 120,000 x g for 45 minutes in the XL 90 to remove any insolubilized materials and collect the supernatants.

2.10.3 Ni-NTA Affinity Purification

2.10.3.1 Ni-NTA Affinity Preparation

HisPur Ni-NTA Resin (Thermo) was prepared by adding the required volume of the resin (1 mL / 10 mL of membrane) to a 15 mL conical centrifuge tube and centrifuged for 2 minutes at 4,000 rpm in an Eppendorf Multi-purpose centrifuge at 4°C. The supernatants were removed, and the resin was washed twice with equal volume with deionised water. Then, the resin was resuspended as a 50% (v/v) suspension in solubilisation buffer (buffer B). This suspension was added to the supernatants from section 2.9.2 in 1:50 ratio (1 mL 50% resin suspension to 50 mL supernatant) and the resin/supernatants mixture were incubated overnight at 4°C in rotor.

2.10.3.2 Ni-NTA Affinity Purification

The Ni-NTA affinity purification involved three sets of protein pairs: Rv1273c and Rv1272c, Rv3781c and Rv3783c, and Rv1218c and Rv1217c. Initially, the resin/supernatants mixture obtained from section 2.9.3.1 was poured into a 20 mL gravity flow column, and the flow-through was collected. Following this, the resin underwent four washes using 5 mL of solubilization buffer (Buffer B), and 1 mL fractions were collected from each wash. To eliminate non-specifically bound proteins, the resin underwent customized Ni-NTA washes for each protein, and 1 mL fractions were collected. Subsequently, elution of the his-tagged proteins bound to the column was conducted using 20 mL of Ni-NTA elution buffer, and the resulting eluates were collected in 10 equal fractions for subsequent analysis via Polyacrylamide Gel Electrophoresis (PAGE).

2.10.4 Size Exclusion Chromatography (SEC)

To further purify the protein samples were separated by virtue of hydrodynamic size using Size Exclusion Chromatography (SEC). SEC of protein samples was performed using a Superdex® 200 Increase 10/300 GL gel filtration using a ÄKTA pure 01 chromatography system (Amersham) controlled by UNICORN 4.0 chromatography programming (Amersham).

The column was equilibrated with 30 ml of the running buffer (50 mM tris, 150 mM NaCl, pH 7.5 r.t) before starting. Samples for SEC purification were collected and concentrated to 500 µL in an Amicon® Ultra Centrifugal Filter (Merck Milipore). A 3,000 molecular weight cut off (MWCO) was used for Rv3781c/Rv3783c and Rv1217c/Rv1218c samples and a 10,000 molecular weight cut off (MWCO) for Rv1272c/Rv1273c. Samples were concentrated by centrifugation at 4,000 rpm, 4°C in an Eppendorf Multi-purpose centrifuge. 300 µL of the concentrated samples were loaded into a 500 µL sample loop and injected onto the column. The column was run for 1 CV at 0.5 mL/min in SEC Running Buffer, with 0.5 ml fractions being collected. Elution absorbance was observed at 280 nm and 254 nm to separately distinguish the elution of protein and SMA.

2.11 Polyacrylamide Gel Electrophoresis (PAGE)

2.11.1 Sodium Dodecyl Sulphate-PAGE (SDS-PAGE)

10% handmade SDS-PAGE gels were prepared using the recipe shown in table 3 while Precast Mini-PROTEAN TGX Stain-Free Gels (BIO-RAD) were used for 4-15 % and 4 -20 % gradient gels. SDS-PAGE gels were mounted into a vertical electrophoresis tank (Bio-Rad) and immersed in 1X running buffer (Tris-Glycine-SDS buffer). Samples were prepared by adding 15 µL of each sample into a micro-centrifuge tube with 5 µL of 4X NuPAGE® LDS sample buffer (Novex®, ThermoFisher Scientific) and 2 µL of 10X NuPAGE® sample reducing agent (Novex®, ThermoFisher Scientific). Then, 15 µL of each sample was loaded into each well alongside 5 µL of Protein marker VI (pre-stained), peqGOLD (VWR Peqlab) (figure 2.4). Gels were run at 180 V for 50 minutes or until sample dye reached the bottom of the gel.

	Stacking Gel	**Resolving Gel**
	4%	10 %
30% Acrylamide/bis	1.33 mL	5 mL
0.5M Tris-HCl, 0.4% SDS pH 6.8	2.5 mL	-
1.5M Tris HCl, 0.4% SDS pH 8.8	-	3.75 mL
dH2O	6 mL	6 mL
TEMED	10 µL	15 µL
10% APS	100 µL	150 µL
Total Volume	10 mL	15 mL

Table 2.3: SDS-PAGE gel recipe.

2.11.2 Coomassie Staining

To visualized proteins bands, gels were stained with Coomassie staining. Gels were removed out from their glass plates and stained with Quick Coomassie Stain (QC Stain) (Neo-Biotech) following the manufacturer's instructions. In summary, gels were washed with deionised water, then stained with around 25 ml of QC Stain and incubated on rocker overnight at room temperature. After the overnight incubation QC Stain was removed and exchanged with deionised water and distained on rocker for 2 hours.

2.11.3 Western Blotting

To visualized proteins by western blotting, gels were removed from their glass plates and put into deionised water. Proteins were transferred from the gels to a nitrocellulose membrane using the Trans-Blot Turbo Transfer System (Bio-Rad) with Trans- Blot® TurboTM Mini Nitrocellulose Transfer Packs (Bio-Rad). Transfer was accomplished using the preprogramed 'Mixed MW' protocol (Trans-Blot Turbo Transfer System (Bio-Rad)) which applied 1.3 V up to 25 V for 7 minutes. After transfer, 30 ml of blocking buffer (5% Skim Milk Powder in Phosphate buffer saline (PBS), 1 tablet in 100 mL of distilled water (Thermo Scientific Oxoid) with 0.1 % Tween 20) was added to the membrane and incubated for 30 minutes on a rocker at room temperature. After the incubation, blocking buffer was discarded and substituted with 2 μL of 6xHis monoclonal antibody (Clontech) primary antibody in 10 ml of 5% Skim Milk Powder in PBS with 0.1 % Tween 20, and incubated on a rocker overnight at 4°C. Then membrane was washed 3 times with PBS with 0.1 % Tween 20 for 10 minutes on a rocker at room temperature. After the washes, 2 μL of Goat Anti-Mouse IgG Antibody [HRP] (2BScientific) secondary antibody in 10 ml of 5% Skim Milk Powder in PBS with 0.1 % Tween 20 was added to the membrane and incubated on rocker at room temperature for 2 hours. Then membrane was washed again 3 times with PBS with 0.1 % Tween 20 for 10 minutes on a rocker at room temperature. To visualize his tag proteins EZ-ECL Enhanced Chemiluminescence Detection Kit for HRP (Biological Industries) was used. In order to make the detection solution, 1 ml of each supplied solution was combined and

allowed to equilibrate for at least 5 minutes. The solution was directly poured on top of the membrane and incubated for 1 minutes, before visualisation of the bands using chemiluminescence imaging on a Uvitec Gel Imaging System.

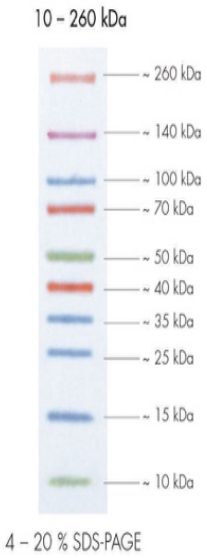


Figure 2.4. Protein marker VI (pre-stained), peqGOLD (VWR)

2.12 Biophysical Analysis

2.12.1 Circular Dichroism (CD)

CD measurements were made using a JASCO J-1500 Circular Dichroism Spectrophotometer controlled by Spectra Manager Software. A 1 mm pathlength quartz cuvette was used for standard spectra measurements. CD samples were prepared by dialyzing the sample into 20 mM fluoride phosphate (to reduce absorbance at low wavelengths) and 100 mM fluoride acetate pH 7.5, at a concentration of 0.1 mg/mL. CD data was collected (see table for parameters) between X and Y nm data was cropped for wavelengths were the high-tension voltage exceeded 600 V.

Channel 1	CD
Channel 2	HT
Channel 3	Abs
Start	350 nm
End	190 nm
Data pitch	1.0 nm
Scanning mood	Continues
Scanning speed	100 nm/min
D.I.T	1 second
Bandwidth	1 nm
Number of accumulations	3

Table 2.4. Parameters for circular dichroism (CD) analysis

Baseline spectra was measured using the same buffer that is been used to prepare sample and similarly average. The final spectra were obtained by subtracting the baseline spectra from sample spectra which been smoothed by the Savitzky-Golay method within the Spectra Manager software.

2.12.2 Analytical Ultracentrifugation (AUC)

Sedimentation velocity experiments were implemented by Pooja Sridhar at the Birmingham Biophysical Characterisation Facility, University of Birmingham. Analysed AUC samples were prepared in in 50 mM Tris, 150 mM NaCl, pH 7.5. The experiments were conducted using a Beckman Coulter Proteome Lab XL-I Analytical Ultracentrifuge equipped with AN-Ti50 rotor, operating at 20°C, and set at speeds at 40,000 rpm, depending on the estimated molecular weight of the sample. Monitoring of the samples was performed through absorbance measurements at 280 nm. Data was analysed by SEDFIT software (Schuck, 2000).

2.13 *Mycobacteria smegmatis* culture and Radio Label

2.13.1 Culture Media

2.13.1.1 Tryptic Soy Broth (TSB)

Tryptic Soy Broth (TSB) (Merck Milipore) was made as per the manufacture's protocol using deionised water. The media was autoclaved in order for 20 minutes at 121^{o}C, 1.4 bar, and then supplemented with 10 % Tween 80 (V/V). Media was left in room temperature to cool down before addition with antibiotics and inoculation.

2.13.1.2 Tryptic Soy Broth (TSB) agar

TSB agar was made by dissolving TSB broth as mentioned above with the addition of 12 g/L of agar granules (Merck Milipore). TSB agar was autoclaved for 20 minutes at 121^{o}C, 1.4 bar. TSB agar plates were poured and left to set on the bench under sterile conditions and stored at 4 ^{o}C.

2.13.1.3 Middlebrook 7h9 Broth

Middlebrook 7h9 Broth powder (Merck Milipore) was made as per the manufacture's protocol in deionised water and autoclaved for 20 minutes at 121^{o}C, 1.4 bar. Media was left in room temperature to cool down before addition with antibiotics and inoculation.

2.14 Bacteria Culture

2.14.1 Growing *M. smegmatis*

A TSB agar plate with was streaked with *M. smegmatis W.T* and incubated for 2-3 days at 37 °C. A single colony from this plate was then used to inoculate in 5 mL of TSB Broth and incubated for 2 days at 37 °C. 1 mL of the culture was added to a fresh 100 mL of Middlebrook 7h9 Broth. OD_{600} readings were taken immediately, OD_{600} should be around 0.05-0.1 at this point. Both cultures were incubated at 37°C with shaking at 180 rpm and mentored until the OD_{600} reached between 0.2-0.4.

2.15 Lipids analysis

2.15.1 Radioactive labelling of *M. smegmatis* cultures

M. smegmatis cultures were grown following the protocol in section 15.1. *M. smegmatis* cultures were labelled with ^{14}C acetic acid (1 µCi/µl Pekin Elmer) when the culture OD_{600} reached between 0.2 – 0.4 and incubated at 37 °C with shaking at 180 rpm for 12 hours. Cultures were harvested by centrifugation at room temperature for 10 minutes at 4,000 rpm in an Eppendorf Multi-purpose centrifuge and stored in – 20 ^{o}C until used.

2.15.2 Extraction of *M. smegmatis* lipids with SMA

Pellets containing *M. smegmatis* labelled with ^{14}C acetic acid were washed with PBS and transferred into 10 ml glass tubes sealed with polytetrafluoroethane (Teflon®)-lined screw cap. Pellets were dried at 55 °C under air flow and resuspended in equal amount of extraction buffers with and without glass beads, acid washed (425 – 600 μM) (SIGMA). The glass tubes were agitated in a rotating mixer at room temperature for 2 hours. The samples were then centrifuged for 10 minutes at 4,000 rpm and supernatants were transferred to a new glass tube.

Extraction Buffer 1	50 mM TrisHCl pH 8.0 (r.t) 10% (v/v) glycerol 500 mM NaCl
Extraction Buffer 2	Extraction Buffer 1 + 5 % (w/v) SMA
Extraction Buffer 3	Extraction Buffer 1 + 2.5 % (w/v) SMA
Extraction Buffer 4	Extraction Buffer 1 + 1 % (w/v) SMA

Table 2.5. Extraction buffers used for *M. smegmatis* lipids extraction.

2.15.3 Total lipids extraction via organic solvent

The dried pellet was treated with 2 mL of a solution consisting of 10:10:3 (chloroform: methanol: water) and incubated at 50°C for 3 hours. Afterward, the tubes were centrifuged at 3,000 rpm for 10 minutes in an Eppendorf Multi-purpose centrifuge, and the resulting supernatant was transferred to a new glass tube. Following this, an additional 2 mL of the 10:10:3 (chloroform: methanol: water) was added to the pellet and incubated at 50°C for 1 hour. The tubes were again centrifuged at 3,000 rpm for 10 minutes, and the obtained supernatant was added to the previous one. To this combined solution, 1.75 mL of chloroform and 0.75 mL of water were added and mixed in a rotor for 15 minutes. It was

then centrifuged for 10 minutes at 3,000 rpm in the Eppendorf Multi-purpose centrifuge. The resulting lower organic phase was transferred to a new glass tube using a glass Pasteur pipet. Subsequently, 2 mL of a solution comprising 3:47:48 (chloroform: methanol: water) were added to the tube containing the lower organic phase, mixed in a rotor for 15 minutes, and centrifuged for 10 minutes at 3,000 rpm. The upper aqueous phase was discarded, and another 2 mL of the 3:47:48 (chloroform: methanol: water) solution was added to the tube, mixed for 2-3 minutes on the rotor, and centrifuged for 15 minutes at 3,000 rpm. The resulting lower organic phase was then transferred to a new glass tube and evaporated and concentrated under compressed air.

2.15.4 Lipids analysis

1600 µl of the supernatants obtained in section 16.2 were transferred to clean glass tubes followed by 800 µl of methanol and 800 µl of chloroform and mixed in a rotating mixer for 15 minutes. Tubes were then centrifuged for 5 minutes at 4,000 rpm in an Eppendorf Multi-purpose centrifuge and transferred to a new glass tube. The lower organic phase containing lipids was transferred using glass Pasteur pipet making sure none of the upper aqueous phase was transferred. Another 800 µl of methanol and 800 µl of chloroform was added to the upper aqueous phase, which was then vortexed followed by centrifugation for 5 minutes at 4,000 rpm in an Eppendorf Multi-purpose centrifuge. the lower organic phase containing lipids was transferred into the previous tube containing lower organic phase from the first extraction. Samples were then evaporated and concentrated under compressed air. 100 µl chloroform: methanol (2:1) were added to the dried lipids in order to resolve on thin layer chromatography.

2.15.5 Thin Layer Chromatography

Radioactive ^{14}C acetic acid samples were loaded on silica TLC plate F254 (Merck). In the case of one-dimensional (1D) TLC, samples were resolved in chloroform: methanol: water (80: 20: 2) solvent system. In order to run two-dimensional (2D) TLC, 6.7 x 6.7 cm square silica plate were used. Samples were resolved in the solvent systems indicated in table 6. When solvent reached 1 cm away of the top, TLCs Plates were removed from the TLC tank and allowed to air dry before exposure to Carestream BIOMAX MR film for 5-7 days.

In the case of non-radioactively labelled lipids, 5-10 µl of samples were loaded into TLC plates and resolved on the appropriate solvent systems. TLC plates were then air dried before visualizing lipids and carbohydrate by spraying with 5% ethanolic molybdophosphoric acid (MPA) and 1% ethanolic 1-Naphthol, respectively, and then dried at 100 °C using a heat gun.

System	Direction	Solvents	Proportions
A	1 (x3)	Petroleum ether (60-80): ethyl acetate	98:2
	2	Hexane: acetone	98:2
B	1 (x3)	Hexane: acetone	92:8
	2	Toluene: acetone	95: 5
C	1	Chloroform: methanol	96: 4
	2	Toluene: acetone	80: 20
D	1	Chloroform: methanol: water	250: 75: 4
	2	Chloroform: acetone: methanol: water	100:120: 5: 6
E	1	Chloroform: methanol: water	10: 5: 1
	2	Chloroform: acetic acid: methanol: water	40: 25: 3: 6

Table 2.6: Solvent Systems used for TLC analysis.

CHAPTER 3

SOLUBILIZATION OF *MYCOBACTERUIM TUBERCULOSIS* ABC-TRANSPORTER Rv1273c AND Rv1272c COMPLEX USING STYRENE MALEIC ACID

3.1 Introduction

Rv1273c and Rv1272c are membrane proteins found in *Mycobacterium tuberculosis* and belong to the ATP-binding cassette (ABC) transporter family (Gupta *et al*., 2020a). They function as exporter ABC transporters involved in the transportation of macrolides and multi-drugs, alongside other genes such as DrrABC, Rv1686c/87c, Rv1456c/57c/58c, and Rv2686c/87c/88c (Cassio Barreto de Oliveira and Balan, 2020). Rv1273c, known as an ATP-binding protein in the ABC transporter, pertains to the MsbA subfamily with a molecular weight of 62.11 kDa (Braibant *et al*., 2000). While the structure of Rv1273c remains unidentified, it is proposed to be a typical ABC transporter, possessing a nucleotide-binding domain (NBD) and a membrane-spanning domain (MSD) (Gupta *et al*., 2020b). Additionally, the sequence of Rv1273c reveals the presence of Walker A and Walker B motifs, along with the 6-amino acid Linton and Higgins motif downstream of the Walker B motif (Holland and Blight, 1999). Although there is no information directly linking Rv1273c to Mycobacterium infection, the absence of information does not decline the possibility of its role or connection (Gupta *et al*., 2020a). Studies also suggest that Rv1273c aids in the survival of Mycobacterium inside macrophages by regulating cell wall architecture and lipid composition (Gupta *et al*., 2020a). With an unidentified function, Rv1273c may be involved in drug transport through the membrane and in the reversible transport of ligands based on cellular energy levels, given its significant sequence identity to MsbA (Balakrishnan *et al*., 2004).

Rv1272c, with a molecular weight of 68.27 kDa, is implicated in the importation of long-chain fatty acids into *Escherichia coli*. There is a suggestion of its homodimer function, supported by a 34% identity with its neighbouring gene, Rv1273c (Martin and Daniel, 2018)). Similar to ABC transporters, Rv1272c shares a nucleotide-binding domain containing canonical motifs such as Walker A, Walker B, ABC Signature, Q-Loop, and H-Loop (Martin and Daniel, 2018).

The aim of this part of the project was to see whether the SMALP method could be used to extract intact complexes of Rv1272c and Rv1273c. Plasmids containing the two proteins were transformed into *E. coli* BL21 (DE3) and protein expression was induced with IPTG. Rv1272c and Rv1273c are proposed to form a hetero dimer which includes Rv1273c in the membrane and Rv1272c forming the extracellular domain (figure 3.1). For these experiments on Rv1273c was tagged with a 12His tag meaning that success in this purification would be defined by extraction and purification of the full complex intact. If both (the tagged and untagged protein) are present in the final pure sample, then the SMALP contains the full complex. The experiment could have been set up with both proteins tagged, however in this case if both proteins appear in the final sample there is no evidence that they have formed a complex and instead might be present as individual proteins in solution.

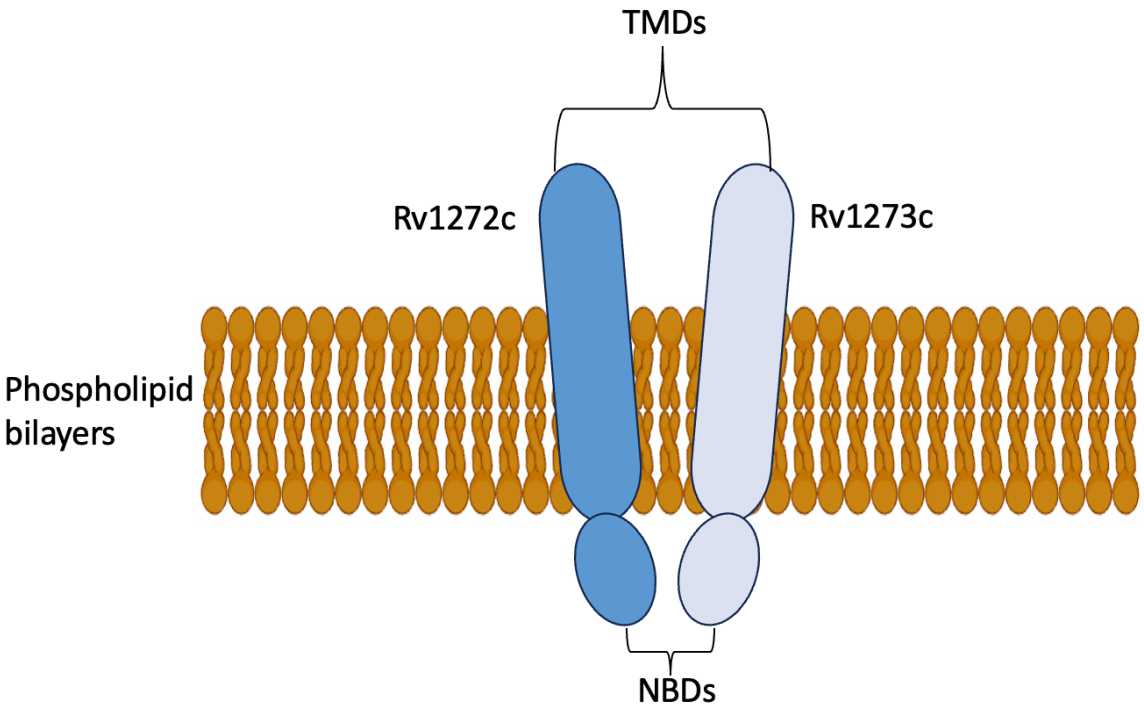


Figure 3.1: Schematic figure of ABC transporters showing the two proteins and the proposed functioning transporter Rv1272c/Rv1273c.

To achieve samples of high purity (Ideally >80%) for future studies a purification protocol had to be developed. Conventionally proteins in SMALPs are purified using a combination of Ni-NTA chromatography followed by SEC (Lee *et al*., 2016). In this chapter I show the development of a purification protocol that provide Rv1273c at high purity and show that maintaining the integrity of the complex between Rv1272c and Rv1273c is challenging using this system.

3.2 Solubilizing Rv1273c/Rv1272 using styrene maleic acid:

3.2.1 Initial purification of Rv1273c/Rv1272c:

Proteins Rv1273c and Rv1272c were both co-expressed in the BL21 (DE3) cell strain. It is important to note that only Rv1273c was tagged with 12 histidine residues while Rv1272c remained untagged. The approach to purification being to attempt to purify the complex intact via the tag on one of the 2 proteins in the complex. Protein induction was initiated when the OD_{600} reached approximately 0.6. Post induction the culture was left to grow at (18°C) overnight. Following induction, cells were harvested through centrifugation and processed to generate a membrane fraction which was resuspended and mixed with SMA, achieving a final concentration of 2.5% SMA. The SMA-solubilized membrane fraction was bound to pre-prepared Ni-NTA overnight at 4°C. To purify Rv1273c and Rv1272c from other solubilized proteins, the resin was transferred to a gravity flow column and any unbound material was collect. The resin was washed with 20 mL of solubilization buffer B, followed by wash with 10 mL of low imidazole concentration (20 mM) eliminate non-specifically bound contaminants. We collected the respective fractions during these wash steps. The His-tagged proteins were subsequently eluted with 10 mL using a higher imidazole concentration of 500 mM at pH 8. Eluted fractions were collected and analysed using SDS-PAGE (Figure 3.2), including the flow-through and the initial two washes using buffer free imidazole.

Notably, a significant amount of the solubilized material did not bind to the resin, as observed in the flow-through fraction (FT). The washes with free imidazole (W1-2) and the low imidazole concentration wash (data not shown) removed minimal protein from the column indicating that most non-specific protein had been removed from the column in the first stage. The elution fractions with a high imidazole concentration of 500 mM (E1-2) did not elute much or any of the protein of interest, which has a molecular weight of 62.11 kDa. However, these fractions did contain two proteins with molecular weights around 100 kDa

and 45 kDa. In contrast, elution fractions (E3-6) successfully eluted protein with a mass of around the 60 kDa mark, along with a higher yield of the two proteins around 100 kDa and 45 kDa. The remaining elution fractions (E7-8) showed a decreasing presence of the three proteins, with no additional proteins eluting in the subsequent fractions.

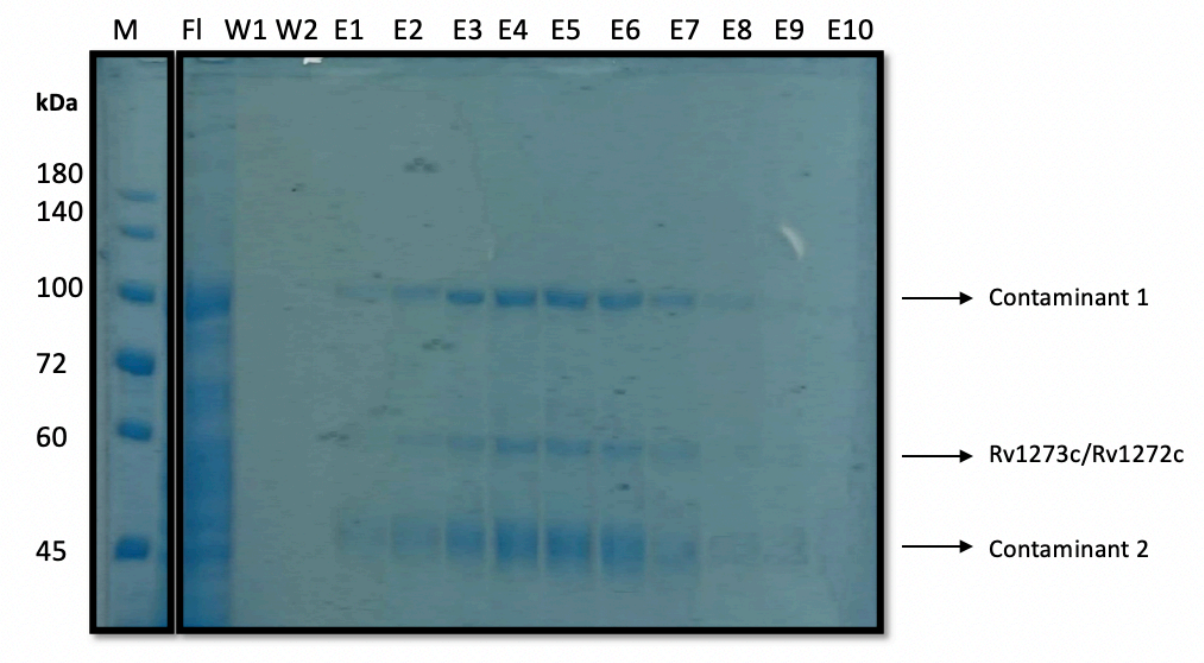


Figure 3.2: Rv1273c/ Rv1272c Ni-NTA Purification. SMA solubilized membrane were incubated with Ni-NTA overnight, then transferred to a gravity flow column, and flow through (FT) collected. Resin then washed 4 times with 5 ml solubilization buffer and collected 4 washes (shown the first two wash W1 and W2), then washed with 10ml of 20 mM imidazole (data not shown). Then His-Tagged proteins were then eluted with 10 ml of 500 mM imidazole buffer (1- 10). Fractions were analysed by electrophoresis on an 10 % handmade SDS-PAGE and visualized by Coomassie stained.

The initial purification protocol has shown promising results in solubilizing the protein, which was subsequently eluted from the Ni-NTA column. The Coomassie-stained gel revealed the protein of interest alongside two contaminating proteins. My plan was to enhance the purification in the next step by modifying the initial protocol. This modification involved reducing the concentration of the imidazole wash and adjusting the volume to remove as much of the contaminants as possible.

3.2.2 Reducing imidazole wash concentration:

Rv1273c and Rv1272c were expressed and solubilized using the same protocol as described in Section 3.2.1. Subsequently, the SMA-solubilized membrane fraction was subjected to binding with pre-prepared Ni-NTA overnight at 4°C. The resin underwent a series of four washes with free imidazole buffer (B), with each wash using 5 ml of buffer. This was followed by a single wash with a low imidazole concentration (10 mM) involving 5 ml of solution, intended to remove non-specifically bound contaminants. The respective fractions were collected during these wash steps. The His-tagged protein was subsequently eluted using a higher imidazole concentration of 500 mM at pH 8. Eluted fractions were collected and analysed using SDS-PAGE (Figure 3.3), alongside the flow-through and the initial three washes using buffer free imidazole. The results indicated that the flow-through (FT) and the first wash (W1) contained a significant amount of solubilized material that did not bind to the resin. The subsequent washes with buffer without imidazole (W2-3) effectively removed any untagged proteins, while it is important to note that no detectable proteins were removed from the column. In contrast, elution with a high concentration of imidazole (500 mM) (E2-5) resulted in the elution of proteins around 60 kDa, 100 kDa, and 45 kDa. However, the first elution (E1) did not yield any proteins, and elution from E6 to E10 demonstrated a reduction in the higher protein band around 100 kDa and the lower protein band around 45 kDa, along with a decreased amount of the protein at the 60 kDa.

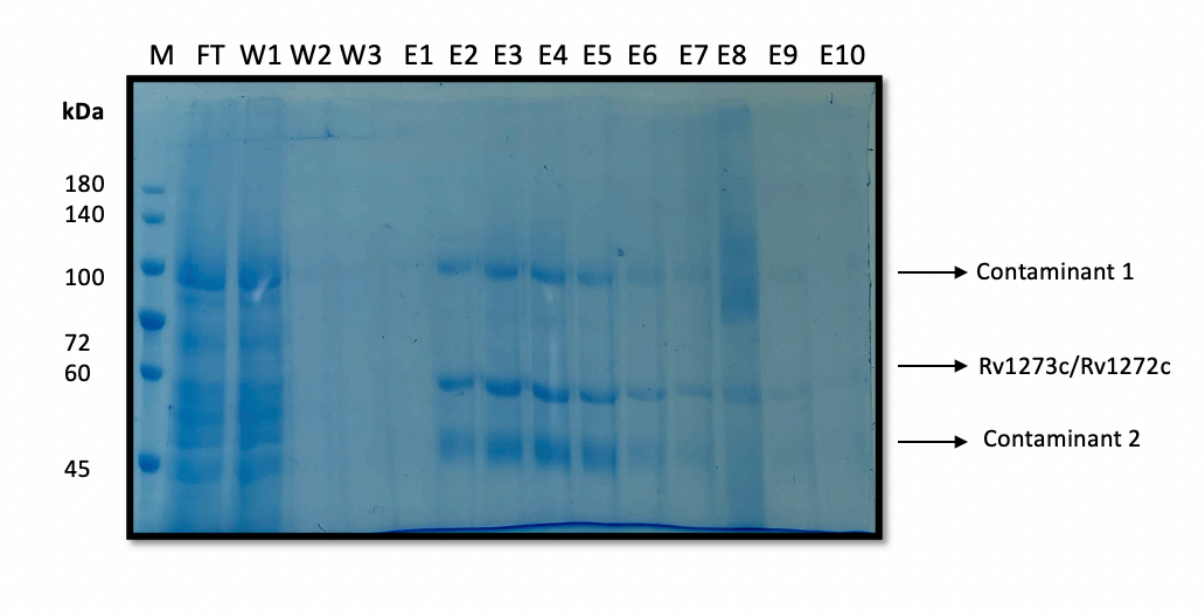


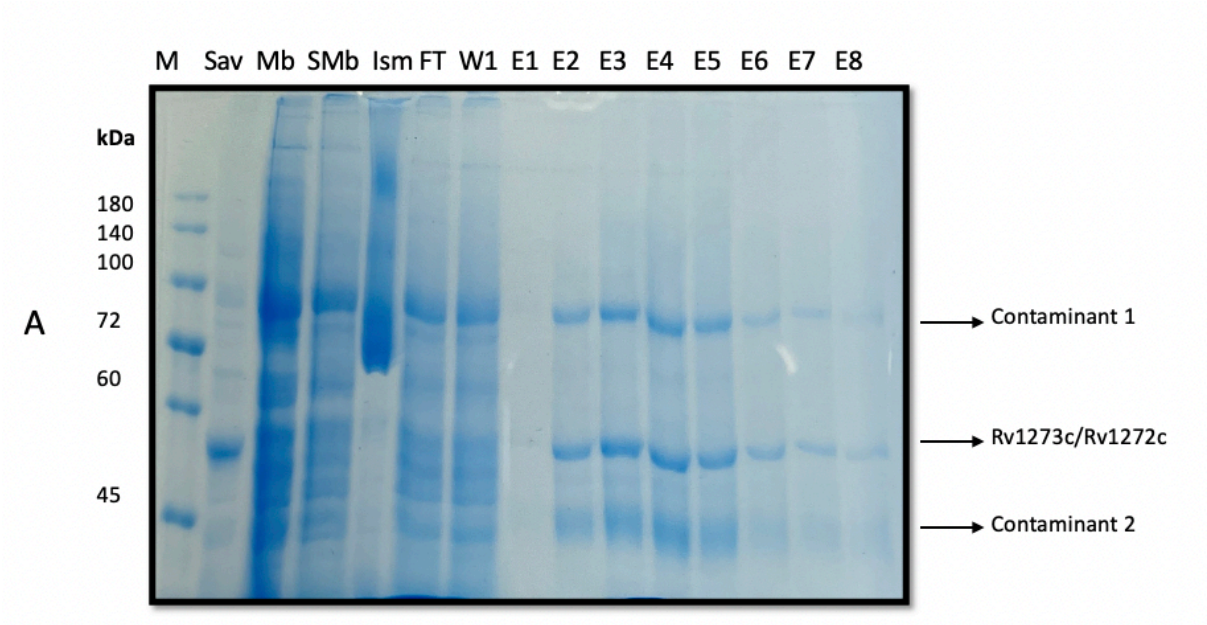
Figure 3.3: Rv1273c/ Rv1272c Ni-NTA Purification. SMA solubilized membrane were incubated with Ni-NTA overnight, then transferred to a gravity flow column, and flow through (FT) collected. Resin then washed 4 times with 5 ml solubilization buffer and collected 4 washes (shown the first three washes), then washed with 5 ml of 10 mM imidazole pH 8 (data not shown). Then His-Tagged proteins were then eluted with 10 ml of 500 mM imidazole buffer (1- 10). Fractions were analysed by electrophoresis on an 10 % handmade SDS- PAGE and visualized by Coomassie stained.

This minor protocol adjustment has resulted in an increase in protein yield, and it exhibits a similar banding pattern as the initial protocol described in section 3.2.1, especially in the early elution fractions. Given that the molecular weight of Rv1273c is 62.11 kDa, we are confident that the band around 60 kDa represents our target protein. To confirm this western blot was carried out using an anti-6XHis antibody to specifically identify the target.

3.2.2.1 Anti-6XHis antibody via western blot

The objective was to determine whether SMA could effectively solubilize the Rv1273c and Rv1272c proteins and to identify if the protein band appearing around 60 kDa represented the protein of interest. For this analysis, elution fractions (E1-8), along with the membrane (Mb), solubilized membrane (SMb), Insoluble material (Ism), and 6XHis-Sav1866 (Sav1866), were subjected to analysis via SDS-PAGE (Figure 3.4a). Subsequently, the proteins were

transferred to a nitrocellulose membrane for western blotting (Figure 3.4b) using anti-His Tag antibodies. The SDS-PAGE results revealed that a substantial amount of the expressed protein was located in the membrane, accompanied by other proteins. Furthermore, it was evident that the amount of expressed protein that had been solubilized was less than half of the protein expressed in the membrane. This observation was based on the presence of two bands of membrane (Mb) and solubilized membrane (SMb) of the western blot, respectively. Additionally, the western blot indicated the absence of the lower band around 45 kDa, which had been visible on the SDS-PAGE. In terms of proteins around 60 kDa or lower, faint bands were observed on the western blot, corresponding to the same bands seen on the SDS-PAGE in elution fractions (E2-4). There were also faint bands around 120 kDa, which could potentially indicate oligomerization. It is worth noting that the faint appearance of these bands might be due to the strong signal of the positive control or the lower abundance of protein. In the elution fractions (E5-8), the bands around 60 kDa and 100 kDa disappeared on the western blot (Figure 3.4b), in contrast to their presence on the SDS-PAGE (Figure 3.4a).



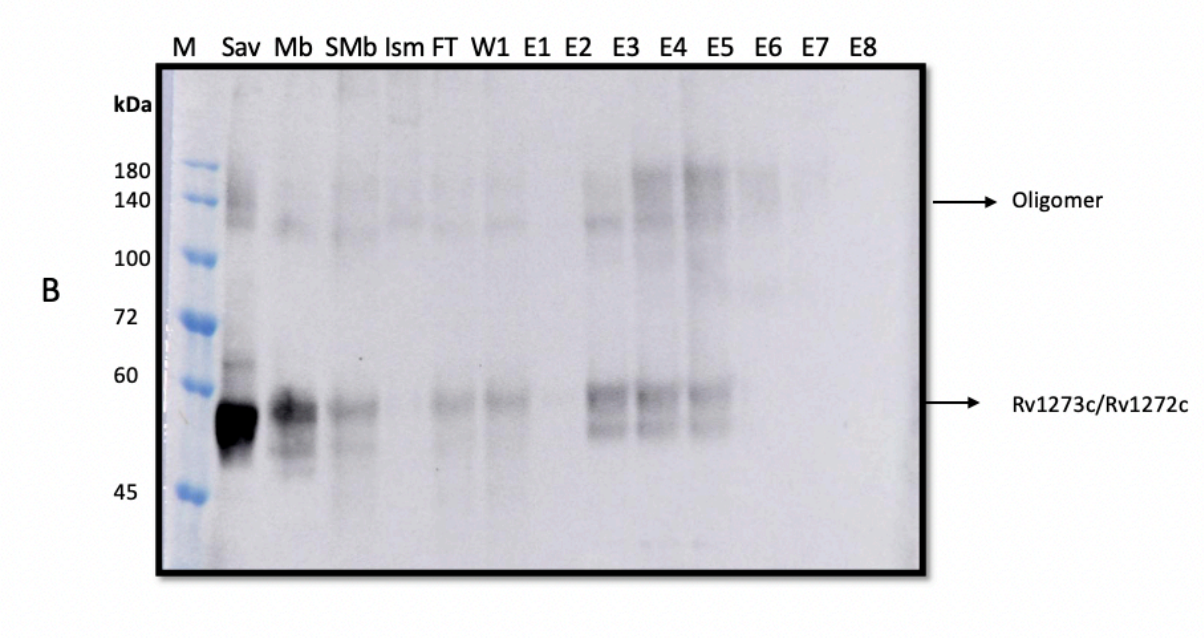


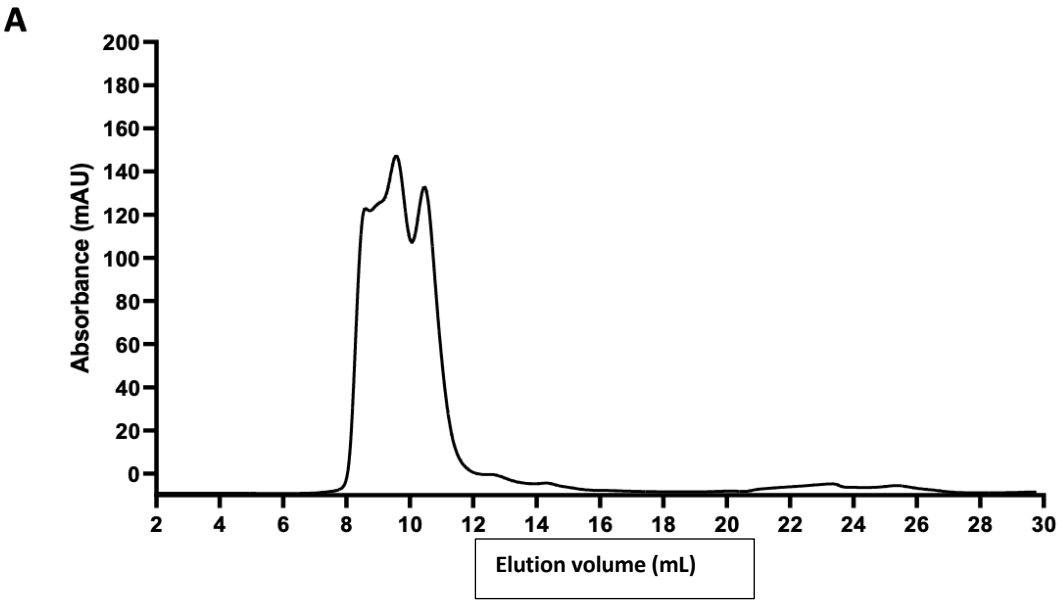
Figure 3.4: Rv1273c/ Rv1272c Ni-NTA Purification. A) SMA solubilized membrane were incubated with Ni-NTA overnight, then transferred to a gravity flow column, and flow through (FT) collected. Resin then washed 4 times with 5 ml solubilization buffer and collected 4 washes (shown the first wash), then washed with 5 ml of 10 mM imidazole pH 8 (data not shown). Then His-Tagged proteins were then eluted with 10 ml of 500 mM imidazole buffer pH 8 (E1- E8). Fractions were analysed by electrophoresis on an 10 % handmade SDS- PAGE and visualized by Coomassie stained. Membrane (Mb), Solubilized membrane (SMb), Insoluble material (Ism). B) The SDS-PAGE was also transferred to a nitrocellulose membrane and probed with an anti-His Tag antibody**.**

The western blot has confirmed that the band around 60 kDa is likely to be the target protein. It also detected some proteins that might be oligomers, possibly including Rv1273c. To address this and remove contaminants, it was proposed that size exclusion chromatography (SEC) should be used to separate the chosen protein target by virtue of size.

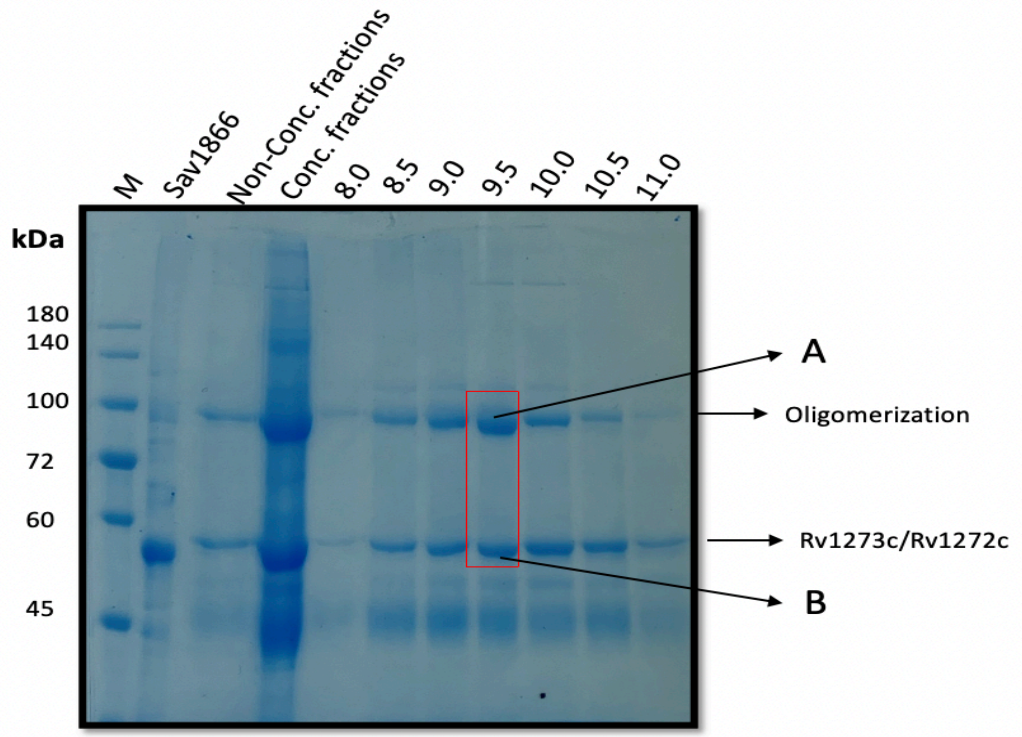
3.2.2.2 Size Exclusion Chromatography (SEC)

In the subsequent purification step, elution fractions were combined and concentrated using a 50,000 molecular weight cut-off (MWCO) concentrator. The resulting sample was then subjected to separation using a Superdex 200 Increase 10/300 size exclusion column.

The elution profile of the sample at 280 nm, as shown in (Figure 3.5a), revealed the elution of proteins. During this separation, the protein eluted across a moderate range of elution volumes, displaying four major peaks at 8.5 mL, 9 mL, 9.5 mL, and 11 mL. Fractions spanning this range were subjected to further analysis through SDS-PAGE and western blotting, as illustrated in (Figure 3.5b and 3.5c). Both Coomassie-stained gel and western blot confirmed that Rv1273c and Rv1272c eluted between 8.5 mL and 11 mL, detected by the presence of the His-Tag associated with Rv1273c within this elution range. The western blot analysis revealed two distinct bands for Rv1273c. One of these bands appeared at approximately 60 kDa, which closely aligns with the known molecular weight of Rv1273c, measuring at 62.11 kDa. The higher band observed above 100 kDa suggests the possibility of an oligomer that includes Rv1273c. Furthermore, the Coomassie-stained gel showed a corresponding band at the correct molecular weight for Rv1273c. Given that the two proteins, Rv1273c and Rv1272c, have similar molecular weights, they may not completely separate on the Coomassie gel. The mass spectrum data analysis of the two bands eluted in the 9.5 mL fraction reveals the presence of peptides from both Rv1272c and a variety of *E. coli* proteins known to bind to nickel resin in both bands labelled as "a" and "b."



B


M
Sav1866
Non-Conc. fractions
Conc. fractions
8.0
8.5
9.0
9.5
10.0
10.5
11.0
kDa
180
140
100
72
60
45
A
Oligomerization
Rv1273c/Rv1272c
B

C

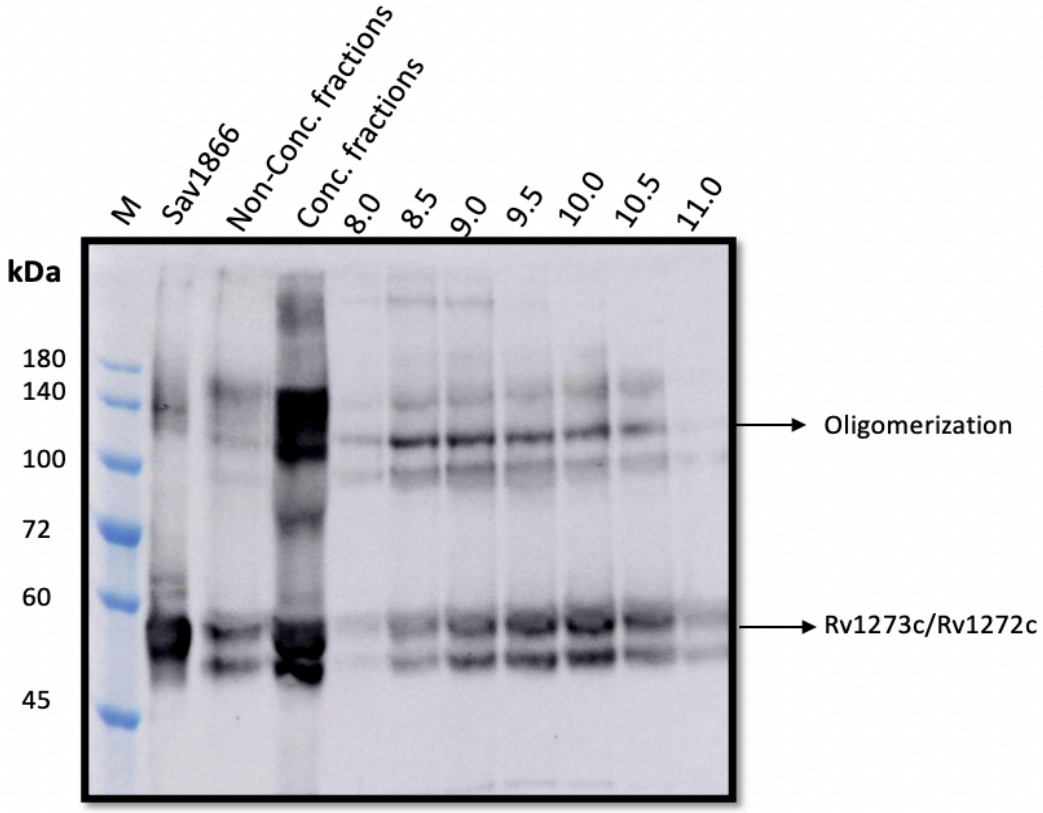

M
Sav1866
Non-Conc. fractions
Conc. fractions
8.0
8.5
9.0
9.5
10.0
10.5
11.0
kDa
180
140
100
72
60
45
Oligomerization
Rv1273c/Rv1272c

Figure 3.5: SEC purification of Rv1273c/ Rv1272c Elution fraction fractions from Ni- NTA purification were collected, concentrated, and injected onto a Superdex 200 Increase 10/300 GL column. The column was run at 0.5 mL/min collecting 0.5 mL fractions. **A)** SEC elution profile. Elution of protein from the column was monitored by absorbance at 280nm. **B)** SDS-PAGE of elution profile 10 % handmade SDS-PAGE of the SEC elution fractions analysed by electrophoresis and protein bands visualized by Coomassie staining**. C)** The SDS-PAGE was also transferred to a nitrocellulose membrane and probed with an anti-His-Tagged antibody**.**

The size exclusion chromatography profile displays four peaks with high absorbance, but there is no clear separation between them. Upon analysing the fractions, it is evident that there is a high yield of two proteins, one of which is the target protein and the other is likely an oligomer containing Rv1273. Additionally, when analysing the concentrated protein before size exclusion, two significant bands corresponding to those observed in the fractions are noticeable. While the purification has shown promising signs of improvement in protein yield and purity, it is not yet satisfactory, and further enhancements are needed. In addition to improve visualisation of the results of purifications precast gradient polyacrylamide gels to improve protein separation while following the same protocol as outlined in section 3.2.2.

3.2.3 Improving purification by using precast gradient gels

To improve the purification process, 4–15% precast gradient gels (Mini-Protean TGX, Bio-Rad, U.K.) were used. The expression, solubilization, and binding of Rv1273c/Rv1272c to the resin followed the same protocol outlined in Section 3.2.2. After an overnight incubation with Ni-NTA resin, we washed the resin and eluted the His-tagged proteins using imidazole. The eluted fractions were collected and subjected to analysis via SDS-PAGE and western blot (Figure 3.6a and 3.6b). This analysis was conducted alongside samples including the flow-through, the first wash with buffer-free imidazole, solubilized membrane (Sm), insoluble material (Ism), and 6XHis-Sav1866 (Sav1866). The transition to gradient gels had an

immediate impact, enhancing the separation of proteins throughout the gel. This improvement was particularly evident in the elution fractions. Notably, the flow-through (FT) and the first wash (W1) still contained a significant amount of solubilized material that did not bind to the resin. Washes with buffer free imidazole and low imidazole concentration buffer effectively removed any untagged proteins, with no detectable proteins removed from the column. However, it is important to mention that the quantity of solubilized protein remained relatively low when compared to the insolubilized fraction. During elution with a high concentration of imidazole (500 mM) in fractions labelled (E1-9), we observed four distinct bands across the Coomassie gel, representing eluted proteins at approximately 100 kDa, 60 kDa, 45 kDa, and 35 kDa. In terms of the western blot results, the His-tagged proteins were detected at their expected molecular weights, appearing as faint bands around 60 kDa throughout the elution fractions, except for (E9). These faint bands could be attributed to the strong signal of the positive control or potentially the protein's lower abundance. Moreover, in elution fractions (E4-7), there was a faint band around 35 kDa, corresponding to the band observed in the Coomassie gel at the same fractions. Interestingly, there was either a disappearance or less distinct bands observed around the 100 kDa marks on the western blot (Figure 3.6b) compared to their appearance on the SDS-PAGE (Figure 3.6a). Furthermore, the western blot showed a loss of bands around 45 kDa, which were present in the Coomassie gel (Figure 3.6a).

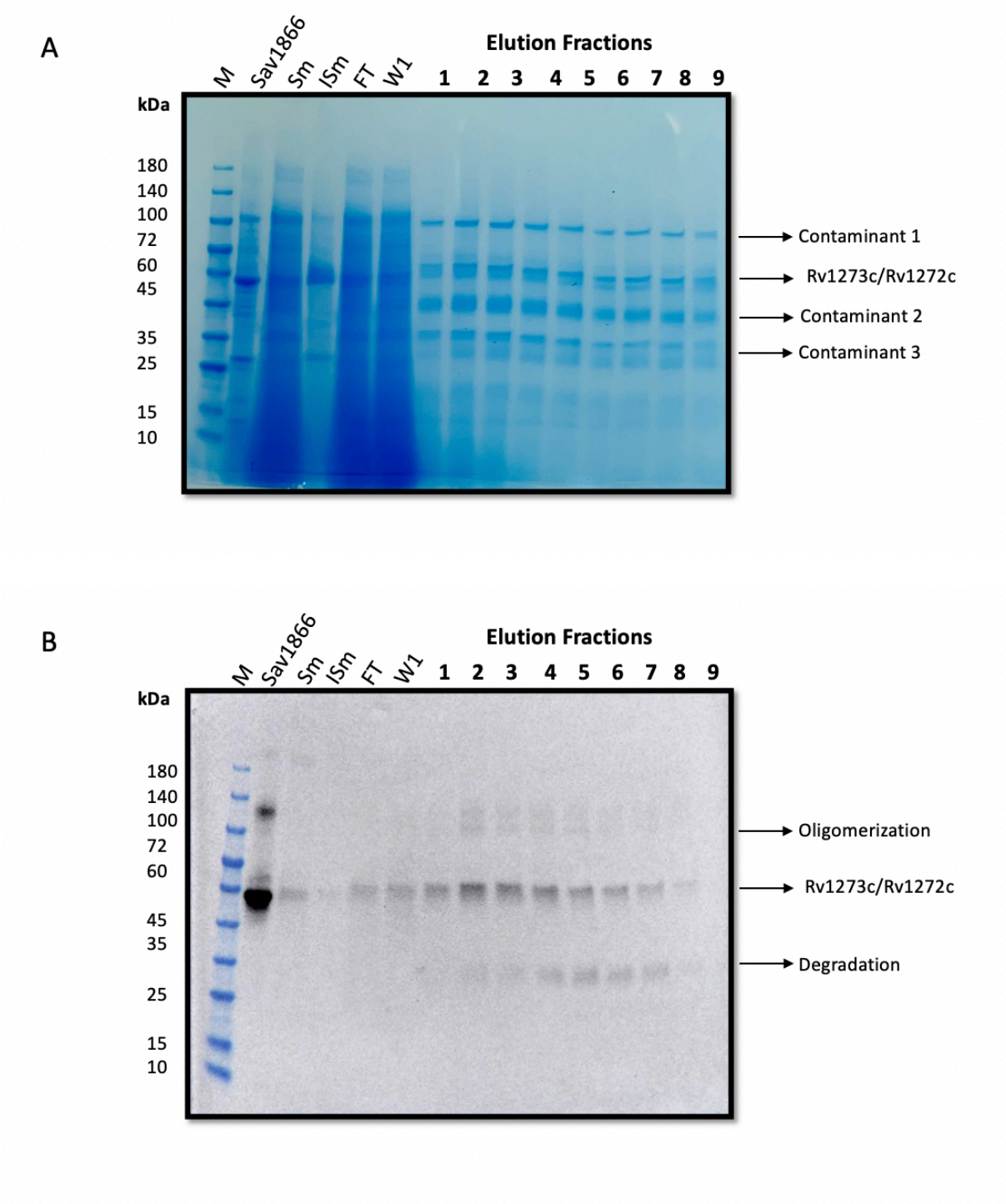
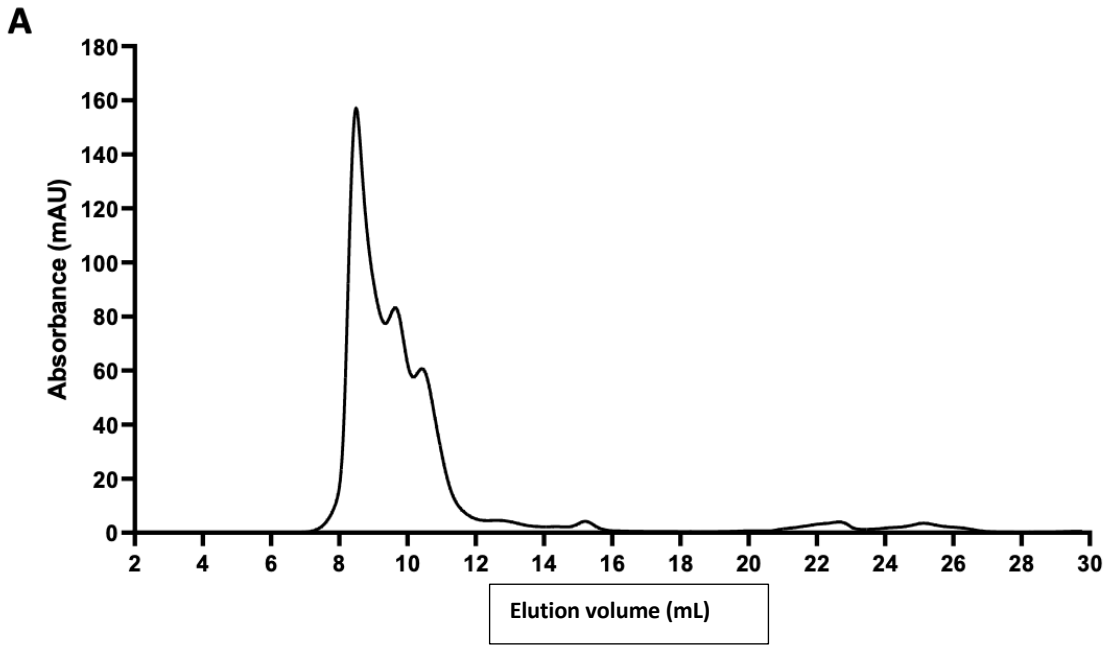


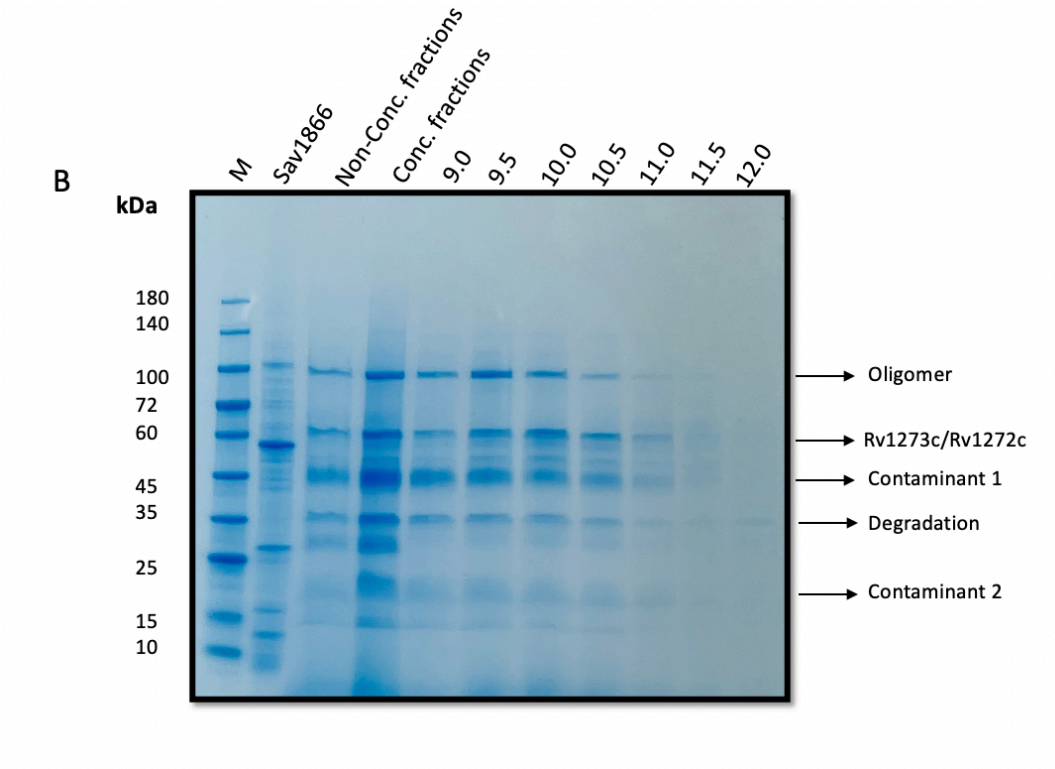
Figure 3.6: Rv1273c/ Rv1272c Ni-NTA Purification. A) SMA solubilized membrane were incubated with Ni-NTA overnight, then transferred to a gravity flow column, and flow through (FT) collected. Resin then washed 4 times with 5 ml solubilization buffer and collected 4 washes (shown the first wash), then washed with 5 ml of 10 mM imidazole pH 8 (data not shown). Then His-Tagged proteins were then eluted with 10 ml of 500 mM imidazole buffer pH 8 (E1- E9). Fractions were analysed by electrophoresis on an 4-15 % precast SDS- PAGE and visualized by Coomassie stained. Solubilized membrane (Sm), Insoluble material (Ism). B) The SDS-PAGE was also transferred to a nitrocellulose membrane and probed with an anti-His-Tagged antibody**.**

Switching to a gradient gel has demonstrated effective separation of the proteins in the elution fractions, with consistently high band intensity in the Coomassie-stained gel for all four eluted proteins. Unfortunately, this was not reflected in the western blot results, especially for the target protein. To address this issue and improve both protein purity and yield size exclusion chromatography was carried out in an attempt to remove some contaminants.

3.2.3.1 Size Exclusion Chromatography (SEC)

In the next phase of the purification process, fractions were pooled, concentrated and applied to Superdex 200 Increase 10/300 size exclusion column. The elution profile of the sample at 280 nm, as shown in (Figure 3.7a), indicates the elution of proteins. The proteins eluted over a moderate range of elution volumes, displaying three major peaks at 8.5 mL, 9.5 mL, and 10.5 mL. Fractions from this range were subjected to further analysis through SDS-PAGE and western blotting, as demonstrated in (Figure 3.7b and 3.7c). Both the Coomassie-stained gel and the western blot confirmed that Rv1273c and Rv1272c eluted between 9.0 mL and 11 mL, detected by the presence of the His-Tag associated with Rv1273c during the elution. The western blot analysis revealed two bands for both Rv1273c and Rv1272c, appearing on top of each other. These bands were observed at approximately 60 kDa, which closely aligns with the known molecular weight of Rv1273c at 62.11 kDa and the molecular weight of Rv1272c at 68.27 kDa. Additionally, there was a higher band observed above 100 kDa, suggesting the possibility of an oligomer that includes Rv1273c. Furthermore, the Coomassie-stained gel displayed a corresponding band at the correct molecular weight for Rv1273c. Given the similarity in molecular weights between the two proteins, Rv1273c and Rv1272c, it's possible that they do not fully separate on the Coomassie gel. Even with the second step of purification, there were still some contaminants observed on the Coomassie gel, indicating the need for further refinement in the purification process.


A
Absorbance (mAU)
180
160
140
120
100
80
60
40
20
0
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
Elution volume (mL)


B
kDa
M
Sav1866
Non-Conc. fractions
Conc. fractions
9.0
9.5
10.0
10.5
11.0
11.5
12.0
180
140
100
72
60
45
35
25
15
10
Oligomer
Rv1273c/Rv1272c
Contaminant 1
Degradation
Contaminant 2

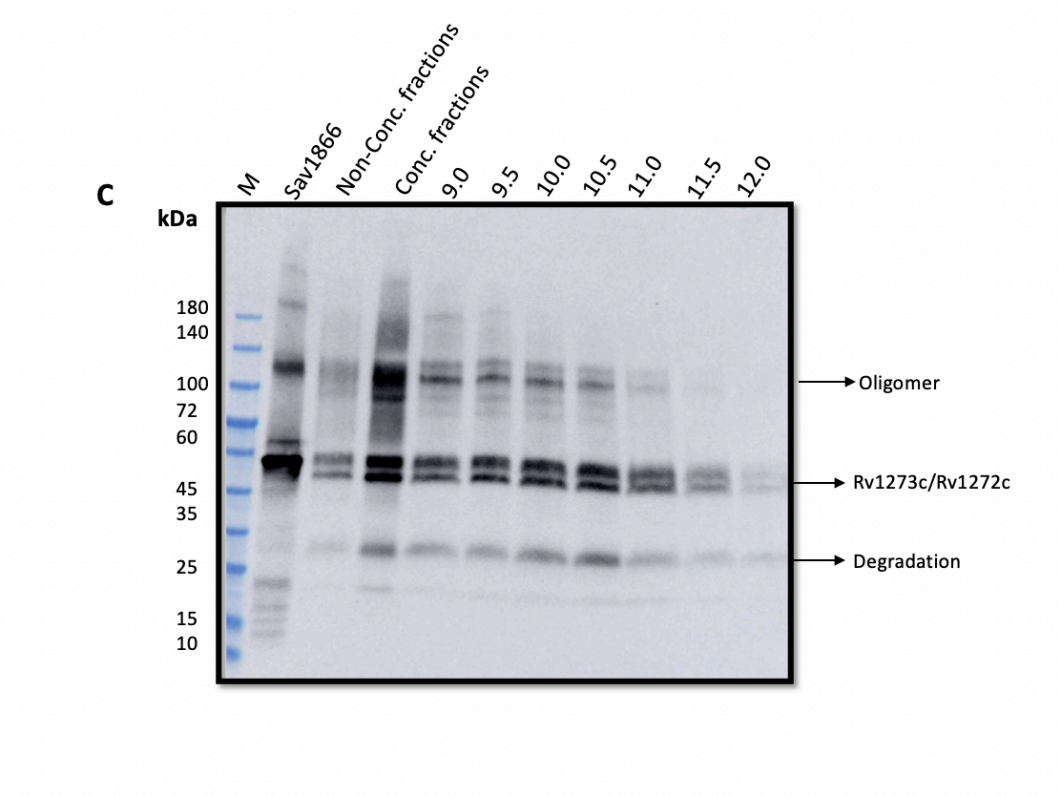


Figure 3.7: SEC purification of Rv1273c/ Rv1272c Elution fraction fractions from Ni- NTA purification were collected, concentrated, and injected onto a Superdex 200 Increase 10/300 GL column. The column was run at 0.5 mL/min collecting 0.5 mL fractions. **A)** SEC elution profile. Elution of protein from the column was monitored by absorbance at 280nm. **B)** SDS-PAGE of elution profile 4 -15 % Precast SDS-PAGE of the SEC elution fractions analysed by electrophoresis and protein bands visualized by Coomassie staining**. C)** The SDS-PAGE was also transferred to a nitrocellulose membrane and probed with an anti-His-Tagged antibody**.**

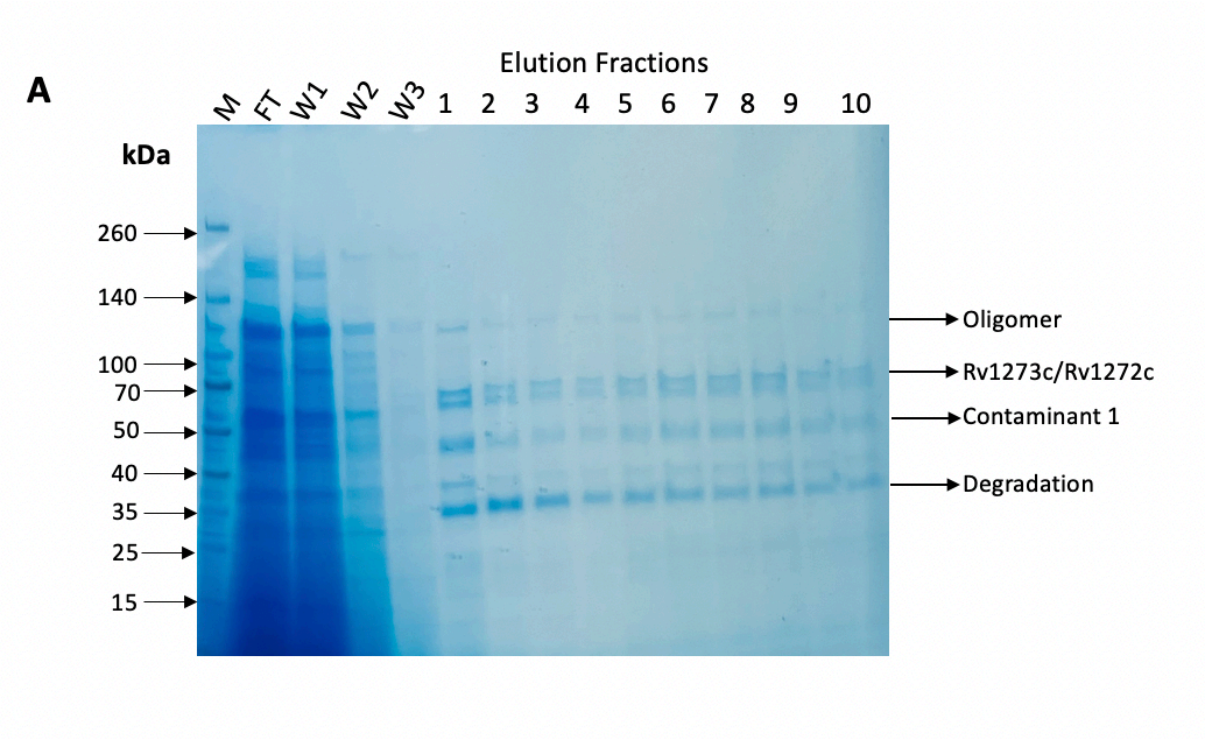
The size exclusion chromatography profile displayed three peaks with high absorbance. However, upon analysing the fractions within this profile, did not observe a significant improvement in purification. The yield of the target protein is acceptable, but still unable to remove any of the contaminants present from the Ni-NTA purification. To address this issue the protocol was enhanced as outlined in section 3.2.2. This will involve introducing additional wash steps with low imidazole concentration to effectively remove more of the contaminating proteins before proceeding with elution using a high imidazole concentration. Additionally, the induction process for protein expression was altered to

induce bacterial cell growth when the falls within the range of 0.6 to 0.8. This approach is aimed at accelerating cell growth, which, in turn, should enhance the potential production of the protein of interest.

3.2.4 Imidazole washes optimization:

Rv1273c and Rv1272c were both overexpressed in the BL21 (DE3) cell strain. The induction was carried out overnight, beginning when the OD_{600} reached a range between 0.6 and 0.8. Cells were harvested and membranes were prepared as describe in section 3.2.1. The resin underwent four washes with free imidazole buffer (B), each involving 5 ml of buffer, followed by a single wash with low imidazole concentrations (10 mM, 50 mM, 100 mM, pH 8 respectively) using 5 mL of solution. These washes effectively removed non-specifically bound contaminants, and we collected the respective fractions. The His-tagged protein was subsequently eluted using a higher imidazole concentration of 500 mM at pH 8. Eluted fractions were collected and analysed via SDS-PAGE and western blot (Figure 3.8a and 3.8b), alongside the flow-through and the initial three washes using buffer without imidazole. Notably, a significant amount of the solubilized material did not bind to the resin, as demonstrated by the flow-through fraction (FT). The free imidazole washes (W1-3) and the low imidazole concentration wash (data not shown) effectively removed any untagged proteins, and it is important to note that no detectable protein was removed from the column. Moreover, elution with a high concentration of imidazole (500 mM) in fractions labelled (E1-10) revealed four bands across the Coomassie gel. The first two bands appeared on top of each other, around 60 kDa, 45 kDa, and 35 kDa. The western blot detected the His-tagged proteins at the expected molecular weight of Rv1273c, appearing as clear bands around 60 kDa across the entire elution fractions, corresponding to one of the two bands observed on top of each other on the SDS-PAGE (E1-10). These faint bands on the Coomassie gel may be due to the amount of protein present during purification. The western blot also showed faint bands below the 140 kDa mark in elution fractions (E1-10), likely indicating the presence of an oligomer including Rv1273c. Additionally, there was a

faint band around 35 kDa, corresponding to the one observed in the Coomassie gel at the same fractions, which could be a sign of protein degradation.



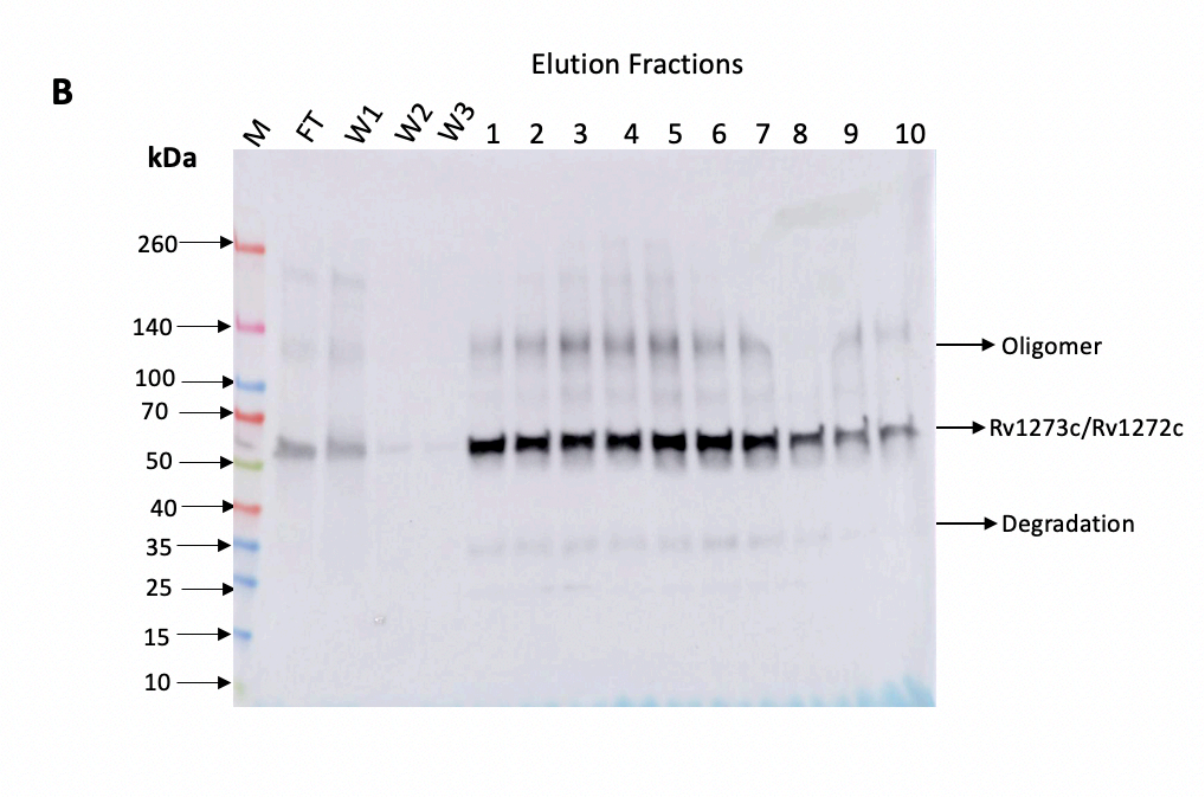
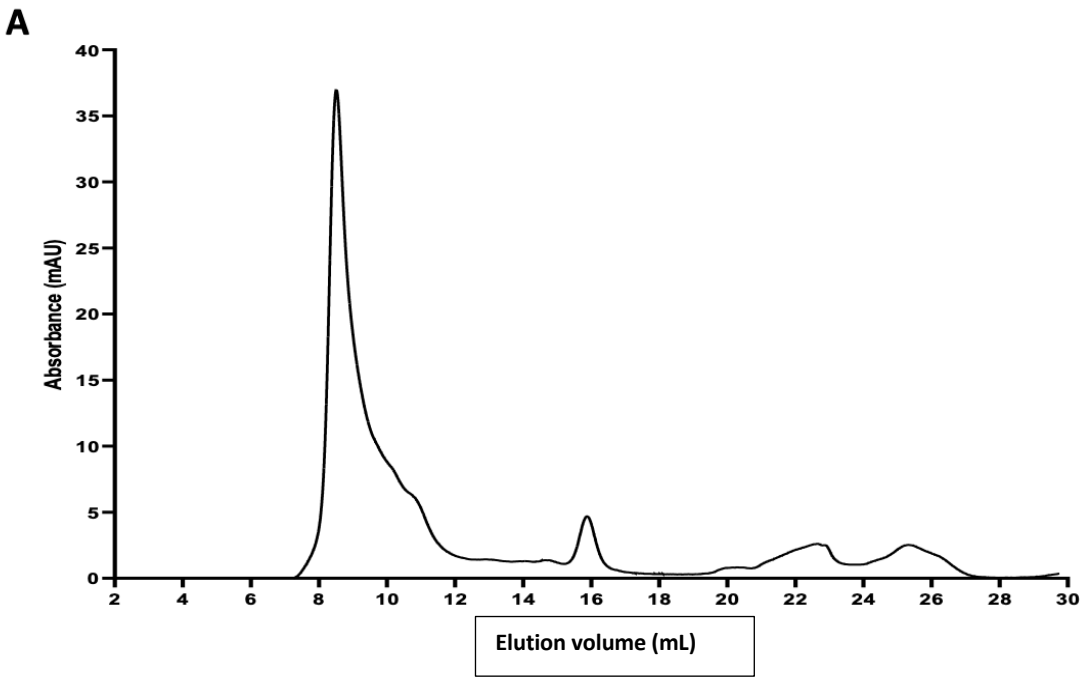


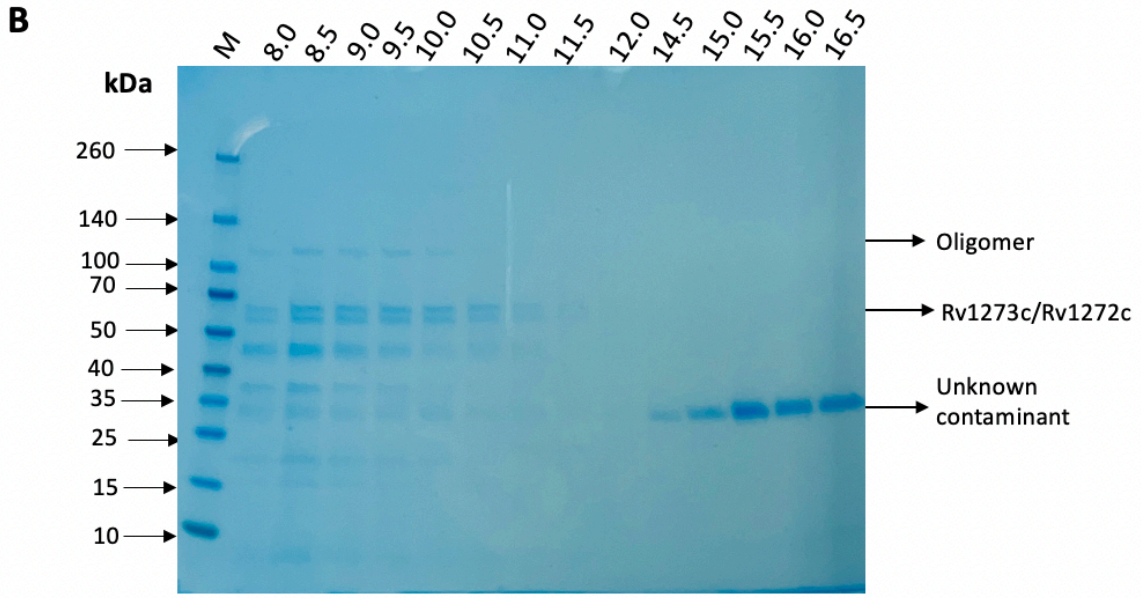
Figure 3.8: Rv1273c/ Rv1272c Ni-NTA Purification. A) SMA solubilized membrane were incubated with Ni-NTA overnight, then transferred to a gravity flow column, and flow through (FT) collected. Resin then washed 4 times with 5 ml solubilization buffer and collected 4 washes (shown the first three washes), then washed with 5 ml of 10 mM, 50 mM, and 100 mM imidazole pH 8 respectively (data not shown). Then His-Tagged proteins were then eluted with 10 ml of 500 mM imidazole buffer pH 8 (E1- E10). Fractions were analysed by electrophoresis on an 4-15 % precast SDS- PAGE and visualized by Coomassie stained. **B**) The SDS-PAGE was also transferred to a nitrocellulose membrane and probed with an anti-His-Tagged antibody**.**

Incorporating these adjustments into the protocol has resulted in a slight improvement in the Ni-NTA purification. However, there are still some contaminants presents. The western blot has shown that some of the bands, which could potentially be contaminants, are actually detected by the 6XHis-antibody. This suggests that they might be oligomers, including the target protein, or indicate protein degradation. To further enhance the purification process and achieve pure protein with a good yield, SEC was again applied to the samples from NTA containing the target.

3.2.4.1 Size Exclusion Chromatography (SEC)

The combined and concentrated the elution fractions from the NTA purification were subjected to separation using a Superdex 200 Increase 10/300 size exclusion column. Figure 9a illustrates the elution profile of the sample at 280 nm, highlighting the protein elution. Proteins eluted over a moderate range of elution volumes, with one major peak at around 8.5 mL, two shoulder peaks between 9 mL and 12 mL, and a smaller peak between 14.5 mL and 16.5 mL. Fractions within this profile underwent testing through SDS-PAGE and western blotting, as demonstrated in (Figure 3.9b and 3.9c). Both the Coomassie-stained gel and the western blot confirmed that Rv1273c and Rv1272c were eluted between 8.0 mL and 11 mL, detected by the presence of the His-Tag associated with Rv1273c during elution. The Coomassie-stained gel displayed the protein fractions along with some contaminating proteins, which appeared to be consistent with contaminants observed throughout the purification process. However, these contaminating proteins disappeared between the 10 mL and 11 mL fraction, leaving only the band representing the correct molecular weight of Rv1273c and Rv1272c. The western blot analysis detected a band for Rv1273c, which was observed at approximately 60 kDa, close to the known molecular weight of Rv1273c at 62.11 kDa. Additionally, there was a higher band observed at around 120 kDa, likely indicating the presence of an oligomer that includes Rv1273c. Furthermore, the Coomassie-stained gel exhibited a corresponding band at the correct molecular weight for Rv1273c. Given the similarity in molecular weights between the two proteins, Rv1273c and Rv1272c, it's possible that they do not fully separate on the Coomassie gel. Interestingly, the Coomassie stain gel showed a pure protein fraction between 14.5 mL and 16.5 mL that was not detected on the western blot.


A
Absorbance (mAU)
0
5
10
15
20
25
30
35
40
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
Elution volume (mL)


B
M
8.0
8.5
9.0
9.5
10.0
10.5
11.0
11.5
12.0
14.5
15.0
15.5
16.0
16.5
kDa
260
140
100
70
50
40
35
25
15
10
Oligomer
Rv1273c/Rv1272c
Unknown
contaminant

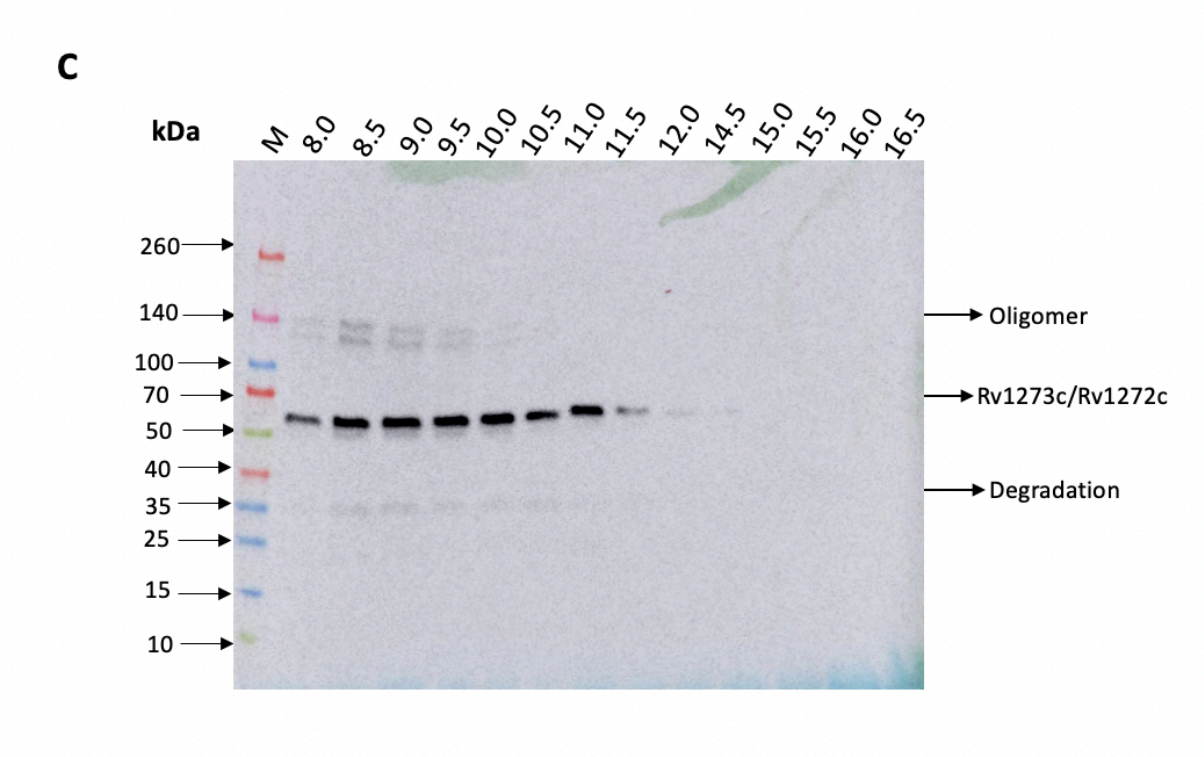


Figure 3.9: SEC purification of Rv1273c/ Rv1272c Elution fraction fractions from Ni- NTA purification were collected, concentrated, and injected onto a Superdex 200 Increase 10/300 GL column. The column was run at 0.5 mL/min collecting 0.5 mL fractions. **A)** SEC elution profile. Elution of protein from the column was monitored by absorbance at 280nm. **B)** SDS-PAGE of elution profile 4 -15 % Precast SDS-PAGE of the SEC elution fractions analysed by electrophoresis and protein bands visualized by Coomassie staining**. C)** The SDS-PAGE was also transferred to a nitrocellulose membrane and probed with an anti-His-Tagged antibody**.**

The size exclusion chromatography profile displays one major peak and two shoulder peaks with low absorbance. When analysing the fractions within this profile, there are positive indicators of target protein purity, although this has a slight impact on protein yield, especially noticeable in the 10.0 mL – 11.5 mL fractions. However, it also elutes an unknown protein in the late fraction with high-intensity bands in the Coomassie-stained gel. To overcome these challenges, several modifications to the protocol outlined in section 3.2.4 were made. First, the pH of all buffers containing imidazole change to 7.5 to enhance protein stability and maintain structural integrity. Second, the elution buffer was modified to ensure complete elution of the target protein without any residue left on the column.

Additionally, low imidazole concentration washes were used to remove as many contaminant proteins as possible. Finally, if the Ni-NTA purification is successful, the molecular weight cut-off of the protein concentrator changes to 10,000 in order to improve protein yield.

3.2.5 Changing buffer pH and increasing the elution volume:

3.2.5.1 Revised protocol 4: 5 mL 10 mM imidazole, 50 mM imidazole, 100 mM imidazole, and 20 mL 500 mM elution buffer pH 7.5:

Rv1273c and Rv1272c were both overexpressed in the BL21 (DE3) cell strain, solubilized, and bound to the resin using the same protocol described in Section 3.2.4. Following an overnight incubation with Ni-NTA resin, the resin was washed four times with a free imidazole buffer (B) pH 8, each time using 5 ml. Followed by a single wash with low imidazole concentrations (10 mM, 50 mM, 100 mM, pH 7.5, respectively) using 5 mL of solution. These washes effectively removed non-specifically bound contaminants, and respective fractions were collected. The His-tagged protein was subsequently eluted using a higher imidazole concentration of 500 mM at pH 7.5. Eluted fractions were collected and analysed via SDS-PAGE and western blot (Figure 3.10a and 3.10b), alongside the membrane (Mb), solubilized membrane (Sm), flow-through (FT), and the first wash of buffer-free imidazole. Notably, a significant amount of the solubilized material did not bind to the resin, as demonstrated by the flow-through fraction (FT). The free imidazole washes (W1) and the low imidazole concentration wash (not shown) effectively removed any untagged proteins, and it's important to note that no detectable protein was removed from the column. However, elution with a high concentration of imidazole (500 mM) in fractions labelled (E1-10) did not show an improved yield of the proteins of interest. The western blot detected His-tagged proteins at the correct molecular weight in the membrane, solubilized membrane, flow-through, the first wash of buffer-free imidazole, and the first five elution fractions.

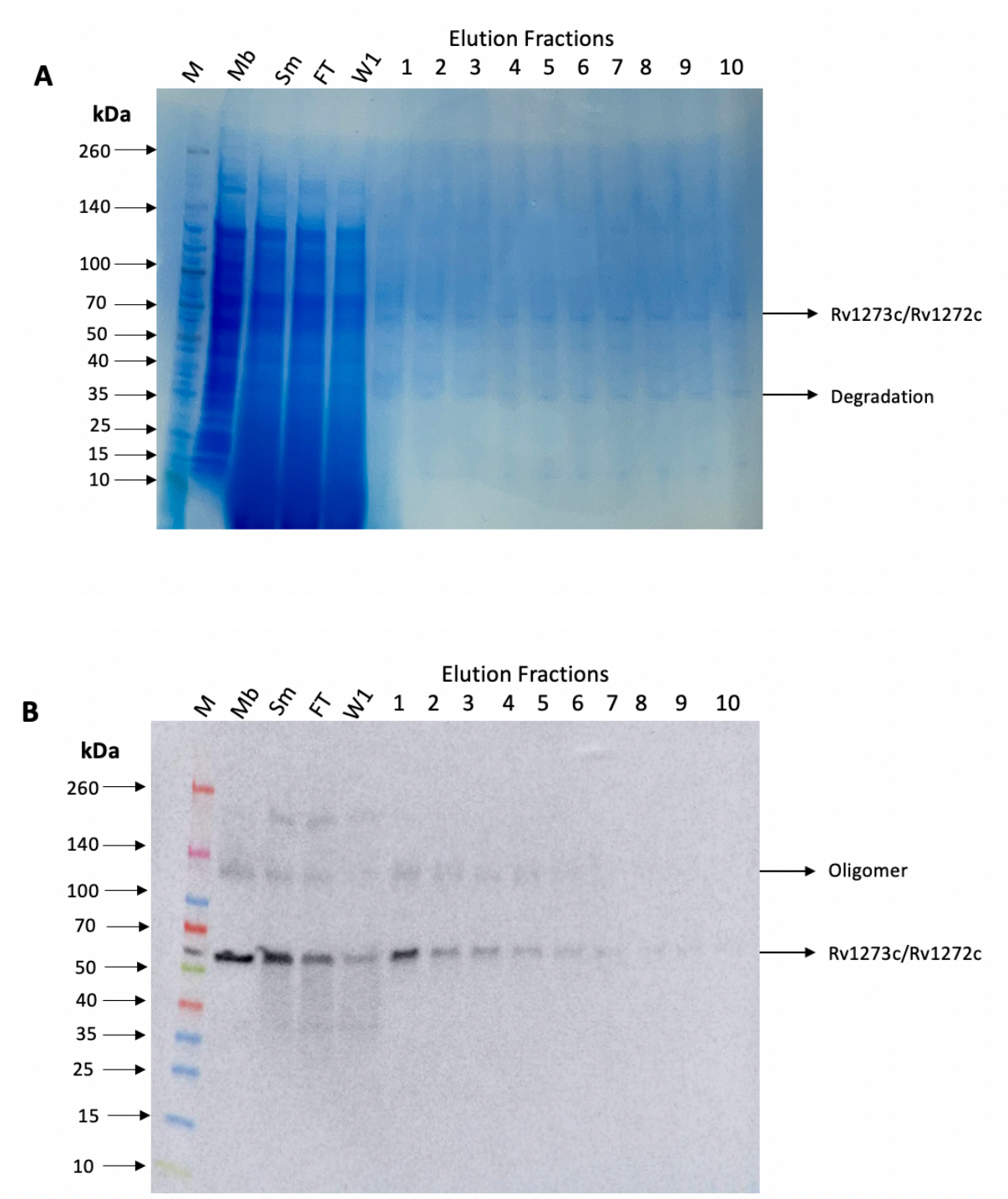


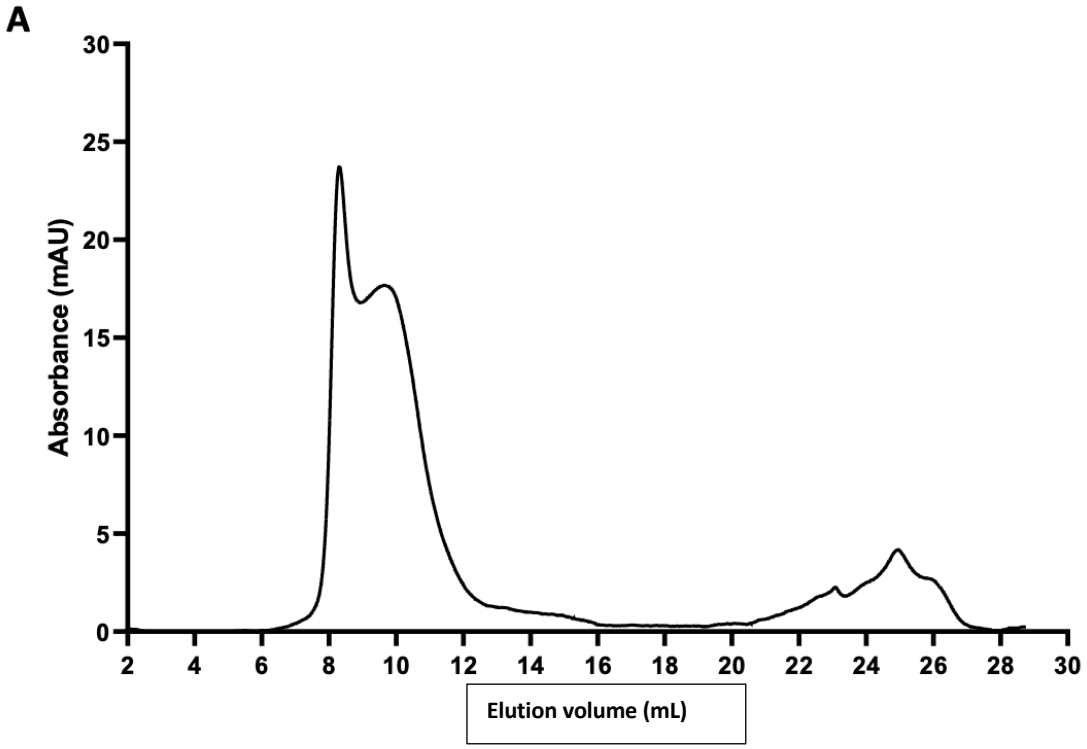
Figure 3.10: Rv1273c/ Rv1272c Ni-NTA Purification. A) SMA solubilized membrane were incubated with Ni-NTA overnight, then transferred to a gravity flow column, and flow through (FT) collected. Resin then washed 4 times with 5 ml solubilization buffer and collected 4 washes (shown the first three washes), then washed

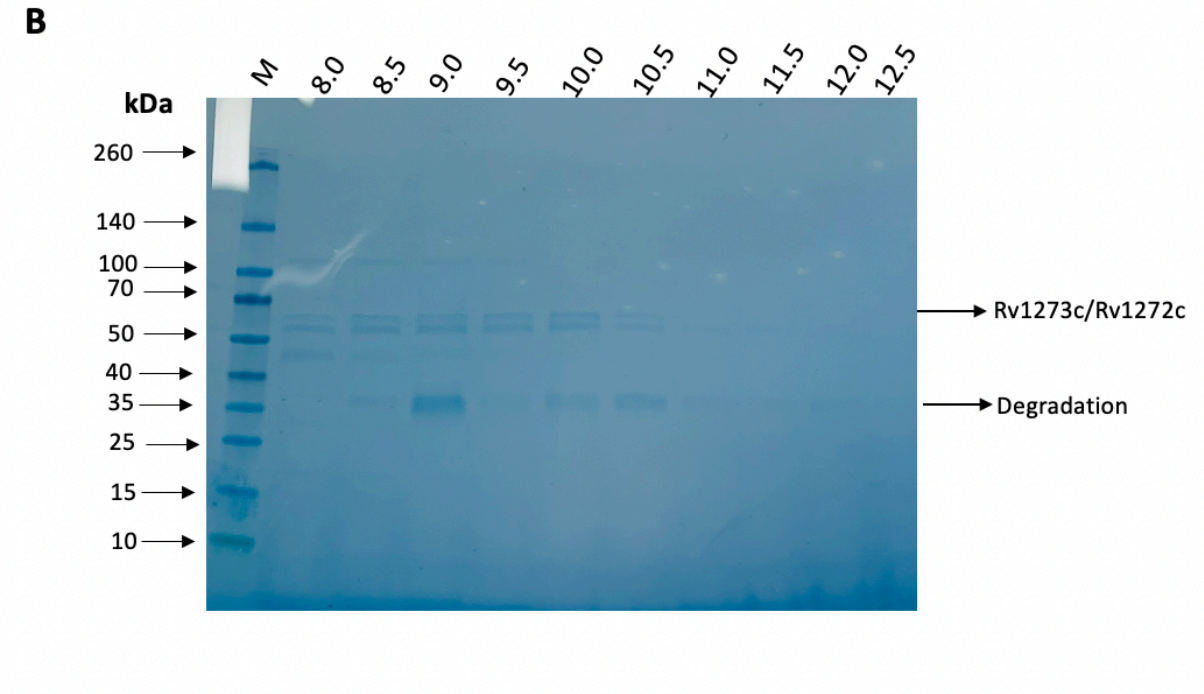
with 5 ml of 10 mM, 50 mM, and 100 mM imidazole pH 7.5 respectively (data not shown). Then His-Tagged proteins were then eluted with 20 ml of 500 mM imidazole buffer pH 7.5 (E1- E10). Fractions were analysed by electrophoresis on 4 -15 % precast SDS- PAGE and visualized by Coomassie stained. B) The SDS-PAGE was also transferred to a nitrocellulose membrane and probed with an anti-His-Tagged antibody.

Unfortunately, the Ni-NTA purification did not show significant improvement; the yield of the target protein remained low, and the purity did not increase notably. It appears that the volume of the low imidazole concentration wash was insufficient to adequately remove contaminants from the column. Despite the low yield the material was applied to size exclusion chromatography. However, it is important to keep in mind that this process may result in some loss of protein, further affecting the yield.

3.2.5.1.1 Size Exclusion Chromatography (SEC)

In the next step of purification, elution fractions were pooled and concentrated using a 10,000 MWCO concentrator and then subjected to separation using a Superdex 200 Increase 10/300 size exclusion column. Figure 3.11a depicts the elution profile of the sample at 280 nm, highlighting protein elution. Proteins eluted over a moderate range of elution volumes, with two major peaks observed between 8 mL and 12.5 mL. Fractions within this profile were subjected to testing via SDS-PAGE and western blotting, as demonstrated in (Figure 3.11b and 3.11c). Both the Coomassie-stained gel and the western blot confirmed that Rv1273c and Rv1272c were eluted between 8.0 mL and 10.5 mL, detected by the presence of the His-Tag associated with Rv1273c during elution. The Coomassie-stained gel showed a slight improvement in purification, with a reduction in the presence of contaminants compared to previous purifications. Moreover, a faint band around the 35 kDa mark was observed, which might be indicative of protein degradation. The western blot analysis revealed a band for Rv1273c, which was observed at approximately 60 kDa within the 8 mL to 10.5 mL range. Importantly, there were no signs of protein oligomerization or degradation in this range.


A
Absorbance (mAU)
30
25
20
15
10
5
0
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
Elution volume (mL)


B
kDa
M
8.0
8.5
9.0
9.5
10.0
10.5
11.0
11.5
12.0
12.5
260
140
100
70
50
40
35
25
15
10
Rv1273c/Rv1272c
Degradation

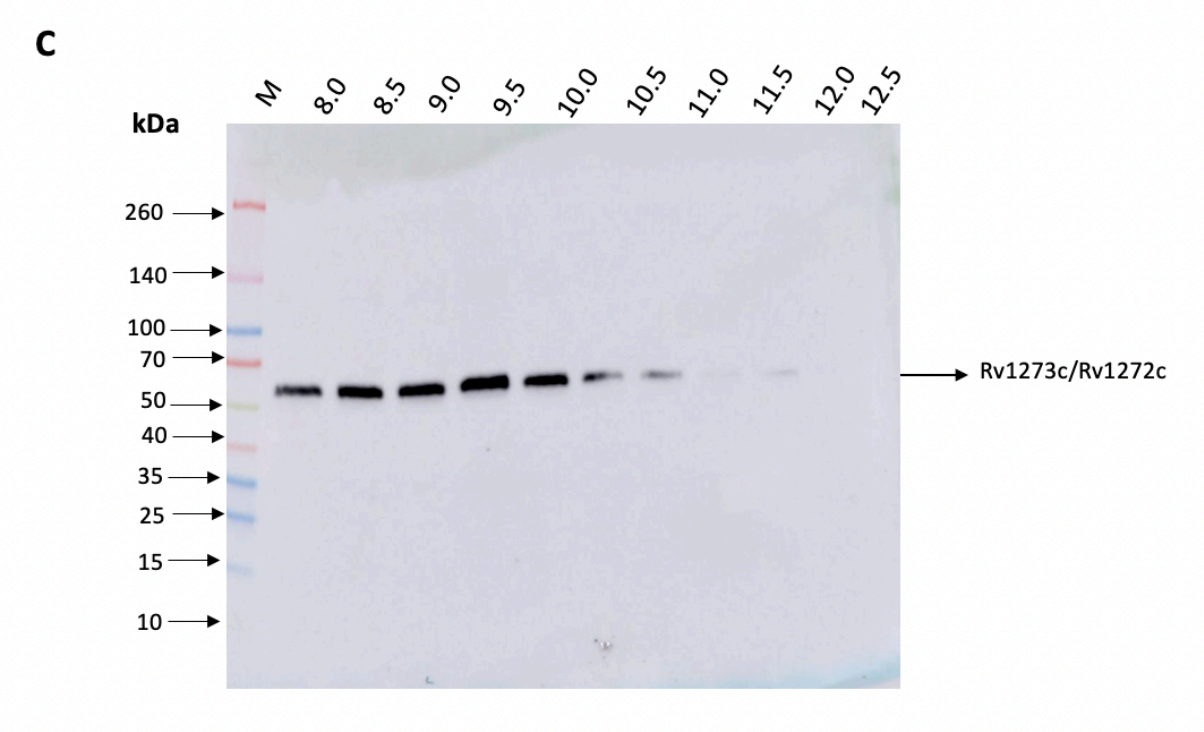
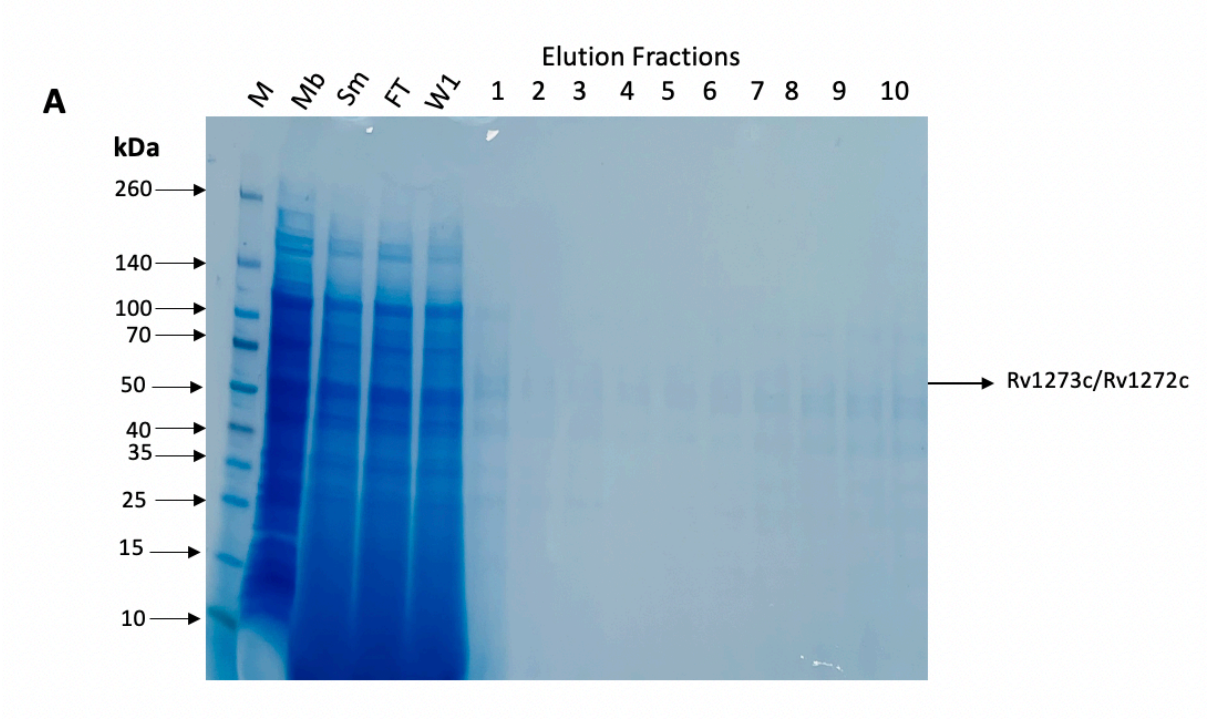


Figure 3.11: SEC purification of Rv1273c/ Rv1272c Elution fraction fractions from Ni- NTA purification were collected, concentrated, and injected onto a Superdex 200 Increase 10/300 GL column. The column was run at 0.5 mL/min collecting 0.5 mL fractions. **A)** SEC elution profile. Elution of protein from the column was monitored by absorbance at 280 nm. **B)** SDS-PAGE of elution profile 4 -15 % Precast SDS-PAGE of the SEC elution fractions analysed by electrophoresis and protein bands visualized by Coomassie staining**. C)** The SDS-PAGE was also transferred to a nitrocellulose membrane and probed with an anti-His-Tagged antibody**.**

The size exclusion chromatography profile displays two peaks with low absorbance. As expected, the yield was low considering the actual protein eluted during the Ni-NTA purification. There was a slight improvement in the purity of the target protein, as indicated by the presence of almost a single band in the 9.5 mL – 10.5 mL fractions. While the overall protein purity was acceptable, the yield fell below expectations. To enhance both aspects, I the volume of the low imidazole concentration washes was increased with the aims to remove as many contaminants as possible and subsequently improve the yield of the target protein.

3.2.5.2 Revised protocol 5: 10 mL 10 mM imidazole, 50 mM imidazole, 100 mM imidazole, and 20 mL 500 mM elution buffer pH 7.5:

Following an overnight incubation of the SMALP solubilised target with Ni-NTA resin, the resin was washed four times with a free imidazole buffer (B) pH 8, each time using 5 ml. Followed by a single wash with low imidazole concentrations (10 mM, 50 mM, 100 mM, pH 7.5, respectively) using 10 mL of solution. These washes effectively removed non-specifically bound contaminants, and fractions were collected. The His-tagged protein was subsequently eluted using a higher imidazole concentration of 500 mM at pH 7.5. Eluted fractions were collected and analysed via SDS-PAGE and western blot (refer to Figure 12a and 12b), alongside the membrane (Mb), solubilized membrane (Sm), flow-through (FT), and the first wash of buffer-free imidazole. It is worth noting that a significant amount of the solubilized material did not bind to the resin, as demonstrated by the flow-through fraction (FT). The free imidazole washes (W1) and the low imidazole concentration wash (not shown) effectively washed away any untagged proteins, and no detectable protein was removed from the column. However, elution with a high concentration of imidazole (500 mM) in fractions labelled (E1-10) did not show any appreciable improvement in the yield of the proteins of interest. Nevertheless, it indicated some protein loss during the washing steps with a large volume of different low concentrations of imidazole compared to the purification in (section 3.2.5.1). The western blot detected His-tagged proteins at the correct molecular weight in the membrane, solubilized membrane, flow-through, the first wash of buffer-free imidazole, and across the entire elution fractions, although these bands appeared faint.



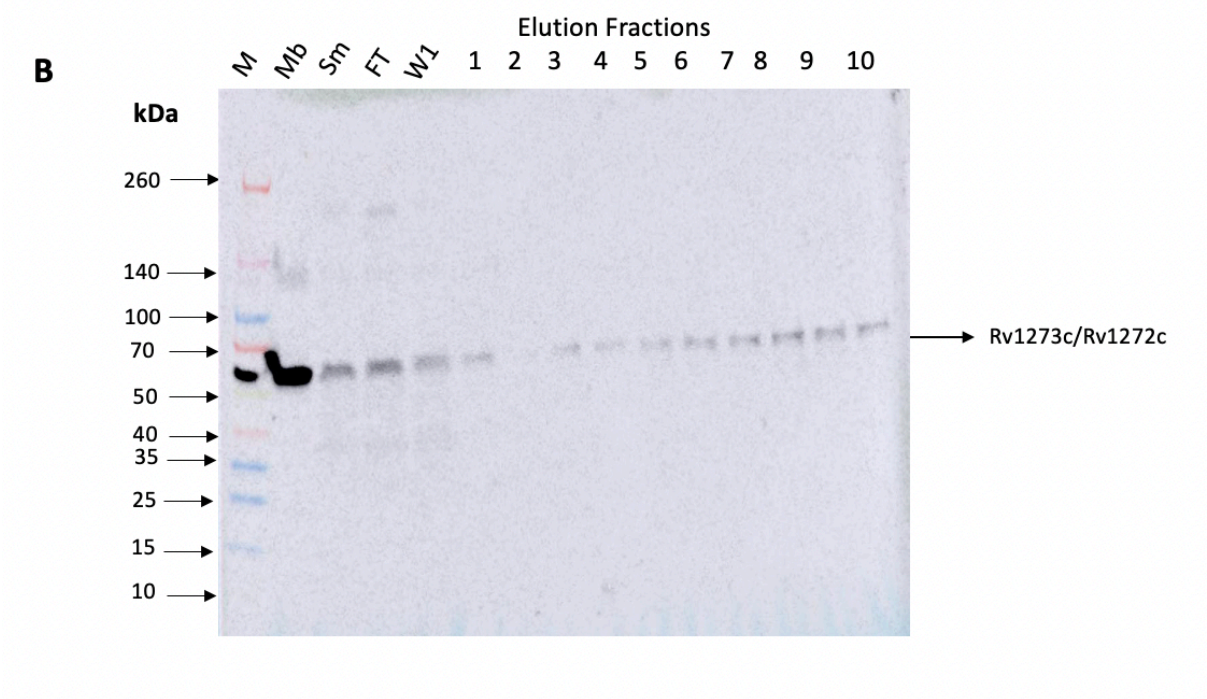


Figure 3.12: **Rv1273c/ Rv1272c Ni-NTA Purification. A)** SMA solubilized membrane were incubated with Ni-NTA overnight, then transferred to a gravity flow column, and flow through (FT) collected. Resin then washed 4 times with 5 ml solubilization buffer and collected 4 washes (shown the first wash), then washed with 10 ml of 10 mM, 50 mM, and 100 mM imidazole pH 7.5 respectively (data not shown). Then His-Tagged proteins

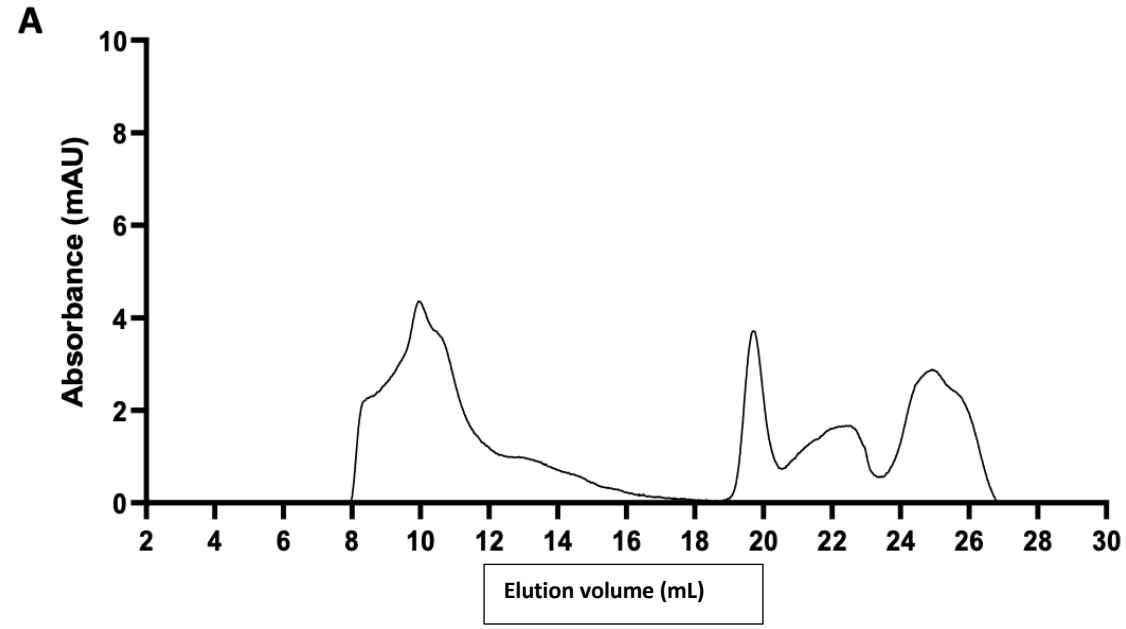
were then eluted with 20 ml of 500 mM imidazole buffer pH 7.5 (E1- E10). Fractions were analysed by electrophoresis on 4 -20 % precast SDS- PAGE and visualized by Coomassie stained. B) The SDS-PAGE was also transferred to a nitrocellulose membrane and probed with an anti-His-Tagged antibody**.**

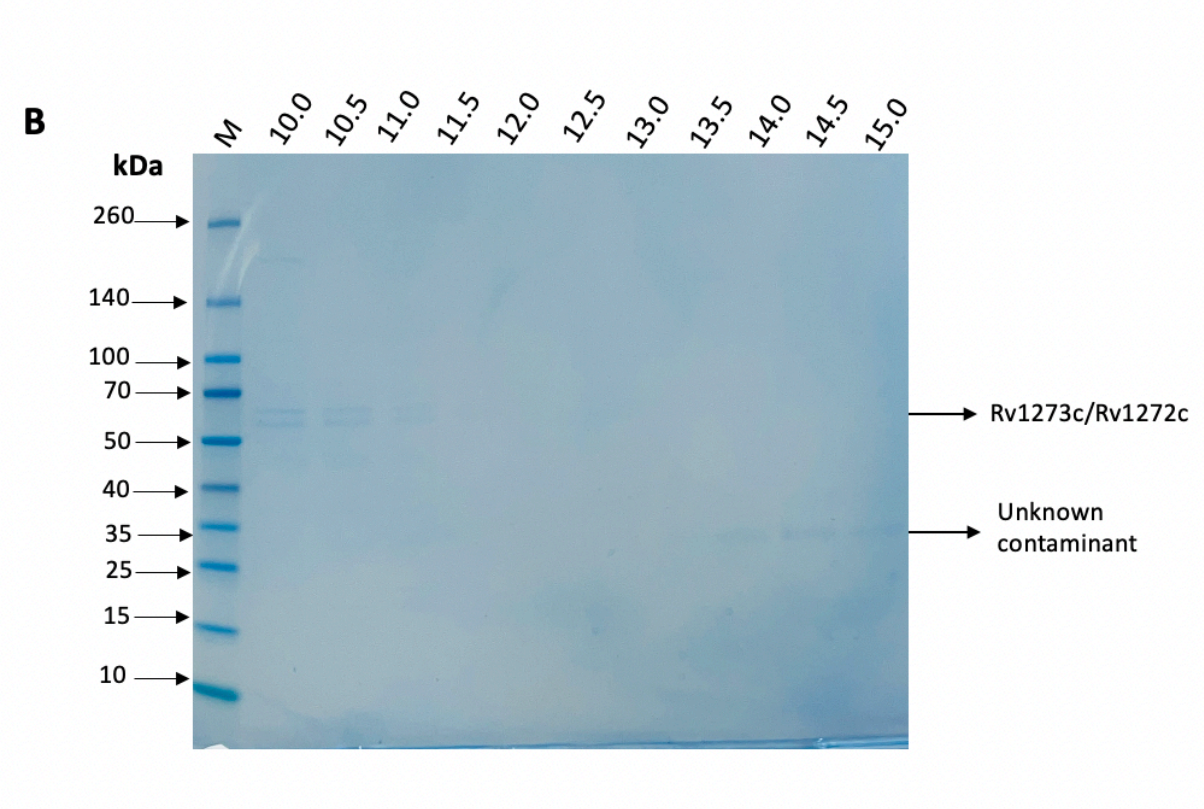
Despite the adjustments made at this stage of Ni-NTA purification, the outcome was unexpected. Unable to recover some of the target protein as originally planned. Protein yield decreased significantly, and the purity is questionable. It seems that increasing the volume of the low imidazole concentration wash likely removed both the target protein and contaminants. The western blot did detect some target protein at the correct molecular weight in the elution fractions.

3.2.5.2.1 Size Exclusion Chromatography (SEC)

Based on the bands observed on the western blot, it was decided to proceed to the next step of purification in order to improve the yield. The elution fractions were combined and concentrated using a 10,000 MWCO concentrator. The sample was then subjected to separation using a Superdex 200 Increase 10/300 size exclusion column, as shown in (Figure 3.13a), which represents the elution profile of the sample at 280 nm, indicating protein elution. Proteins were eluted over a moderate range of elution volumes, with approximately three small peaks observed between 8 mL and 12.5 mL, characterized by low absorbances. Fractions from 10 mL to 15 mL were subjected to testing via SDS-PAGE and western blotting, as demonstrated in (Figure 3.13b and 3.13c). Unfortunately, the Coomassie-stained gel showed faint bands around 10 mL to 11 mL, and these bands were also detected on the western blot, indicating that a small amount of Rv1273c and Rv1272c were eluted between 10.0 mL and 11.5 mL, as detected by the His-Tag of Rv1273c. However, the Coomassie-stained gel did not show any significant improvement in purification. Additionally, faint bands were observed on the gel between 14.0 mL and 15.0 mL, which were the same bands that appeared in a previous purification step mentioned in (Section 3.2.4.1.) and were not detected by the western blot. The western blot analysis did confirm the presence of a band for Rv1273c, which was observed around 60 kDa between 8 mL and 10.5 mL, with no signs

of protein oligomerization or degradation, despite the low absorbance observed on the size exclusion chromatography (SEC) profile.





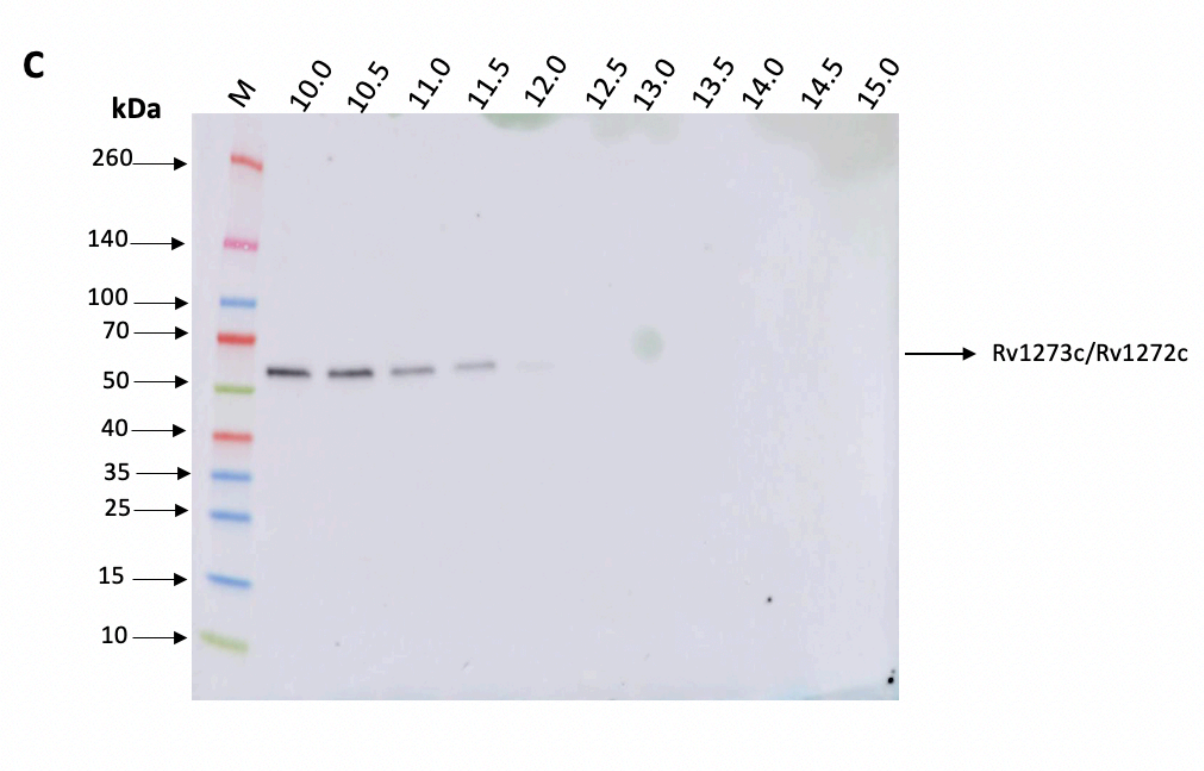
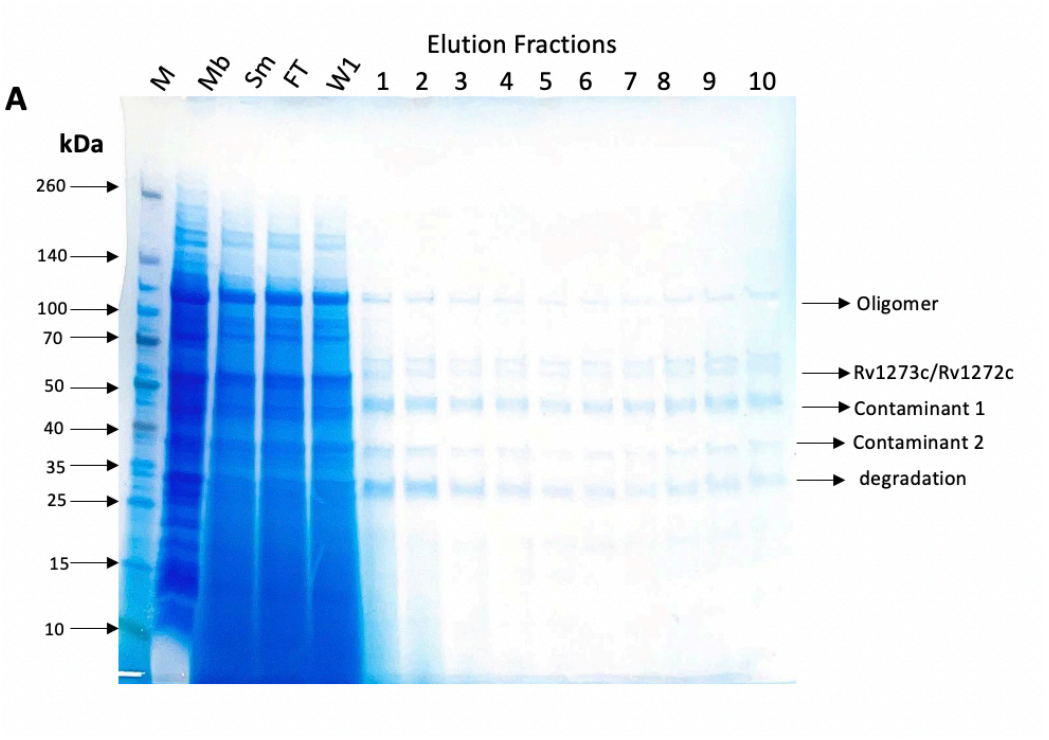


Figure 3.13: SEC purification of Rv1273c/ Rv1272c Elution fraction fractions from Ni- NTA purification were collected, concentrated, and injected onto a Superdex 200 Increase 10/300 GL column. The column was run at 0.5 mL/min collecting 0.5 mL fractions. **A)** SEC elution profile. Elution of protein from the column was monitored by absorbance at 280 nm. **B)** SDS-PAGE of elution profile 4 -20 % Precast SDS-PAGE of the SEC elution fractions analysed by electrophoresis and protein bands visualized by Coomassie staining**. C)** The SDS-PAGE was also transferred to a nitrocellulose membrane and probed with an anti-His-Tagged antibody**.**

As expected, the size exclusion chromatography profile displayed peaks with very low absorbance. Analysing the fractions within the profile did not reveal any significant improvements. Overall, the adjustments made to the protocol were unsuccessful. To address this, the Ni-NTA chromatography protocol was modified by increasing the concentration of the first low imidazole wash to 20 mM and removing the 100 mM imidazole wash. This modification aims to minimize the loss of the target protein during this process, ultimately improving both protein purity and yield.

3.2.5.3 Successful protein purification:

Rv1273c and Rv1272c were both overexpressed in the BL21 (DE3) cell strain, solubilized, and bound to the resin using the same protocol described in Section 3.2.4. Following an overnight incubation with Ni-NTA resin, the resin was washed four times with a free imidazole buffer (B) pH 8, each time using 5 ml. Followed by a single wash with low imidazole concentration (20 mM, 50 mM, pH 7.5) using 10 mL to remove non-specifically bound contaminants, with fractions being collected at each step. The His-tagged protein was then eluted with a higher concentration of imidazole (500 mM, pH 7.5). Eluted fractions were collected and analysed via SDS-PAGE and western blot (as shown in Figure 3.14 a and b), in addition to the analysis of the membrane (Mb), solubilized membrane (Sm), flow-through (FT), and the first wash of buffer-free imidazole. It is evident that a significant amount of the solubilized material did not bind to the resin, as demonstrated by the presence of proteins in the flow-through fraction (FT). The free imidazole washes (W1) and low imidazole concentration washes (not shown) effectively removed untagged proteins, with no detectable protein removal observed during these steps. Upon elution with a high concentration of imidazole (500 mM, E1-10), four bands were observed across the entire Coomassie gel, representing eluted proteins around 120 kDa, 60 kDa, 50 kDa, and 35 kDa. The western blot confirmed the presence of His-tagged proteins at the correct molecular weight (60 kDa) across all elution fractions, corresponding to the bands seen around 60 kDa on the SDS-PAGE (E1-10). Moreover, the western blot revealed faint bands around 120 kDa, which are likely indicative of protein oligomerization, and another band around 35 kDa, suggesting protein degradation.



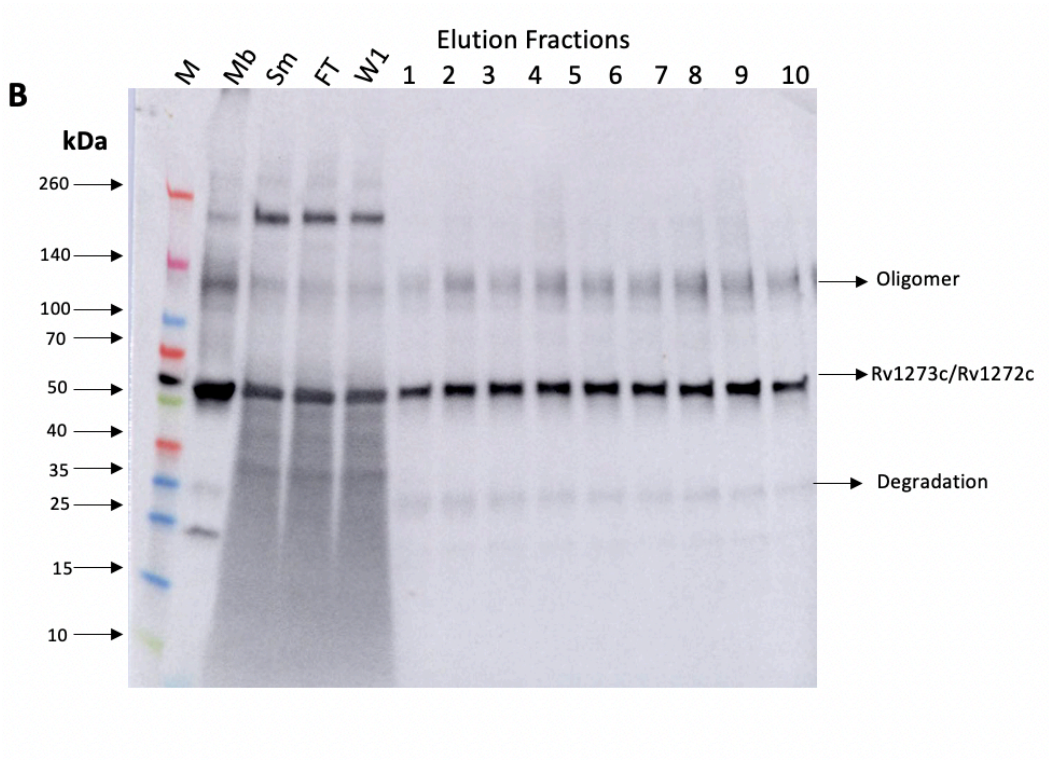


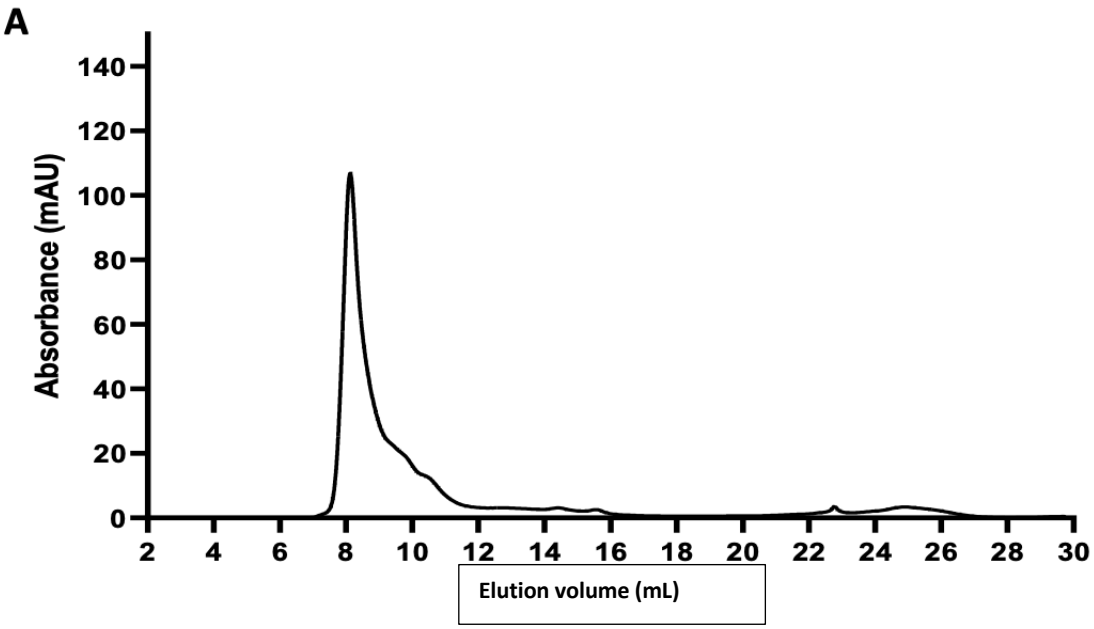
Figure 3.14: Rv1273c/ Rv1272c Ni-NTA Purification. A) SMA solubilized membrane were incubated with Ni-NTA overnight, then transferred to a gravity flow column, and flow through (FT) collected. Resin then washed 4 times with 5 ml solubilization buffer and collected 4 washes (shown the first three washes), then washed with 10 ml of 20 mM and 50 mM imidazole pH 7.5 respectively (data not shown). Then His-Tagged proteins

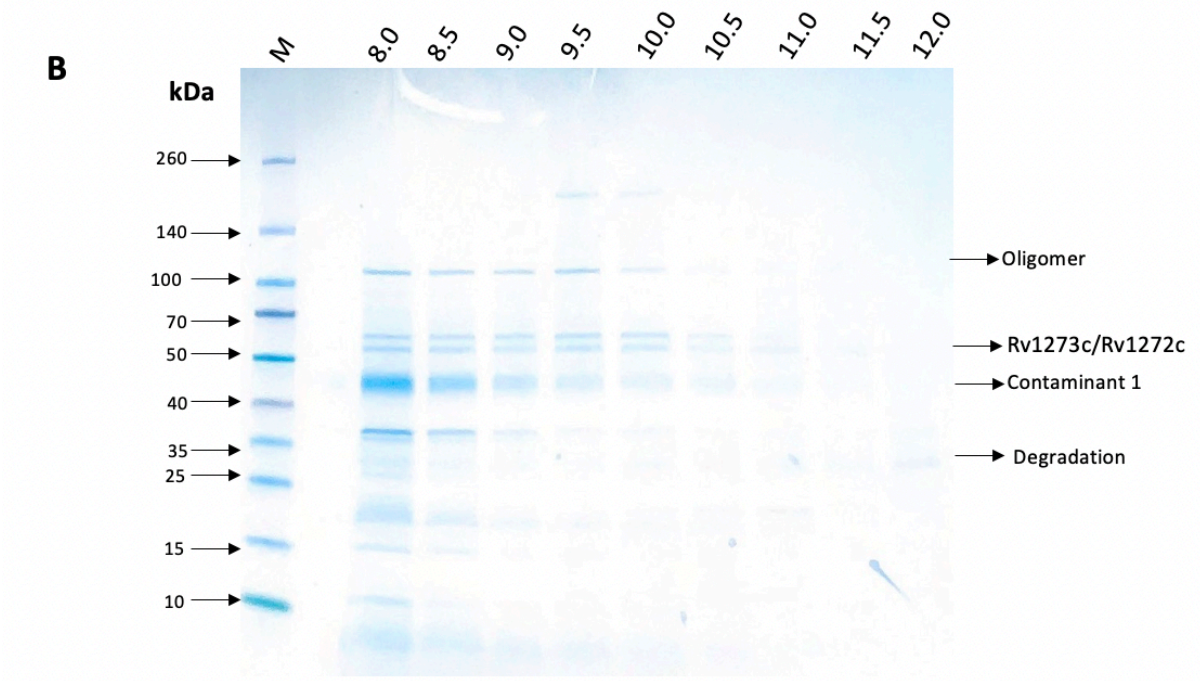
were then eluted with 20 ml of 500 mM imidazole buffer pH 7.5 (E1- E10). Fractions were analysed by electrophoresis on 4 - 20 % precast SDS- PAGE and visualized by Coomassie stained. B) The SDS-PAGE was also transferred to a nitrocellulose membrane and probed with an anti-His-Tagged antibody**.**

The Ni-NTA purification has shown significant improvement compared to the one described in section 3.2.5.2. At this stage, the target protein has been eluted alongside some contaminants, resulting in an increased protein yield, although it is only a slight improvement compared to the previous purification.

3.2.5.3.1 Size Exclusion Chromatography (SEC)

The elution fractions from the NTA chromatography were pooled and concentrated and subjected to separation using a Superdex 200 Increase 10/300 size exclusion column. In Figure 3.15a the elution profile of the sample at 280 nm can be seen, which indicates protein elution. The protein eluted within a medium range of elution volumes, with two major peaks occurring between 8 – 12 mL. Fractions collected across this elution profile were subjected to analysis via SDS-PAGE and western blotting, as depicted in (Figure 3.15 b and c). Coomassie stain gel and western blot analysis confirmed that Rv1273c and Rv1272c eluted between 8.0 mL and 11.5 mL, as detected by the presence of the His-Tag of Rv1273c in the elution. The Coomassie stain gel showed an improvement in the purification and protein yield. However, there was also a faint band around the 35 kDa mark, which might indicate protein degradation, as well as a higher band around 120 kDa. The western blot analysis revealed a band for Rv1273c, which was around 60 kDa and detected between 8.0 – 11.5 mL. Additionally, it detected two bands corresponding to those observed on the Coomassie stain gel, around 120 kDa and 35 kDa. These protein bands likely indicate protein oligomerization and degradation, respectively, as they have been consistently present in most of the protein purifications.


A
Absorbance (mAU)
140
120
100
80
60
40
20
0
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
Elution volume (mL)


B
M
8.0
8.5
9.0
9.5
10.0
10.5
11.0
11.5
12.0
kDa
260
140
100
70
50
40
35
25
15
10
Oligomer
Rv1273c/Rv1272c
Contaminant 1
Degradation

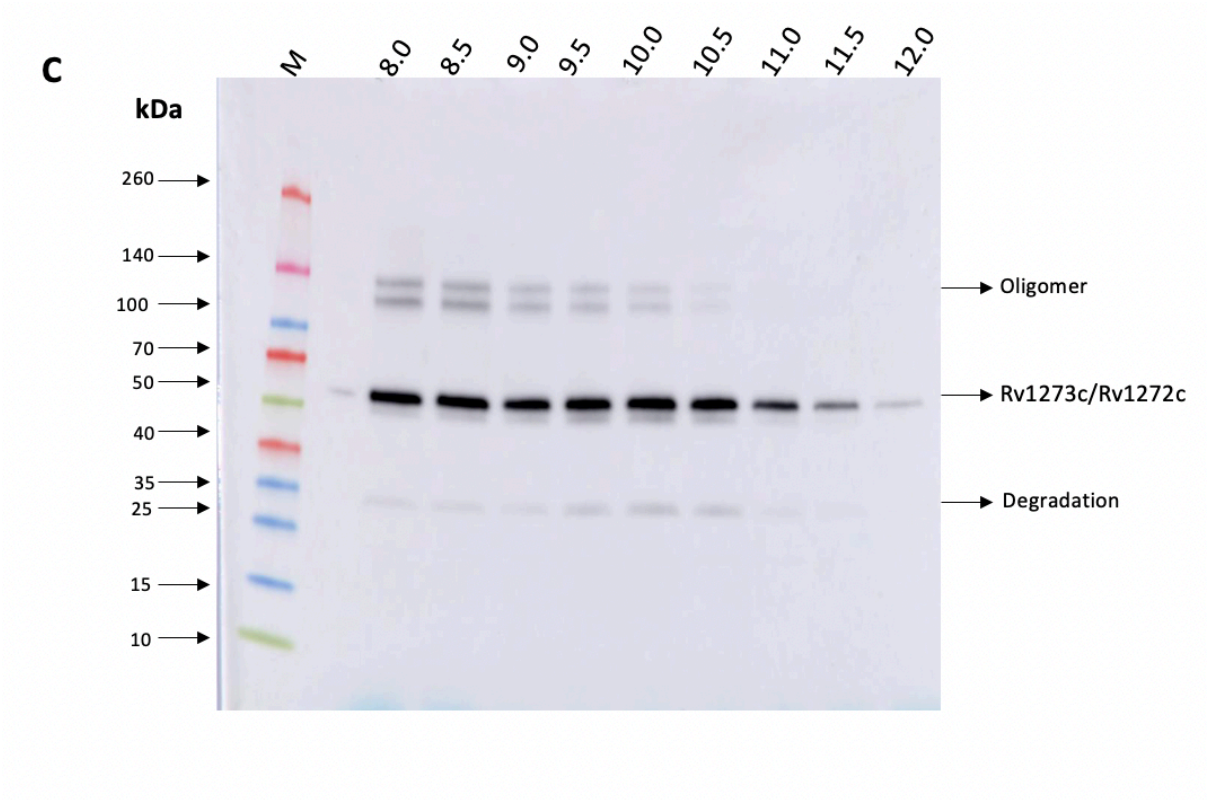


Figure 3.15: SEC purification of Rv1273c/ Rv1272c Elution fraction fractions from Ni- NTA purification were collected, concentrated, and injected onto a Superdex 200 Increase 10/300 GL column. The column was run at 0.5 mL/min collecting 0.5 mL fractions. **A)** SEC elution profile. Elution of protein from the column was monitored by absorbance at 280 nm. **B)** SDS-PAGE of elution profile 4 -20 % Precast SDS-PAGE of the SEC elution fractions analysed by electrophoresis and protein bands visualized by Coomassie staining**. C)** The SDS-PAGE was also transferred to a nitrocellulose membrane and probed with an anti-His-Tagged antibody**.**

The size exclusion chromatography profile displays one major peak alongside two shoulder peaks with high absorbance. This indicates an improvement in protein yield compared to the purification described in section 3.2.5.2.1. Coomassie stain gel and western blot have confirm this improvement in the yield of the target protein. While there was a weak separation of proteins in SEC, obtaining a high percentage of purity and a high yield of the target protein was challenging. Nevertheless, we were able to obtain protein with good quality and purity for further studies.

3.3 CHARACTERISATION OF THE COMPONENTS OF Rv1273c and Rv1272c IN STYRENE MALEIC ACID LIPID PARTICLES

3.3.1 Circular Dichroism of SMA solubilised Rv1273c and Rv1272c

Circular dichroism is a technique used to analyse how molecules absorb different types of circularly polarized light. It is particularly valuable in deciphering protein structures due to the distinctive patterns that emerge during specific transitions in the peptide backbone. For instance, observing the circular dichroism spectrum of an alpha-helix reveals two negative points of reduced absorption at 208 nm and 222 nm. Beta-sheets, however, present less straightforward results because their structures exhibit more variation. Typically, they exhibit negative absorption around 216 nm.

CD data was collected for the sample corresponding to fraction on figure 3.16. This fraction contained Rv1273c which meant that the CD data would provide information on the secondary structure content of that protein. The CD data shows a relatively flat negative minimum between 208 and 222 nm and including 216 nm. This indicates that the sample contains both α-helix (208, 222 nm) and β-sheet (216 nm). To further assess these data the spectrum was deconvolved using Betsel. The Betsel analysis predicted that the protein contains 17.6% Alpha, 23.8% Beta and 58.7% loop/unfolded. Figure 3.17 shows the CD data and the result of the fit from Betsel showing that there is a good agreement between the two.

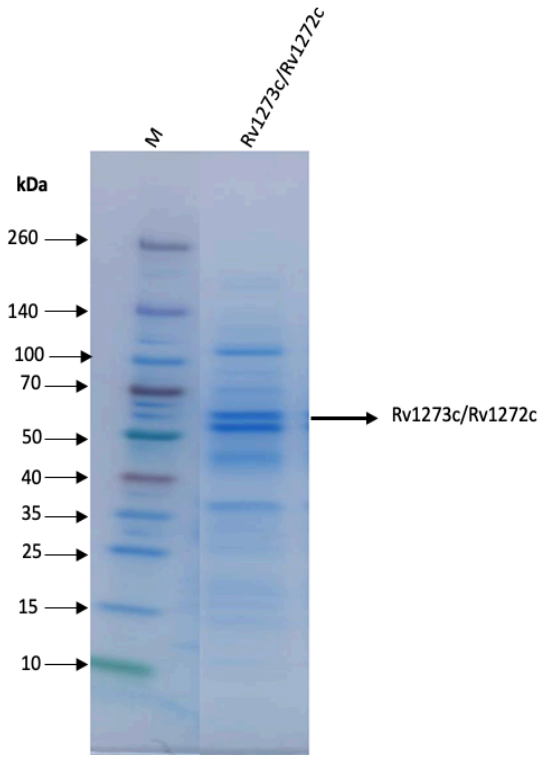


Figure 3.16: SDS-PAGE of Rv1273c and Rv1272c fraction after SEC purification

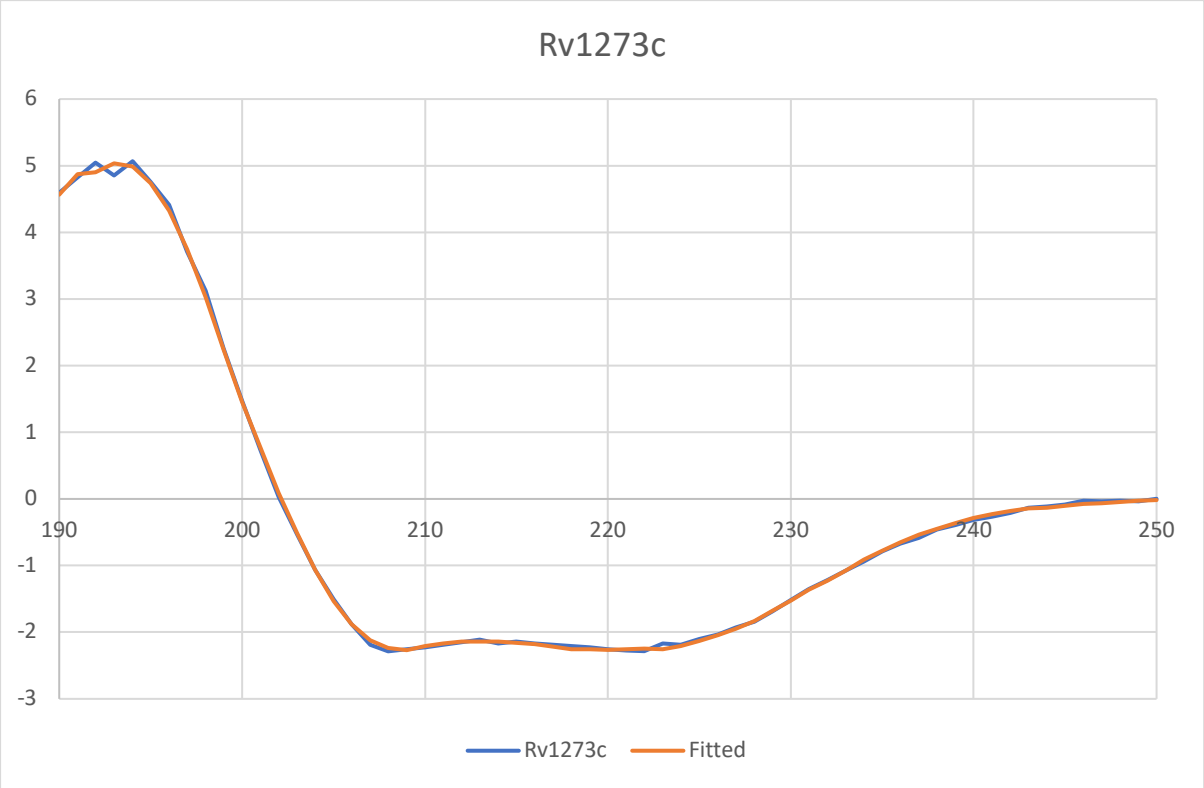


Figure 3.17: CD analysis of Rv1273c on SMALPs. The fractions obtained after the size-exclusion chromatography (SEC) purification of Rv1273c were examined using CD spectroscopy to evaluate its secondary

structure. The analysis was performed using a JASCO J-1500 Circular Dichroism Spectrophotometer operated through Spectra Manager software. Measurements were conducted in a quartz glass cuvette with a pathlength of 1 mm, using a protein concentration of approximately 0.1 mg/mL.

3.3.2 Analytical Ultracentrifugation (AUC)

A sample was taken from fraction on figure 3.18 and concentrated before being analysed using Sedimentation Velocity Analytical Ultracentrifugation (svAUC). svAUC provides a measure of the molecular mass of proteins in solution. These data from the svAUC run were analysed using SEDFIT (Schuck, 2000) and a plot of molecular weight verses concentration generated (see figure 3.19).

These data show one dominant peak at 143 KDa with a second lower peak at around 284 KDa and small peak around the 60 kDa. This first peak is consistent with either a dimer of either protein or a heterodimer containing one of each Rv1273c and Rv1272c. Given the data from the SDS PAGE of the sample which show two proteins might be present in significant quantities these data would be consistent with a heterodimer of Rv1273 and Rv1272c which would be 143 KDa. Also, a tetramer of either protein in 284 kDa.

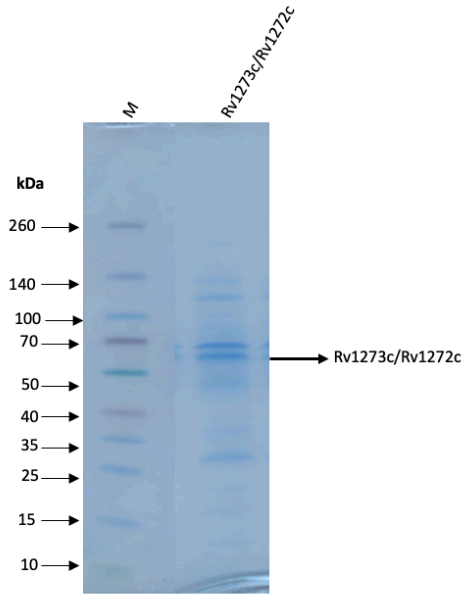


Figure 3.18: SDS-PAGE of Rv1273c and Rv1272c fraction after SEC purification

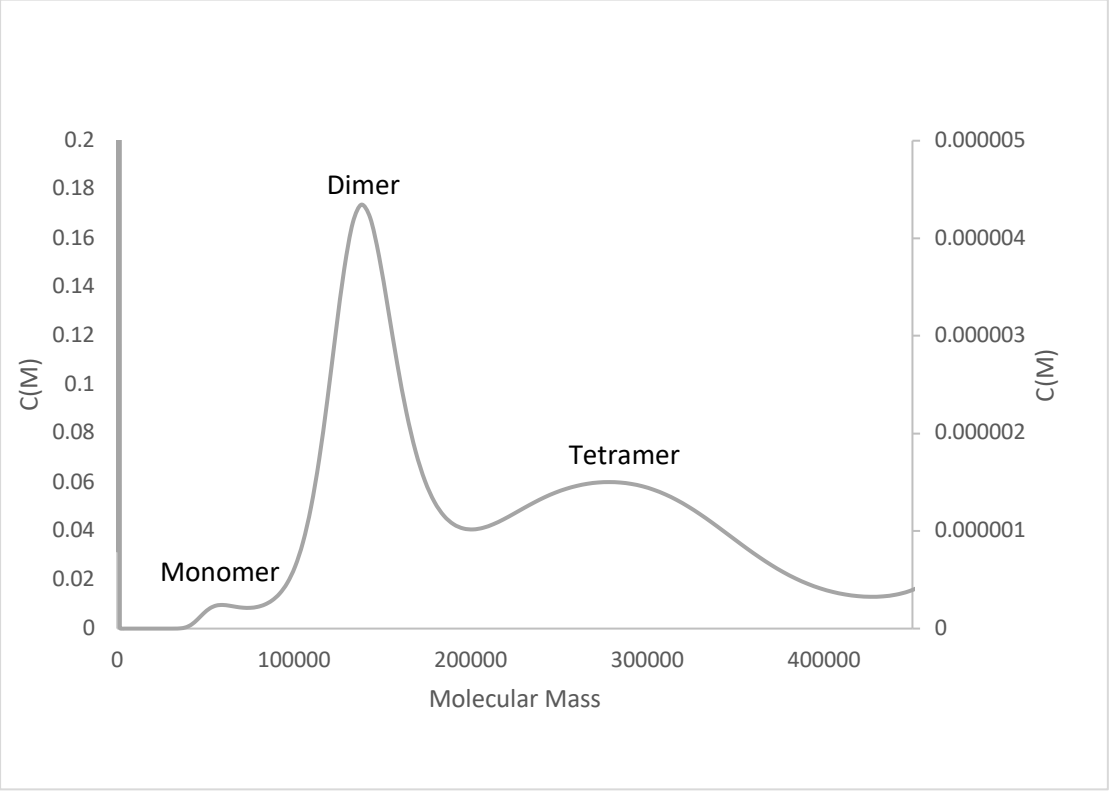


Figure 3.19: AUC Analysis of Rv1273c and Rv1272c SEC fraction

3.4 Discussion

The aim of this chapter was to investigate whether the SMALP method could solubilize the intact complex of *Mycobacterium tuberculosis* ABC-transporter Rv1273c and Rv1272c. Ni-NTA and size exclusion chromatography were used for purification, aiming for an ideally >80% protein purity benchmark. Along the way, protocols showed that SMA effectively solubilizes membrane proteins. Purifying the entire complex with only one tagged protein posed a challenge, particularly in ensuring that the untagged protein was not lost during the purification process. While achieving solubilization, there was some encountered hurdles, notably inefficient binding to the Ni-NTA, possibly due to the nature of SMALP. SMALP's 7-10 nm size might obstruct tag binding to the membrane. Exploring different resins with varying capacities could enhance binding and overall purification.

Additionally, considering additional tags like strep tags or Green Fluorescent Protein (GFP) tags at the opposite terminus might enhance specificity. Common protein contaminants co-eluting with tagged proteins from *E. coli* expression were challenging to remove due to their nickel affinity, creating competition during binding. Purifying the protein complex in *Mycobacterium smegmatis* could address this, offering insights into the complex and its surrounding lipids. Further, studying individual proteins apart from the complex could elucidate their functions and aid future structural determination.

Ensuring high-quality, high-yield protein was a major concern during purification. Low yield hindered robust structural characterization, limiting the application of higher-resolution methods like small angle X-ray scattering and electron microscopy. Increasing protein yield within SMALP is crucial for achieving higher-resolution data in the future. This chapter's data was derived from an eight-litre bacterial culture, suggesting scaling up culture volume as a potential solution. However, this could be time-consuming and challenging for many laboratories lacking the capacity to handle such large volumes of bacterial culture.

CHAPTER 4

SOLUBILIZATION OF *MYCOBACTERUIM TUBERCULOSIS* ABC-TRANSPORTER Rv3781c AND Rv3783c COMPLEX USING STYRENE MALEIC ACID

4.1. Introduction

Rv3781c and Rv3783c are membrane proteins categorized as ABC transporter exporters, potentially engaged in recycling and transporting membrane components and Lipopolysaccharides (Cassio Barreto de Oliveira and Balan, 2020). Rv3781c, identified as RfbE, exhibits nucleotide-binding domains (NBDs), while Rv3783c, also known as RfbD, encompasses membrane-spanning domains (MSDs) similar to those found in O-antigen and Lipopolysaccharides across various species (figure 4.1) (Braibant *et al*., 2000). Rv3781c displays distinctive motifs typical of ABC transporters, such as Walker A, ABC signature, Walker B, and the Linton and Higgins motif. In contrast, Rv3783c lacks the typical motifs characterizing membrane-spanning domains (Braibant *et al*., 2000). These proteins are speculated to play a role in Mycobacterium cell wall biosynthesis, particularly in arabinogalactan, and might also function in exporting lipid intermediates containing the galactan structure before its arabinosylation (Cuthbertson *et al*., 2010). Analysis of the sequence of the proteins using Mycobrowser shows that the peptide chain weights for each protein are 32.33 kDa for Rv3783c and 29.96 kDa for Rv3781c.

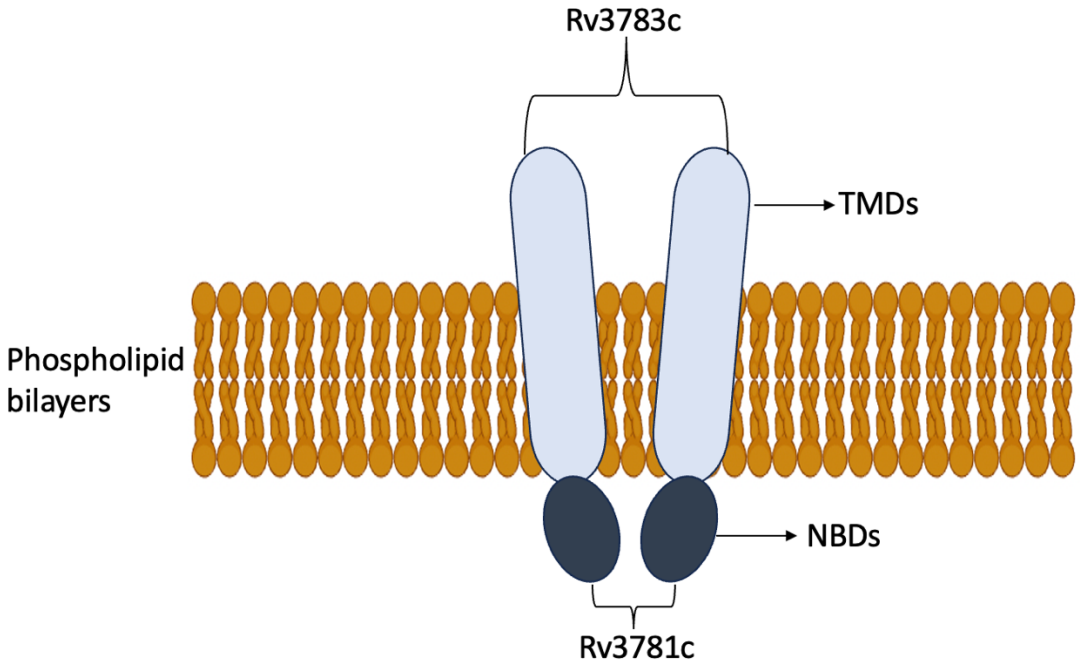


Figure 4.1: Schematic figure of ABC transporters showing the two proteins and the proposed functioning transporter Rv3781c/3783c.

The aim of this part of the project was to use SMALPs to extract intact complexes of Rv3781c and Rv3783c. Plasmids containing the two proteins were transformed into *E. coli* BL21 (DE3) and protein expression was induced with IPTG. Rv3781c and Rv3783c are proposed to form a hetero-tetramer which includes Rv3781c in the membrane and Rv3783c forming the extracellular domain). For these experiments on Rv3781c was tagged with a 12His tag meaning that success in this purification would be defined by extraction and purification of the full complex intact. If both (the tagged and untagged protein) are present in the final pure sample, then the SMALP contains the full complex. The experiment could have been set up with both proteins tagged, however in this case if both proteins appear in the final sample there is no evidence that they have formed a complex and instead might be present as individual proteins in solution.

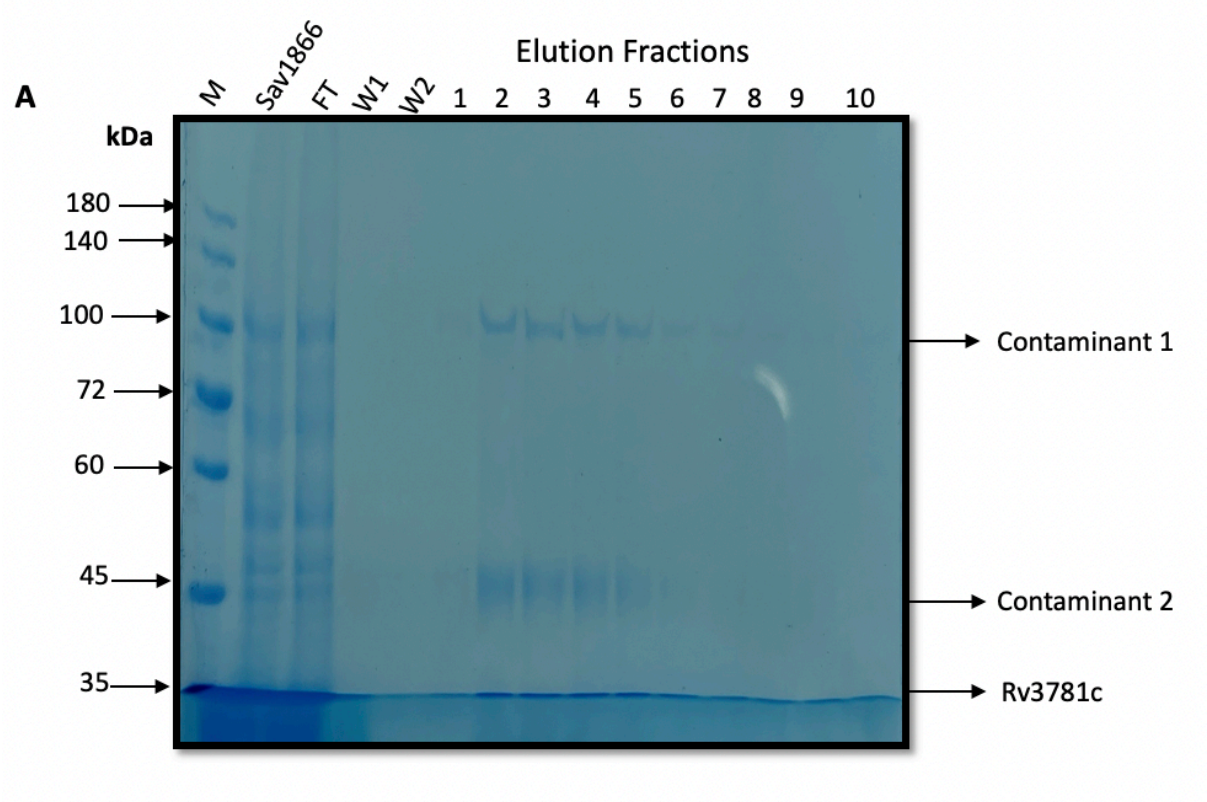
To achieve samples of high purity (Ideally >80%) for future studies a purification protocol had to be developed. Conventionally proteins in SMALPs are purified using a combination of Ni-NTA chromatography followed by SEC (Lee *et al*., 2016). In this chapter I show the development of a purification protocol that provide Rv3781c at high purity and show that maintaining the integrity of the complex between Rv3781c and Rv3783c is challenging using this system.

4.2 Solubilizing Rv3781c/Rv3783c using styrene malic acid:

4.2.1 Initial purification of Rv3781c and Rv3783c

Since these are a new protein complex and have not been purified using the SMALP system, the initial approach to protein purification followed the work of (Lee *et al*., 2016). This included a programme of washing NTA resin loaded with the protein of choice with free imidazole buffer (Buffer B), 20 mM imidazole wash, and finally elute with 500 mM imidazole. Rv3781c and Rv3783c were overexpressed in the BL21 (DE3) cell strain. Only Rv3781c was tagged with 12 histidine residues, and this overexpression was induced overnight when the OD_{600} reached approximately 0.6. Cells were harvested by centrifugation, and membranes were prepared. The membrane fraction was then resuspended and combined with SMA to reach a final concentration of 2.5%. This mixture was incubated at room temperature for 2 hours with gentle rocking. The insoluble fraction was then separated by ultracentrifugation. The SMA-solubilized membrane fraction was bound to prepared Ni-NTA overnight on a rotating platform at 4°C to purify Rv3781c and Rv3783c from other solubilized proteins. The resin was transferred using gravity flow column, and the flow-through was collected. The resin was washed four times with 5 ml each time using a buffer containing free imidazole (Buffer B). This was followed by a wash with a low imidazole concentration (20 mM) once, using 10 ml of buffer, to remove non-specifically bound contaminants, and the fractions were collected. The His-tagged proteins were eluted using a higher concentration of imidazole (500 mM) at pH 8. Eluted fractions were collected and analysed on SDS-PAGE (see Figure 4.2a), along with the flow-through and the first two washes of buffer containing free imidazole. It is evident that a significant amount of the solubilized material did not bind to the resin, as demonstrated by the presence of the target protein in the flow-through fraction (FT), The free imidazole washes (W1-2) and the low imidazole concentration wash (data not shown) effectively removed untagged proteins, confirming that no detectable protein was eliminated from the column during this procedure. In the elution fractions with a high concentration of imidazole (500 mM) (E1-10), there are protein bands on the SDS-PAGE with masses around 35 kDa, which closely match the expected mass of the protein of interest. However, two additional protein bands are observed which eluted in the same washes with masses of around 100 kDa and 45

kDa. Proteins on the SDS-PAGE were transferred onto a nitrocellulose membrane for western blot analysis (see Figure 4.2b). The western blot was probed with an anti-6XHis antibody and confirmed that the bands around 35 kDa (corresponding to those observed in the Coomassie gel) contain a histidine tag. Additionally, it reveals a band around 60 kDa (which also reacts with the anti 6XHis Antibody in the elution fractions (E2-E6) that was not visible in the Coomassie gel. Notably, the two bands around 100 kDa and 45 kDa were absent in the western blot analysis indicating that they do not contain a region that reacts with the 6XHis antibody.



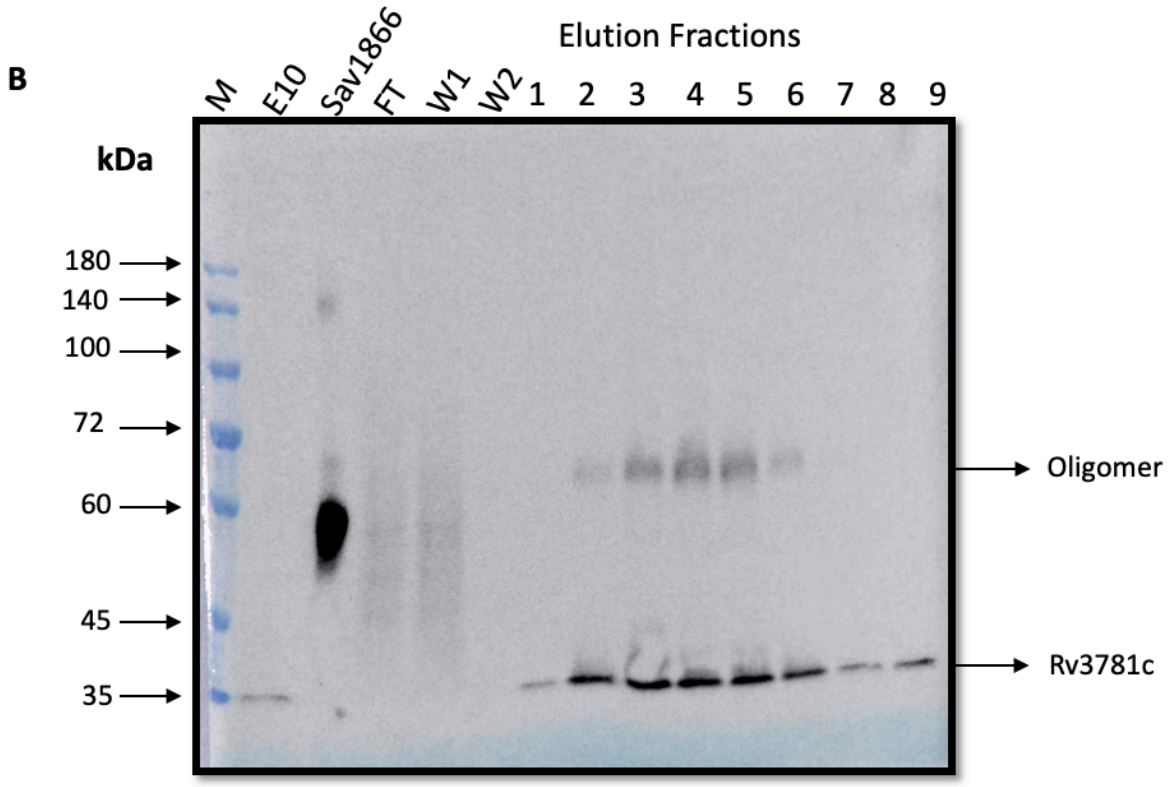


Figure 4.2: Rv3781c/ Rv3783c Ni-NTA Purification. A) SMA solubilized membrane were incubated with Ni-NTA overnight, then transferred to a gravity flow column, and flow through (FT) collected. Resin then washed 4 times with 5 ml solubilization buffer and collected 4 washes (shown the first two wash W1 and W2), then washed with 10ml of 20 mM imidazole (data not shown). Then His-Tagged proteins were then eluted with 10 ml of 500 mM imidazole buffer (1- 10). Fractions were analysed by electrophoresis on an 10 % handmade SDS-PAGE and visualized by Coomassie stained. B) The SDS-PAGE was also transferred to a nitrocellulose membrane and probed with an anti-His-Tagged antibody**.** 6XHis-Sav1866 (Sav1866), elution fraction 10 (E10).

During the initial purification, the results were promising indicating that SMA has the capability to solubilize these proteins. However, the protein yield and purity percentage were less than ideal at this stage, as evidenced by the presence of contaminating proteins eluted alongside the protein of interest. Notably, the protein of interest, Rv3781c, migrated very close to the dye front, making it challenging to visualize on a Coomassie gel stain. To address this issue, it was proposed that lower concentrations of imidazole should be used in washes to enable the recovery of more target protein.

4.2.2 Revised purification protocol 1

Rv3781c/Rv3783c were expressed and solubilized using the same protocol as described in Section 4.2.1. In this revised protocol the resin underwent four washes with buffer (B), each time using 5 ml. This was followed by a single wash with a low imidazole concentration (10 mM) using 5 ml to remove any non-specifically bound contaminants, and the resulting fractions were collected. The His-tagged proteins were eluted using a higher concentration of imidazole (500 mM, pH 8). Eluted fractions (E1-E8) were collected and analysed using SDS-PAGE (Figure 4.3a). The analysis included comparisons with the flow-through (FT), membrane (Mb), solubilized membrane (Smb), insoluble material (Ism), and wash one (W1) containing buffer free imidazole. It is worth noting that the flow-through (FT) and wash one (W1) demonstrated that a significant amount of the solubilized material did not bind to the resin. The free imidazole wash (W1) and the low imidazole concentration wash (data not shown) effectively removed untagged proteins, confirming that no detectable protein was eliminated from the column during this procedure. In the elution fractions with a high concentration of imidazole (500 mM, E2-E8), contaminating proteins were observed at around 100 kDa and 45 kDa. However, in the first elution fraction (E1), no proteins were eluted, and in fractions E6-E10, there was a decrease in the intensity of the higher protein band at around 100 kDa and the lower protein band at around 45 kDa, along with a smaller amount of protein around 60 kDa. Furthermore, bands around 35 kDa, which closely resemble the protein of interest, were visible in the gel, similar to those seen in Figure 4.2a. The western blot (Figure 4.3b) confirmed the presence of faint bands of proteins around 60 kDa or higher in fractions E2-E4, which were not displayed on the SDS-PAGE gel. Additionally, the Western blot reveals a similar band pattern in both the membrane (Mb) and the insoluble material (ISM). Notably, there is a high intensity of bands in the ISM at approximately 100 kDa, 65 kDa, 50 kDa, and 35 kDa. It is evident that a significant amount of the protein of interest remains undissolved in both the membrane and ISM, indicating that a substantial portion of the protein has not been effectively solubilized.

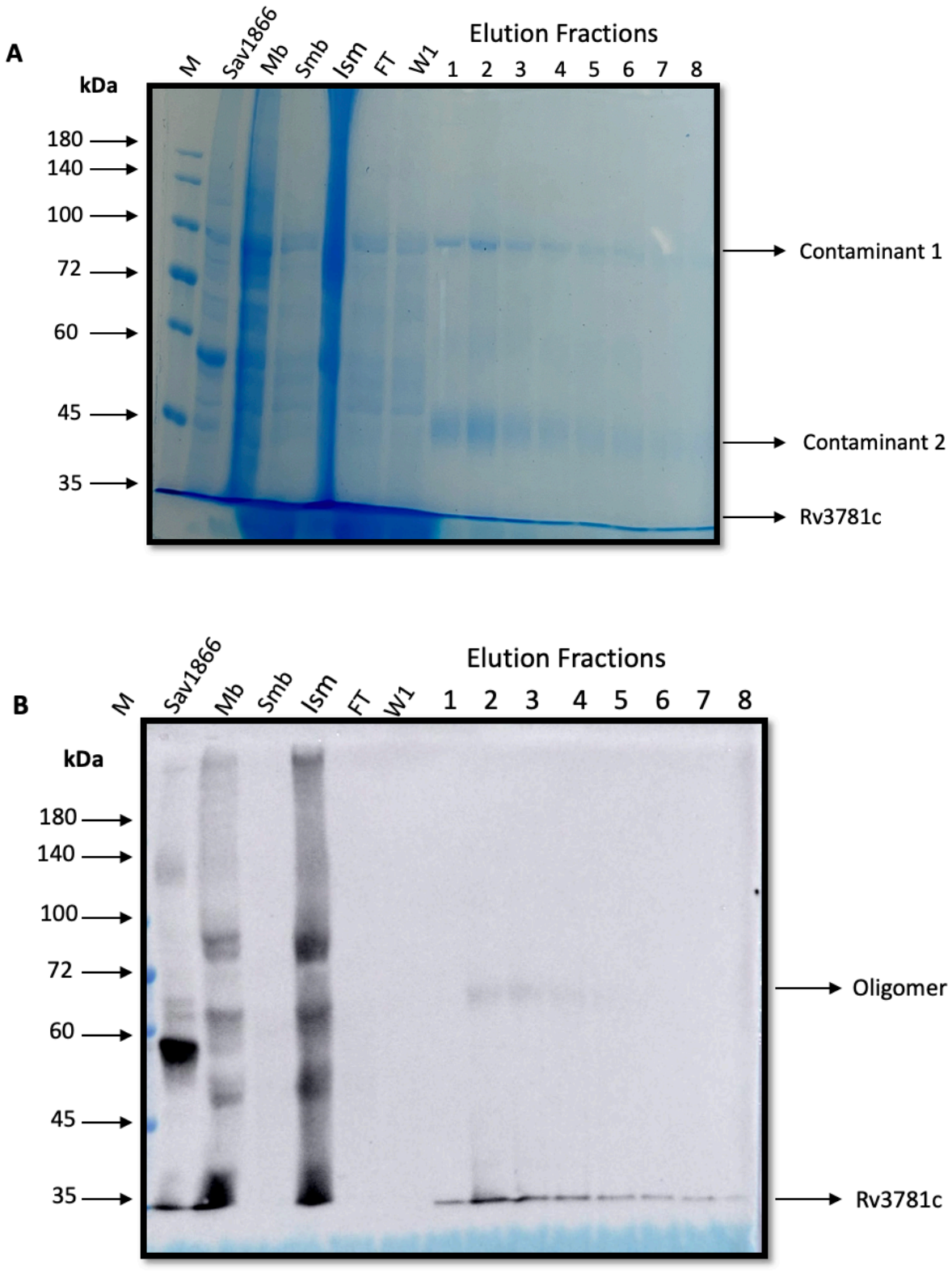


Figure 4.3: Rv3781c/ Rv3783c Ni-NTA Purification. A) SMA solubilized membrane were incubated with Ni-NTA overnight, then transferred to a gravity flow column, and flow through (FT) collected. Resin then washed 4 times with 5 ml solubilization buffer and collected 4 washes (shown the first wash W1), then washed with 5 ml of 10 mM imidazole pH 8 (data not shown). Then His-Tagged proteins were then eluted with 10 ml of 500 mM

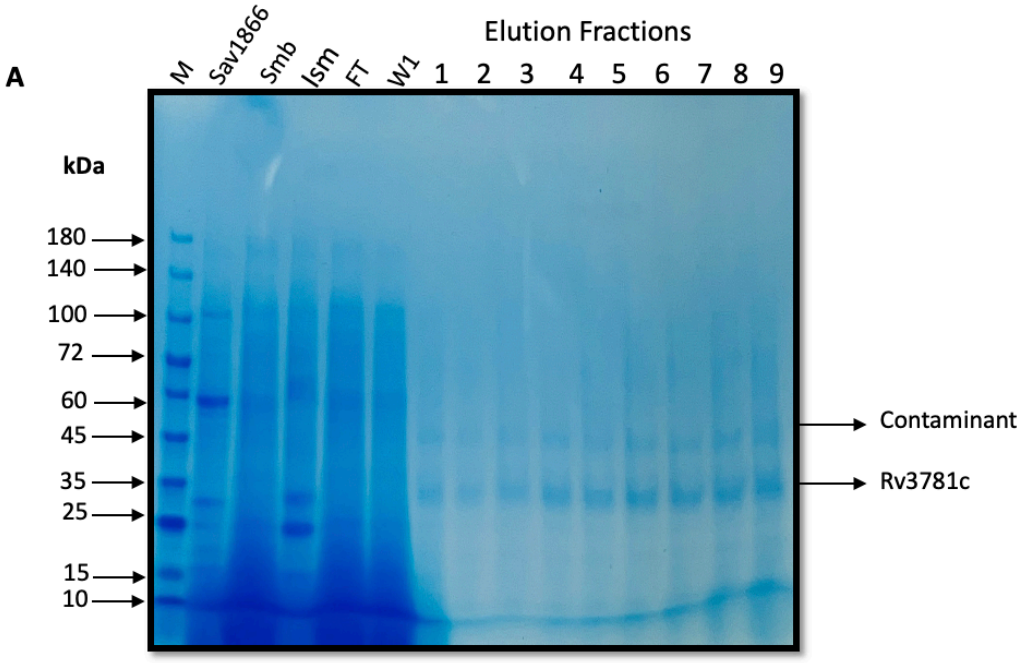
imidazole buffer pH 8 (E1- E8). Fractions were analysed by electrophoresis on an 10 % handmade SDS- PAGE and visualized by Coomassie stained. 6Xhis-Sav1866 (Sav1866), Membrane (Mb), Solubilized membrane (Smb), Insoluble material (Ism). B) The SDS-PAGE was also transferred to a nitrocellulose membrane and probed with an anti-His-Tagged antibody**.**

Reducing the imidazole concentration and adjusting the elution volume has not resulted in any significant change in the outcome of the protein purification. The protein yield remains low, and the purity is still far from ideal. Notably, the protein of interest continues to migrate very close to the dye front during elution, as observed in the initial purification, making it challenging to visualize on the Coomassie-stained gel. To address this issue, precast gradient gels were used in all future analyses to enhance protein separation while following the same protocol outlined in section 4.2.2.

4.2.3 Improving purification by using precast gradient gels:

To enhance analysis of the purification, we switched from hand-cast gels with 10% acrylamide to 4 –15% and 4 – 20 % precast gradient gels (Mini-Protean TGX, BioRad, U.K.). Rv3781 and Rv3783 were expressed, solubilized, and bound to the resin using the same protocol described in Section 4.2.2. Following an overnight incubation with Ni-NTA resin, the resin was washed, and the His-tagged proteins were eluted with imidazole. The eluted fractions were collected and subjected to analysis using SDS-PAGE and western blot (Figure 4.4 a and b), in addition to comparisons with the flow-through, the first wash with buffer-free imidazole, solubilized membrane (SM), insoluble material (ISM), and 6Xhis-Sav1866 (which provides a control for the Western Blot). One significant observation with the use of gradient gels was the improved separation of proteins throughout the gel. This enhanced separation is particularly evident in the elution fractions displayed in the Coomassie gel. The flow-through (FT) and the first wash (W1) highlight the substantial amount of solubilized material that did not bind to the resin. The wash steps with buffer-free imidazole and low imidazole concentration buffer effectively removed any untagged proteins, with no detectable proteins being removed from the column. Furthermore, it is worth noting that the amount of solubilized protein was comparatively low when compared to the

insolubilized fraction. Upon elution with a high concentration of imidazole (500 mM, E1-9), two distinct bands across the Coomassie gel were observed, representing eluted proteins around 45 kDa and 35 kDa. The western blot detected His-tagged proteins close to the molecular weight (Mw) of the expressed protein, appearing as faint bands around 35 kDa across all elution fractions, corresponding to the same bands observed on the SDS-PAGE gel (E1-9). Moreover, the western blot revealed faint bands around 60 kDa that were not visible on the Coomassie gel. Interestingly, there was a loss of bands around 45 kDa in the western blot compared to what was present in the Coomassie gel in Figure 4.3a.



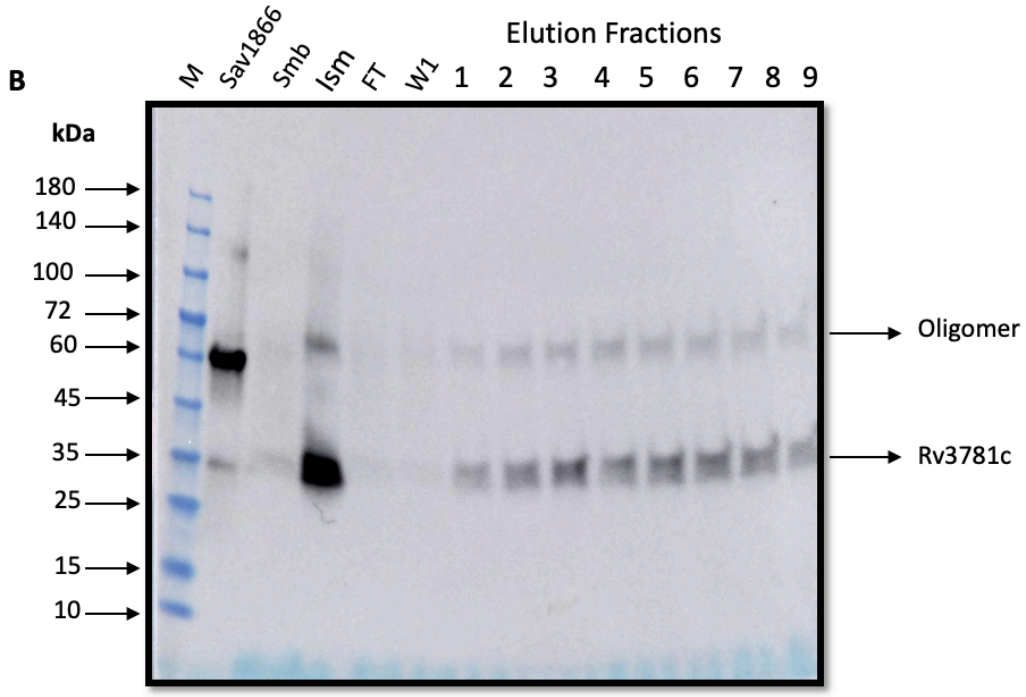
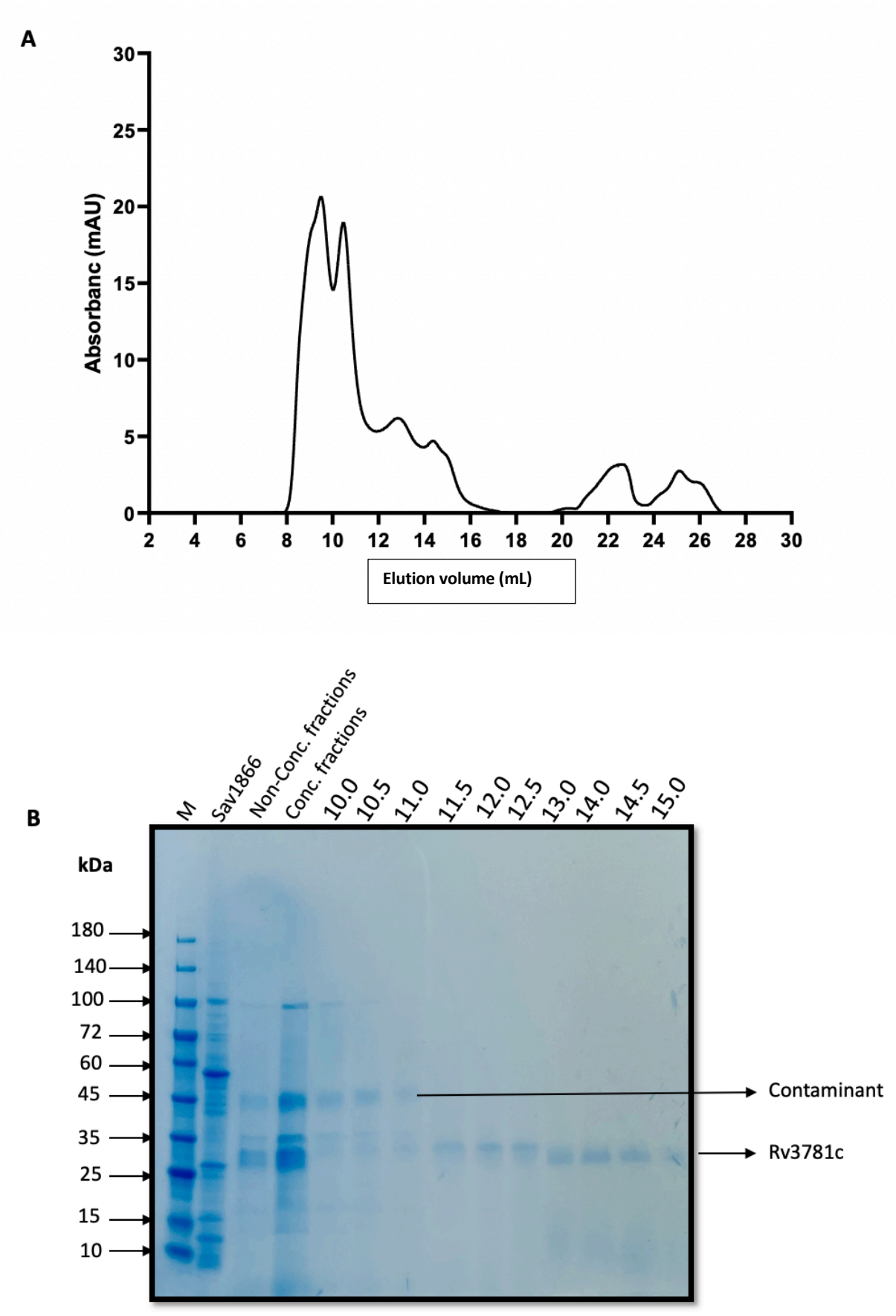


Figure 4.4: Rv3781c/ Rv3783c Ni-NTA Purification. A) SMA solubilized membrane were incubated with Ni-NTA overnight, then transferred to a gravity flow column, and flow through (FT) collected. Resin then washed 4 times with 5 ml solubilization buffer and collected 4 washes (shown the first wash), then washed with 5 ml of 10 mM imidazole pH 8 (data not shown). His-Tagged proteins were then eluted with 10 ml of 500 mM imidazole buffer pH 8 (E1- E9). Fractions were analysed by electrophoresis on an 4-15 % precast SDS- PAGE and visualized by Coomassie stained. 6Xhis-Sav1866 (Sav1866), Solubilized membrane (SM), Insoluble material (Ism). B) The SDS-PAGE was also transferred to a nitrocellulose membrane and probed with an anti-His-Tagged antibody.

Switching to a gradient gel has resulted in a notable improvement the ability to analyse the purification, allowing for a broader range of protein separation. When comparing this purification method to the one described in section 4.2.2, we observed the loss of one of the contaminant bands. Additionally, the protein of interest migrated closer to its expected molecular weight and became clearly visible on the Coomassie-stained gel. Based on the relative success of the NTA purification, Size Exclusion Chromatography (SEC) was added to the protocol post NTA purification to enhance protein purity and yield.

4.2.3.1 Size Exclusion Chromatography (SEC)

In the next step of purification, the elution fractions were combined and concentrated using a 50,000 Molecular weight cut-off (MWCO) concentrator. The concentrated sample was then subjected to separation using a Superdex 200 Increase 10/300 size exclusion column. Figure 4.5a displays the elution profile of the sample at 280 nm, which indicates the elution of proteins. The elution of proteins occurred over the middle range of elution volumes, with four major peaks observed at 9.5 mL, 10.5 mL, 13 mL, and 14.5 mL. Fractions collected from this elution profile were subsequently analysed using SDS-PAGE and western blotting, as demonstrated in Figure 4.5b and c. Both Coomassie stain gel and western blot analysis confirmed that Rv3781c and Rv3783c eluted within the range of 10.0 mL to 13.0 mL. This was verified by the detection of the His-Tag of Rv3781c across the elution fractions. The western blot analysis of the fraction spanning 10.0 mL to 13.0 mL revealed two distinct bands. One band was approximately 35 kDa, which was particularly evident in fraction 12.0 mL – 13.0 mL and closely matched the known molecular weight of Rv3781c, which is 29.96 kDa. The other band detected on the western blot had a molecular weight of around 60 kDa, suggesting it might be an oligomer that includes Rv3781c. Additionally, the Coomassie gel displayed a corresponding band at the correct molecular weight for Rv3781c. Importantly, there was no indication of the 60 kDa band observed in the western blot. Furthermore, the Coomassie gel exhibited a band at approximately 45 kDa in the first three fractions spanning 10.0 mL to 11.0 mL.


A
Absorbanc (mAU)
30
25
20
15
10
5
0
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
Elution volume (mL)
B
M
Sav1866
Non-Conc. fractions
Conc. fractions
10.0
10.5
11.0
11.5
12.0
12.5
13.0
14.0
14.5
15.0
kDa
180
140
100
72
60
45
35
25
15
10
Contaminant
Rv3781c

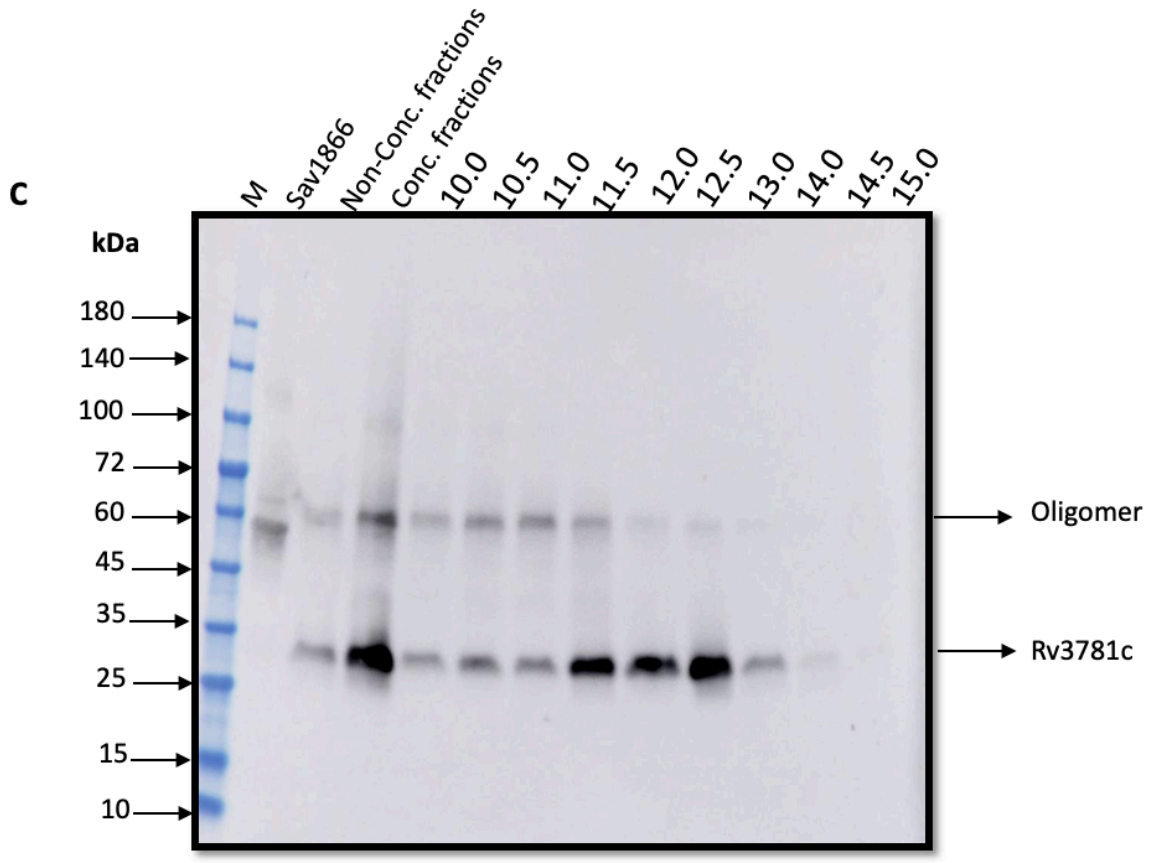


Figure 4.5: SEC purification of Rv3781c/ Rv3783c: Elution fraction fractions from Ni- NTA purification were collected, concentrated, and injected onto a Superdex 200 Increase 10/300 GL column. The column was run at 0.5 mL/min collecting 0.5 mL fractions. **A)** SEC elution profile. Elution of protein from the column was monitored by absorbance at 280nm. **B)** SDS-PAGE of elution profile 4 -15 % Precast SDS-PAGE of the SEC elution fractions analysed by electrophoresis and protein bands visualized by Coomassie staining**. C)** The SDS-PAGE was also transferred to a nitrocellulose membrane and probed with an anti-His-Tagged antibody**.**

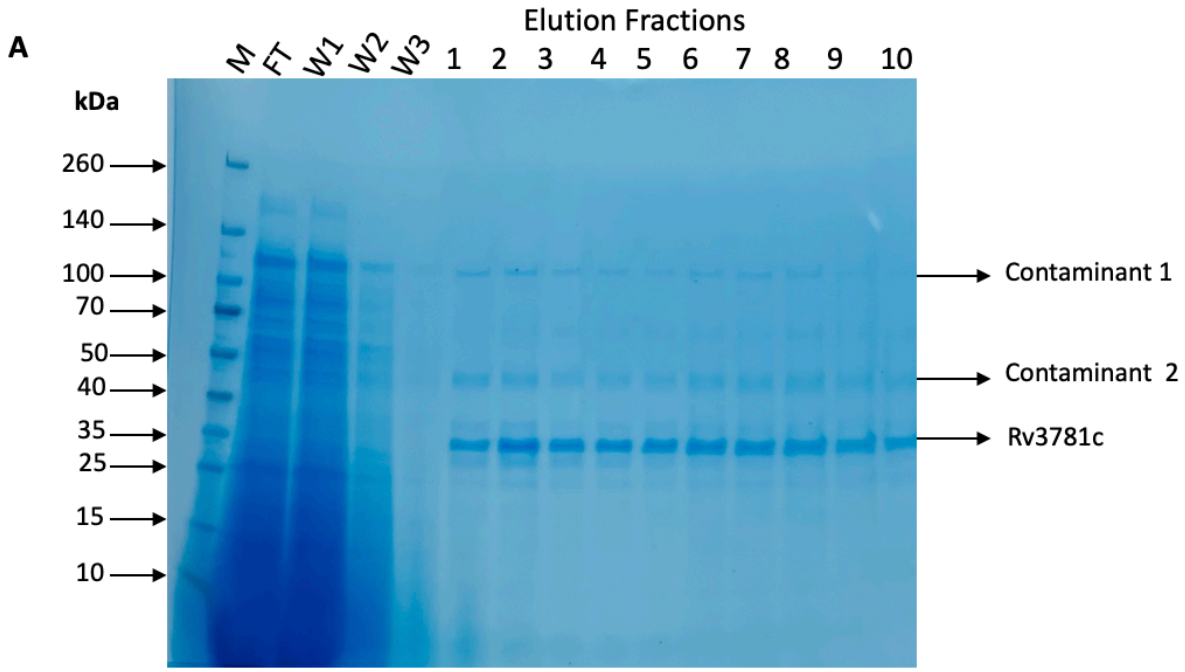
Examining the concentrated protein before injection revealed four distinct bands, approximately around 100 kDa, 45 kDa, 35 kDa, and 30 kDa. However, after running the sample through Size Exclusion Chromatography, the two higher bands disappeared, leaving only a single band that eluted between 11.5 mL and 15.0 mL. Furthermore, the migration of protein bands in the last fractions suggests the possible separation of the protein complex into two distinct bands. While successfully isolating a single band after the second purification step of the protein of interest, the protein yield remains insufficient. To address

this issue, the initial protocol was further modified by introducing additional washes with low imidazole concentration to remove more of the contaminating proteins before eluting with a high imidazole concentration. As a further change to increase yields the bacterial cells where induced when the OD_{600} falls within the range of 0.6 to 0.8. This approach aims to boost the growth rate of the cells, enhancing the potential production of the protein of interest.

4.2.4 Revised protocol 2: more imidazole washes:

Rv3781c and Rv3783c were overexpressed in the BL21 (DE3) cell strain, the induction was carried out overnight when the OD_{600} reached a range between 0.6 and 0.8. Following induction, cells were harvested and membranes were prepared as describe in section 4.2.1. The SMA-solubilized membrane fraction was bound to prepared Ni-NTA resin overnight on a rotating platform at 4°C to purify Rv3781c and Rv3783c, effectively separating them from other solubilized proteins. The resin was then transferred to a gravity flow column, and the flow-through was collected. The resin underwent four washes with a free imidazole buffer (B), each time using 5 ml. This was followed by a single wash with low imidazole concentrations of 10 mM, 50 mM, and 100 mM, using 10 mL in each case, to remove any non-specifically bound contaminants, and the resulting fractions were collected. The His-tagged proteins were eluted using a 10 mL solution with a higher concentration of imidazole (500 mM, pH 8). Eluted fractions were collected and analysed on SDS-PAGE and western blot (Figure 4.6a and b), alongside the flow-through and the first three washes of buffer-free imidazole. It's notable that a significant amount of the solubilized material did not bind to the resin, as evidenced by the flow-through fraction (FT), the free imidazole washes (W1-3), and the low imidazole concentration wash (data not shown). This indicates that no detectable protein was removed from the column during these steps. Furthermore, the Coomassie gel for the elution fractions (E1-10) showed three distinct bands, with the first two being faint, at approximately 120 kDa and between 40-50 kDa, and the third one being a strong band at around 35 kDa. We observe a higher protein yield in this purification

compared to the previous ones, especially in the case of section 4.2.3. Additionally, we've identified a new contaminant band at 120 kDa that wasn't observed in the previous purification. The western blot detected the His-tagged proteins at the expected molecular weight of Rv3781c, appearing as strong bands at approximately 35 kDa across all elution fractions, corresponding to the bands observed in the SDS-PAGE (E1-10). Additionally, the western blot revealed faint bands below the 70 kDa mark in the elution fractions (E2-10), indicating the presence of an oligomer that may include Rv3781c.



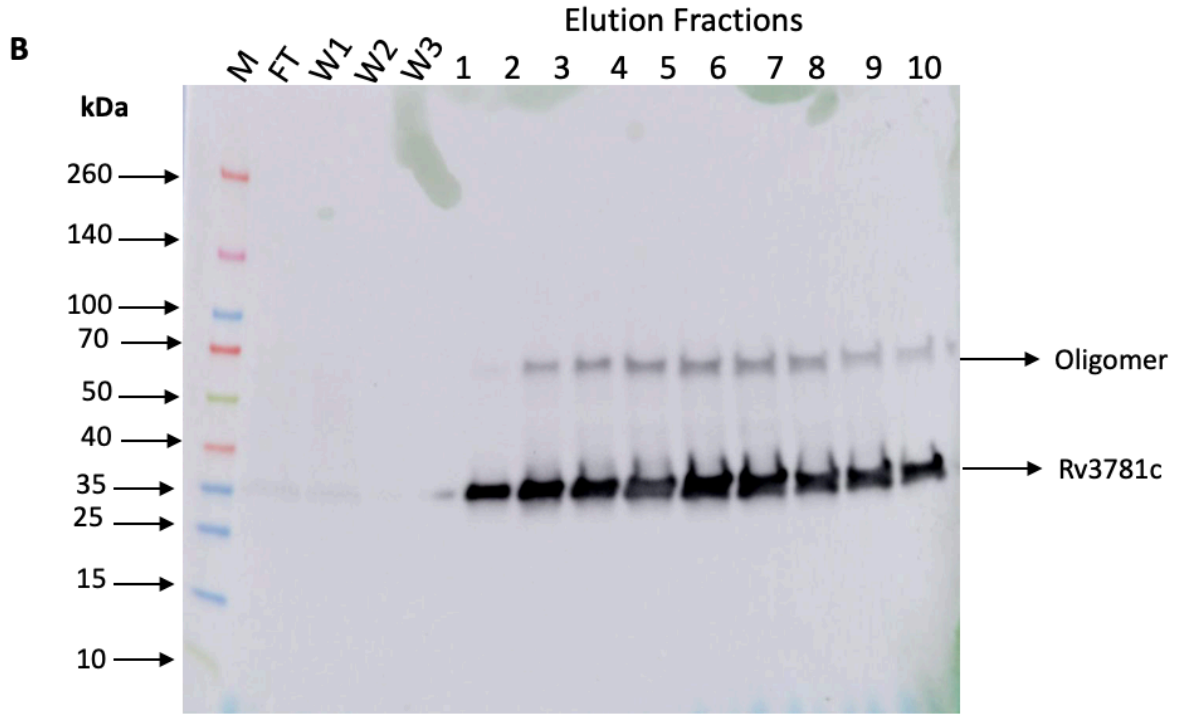


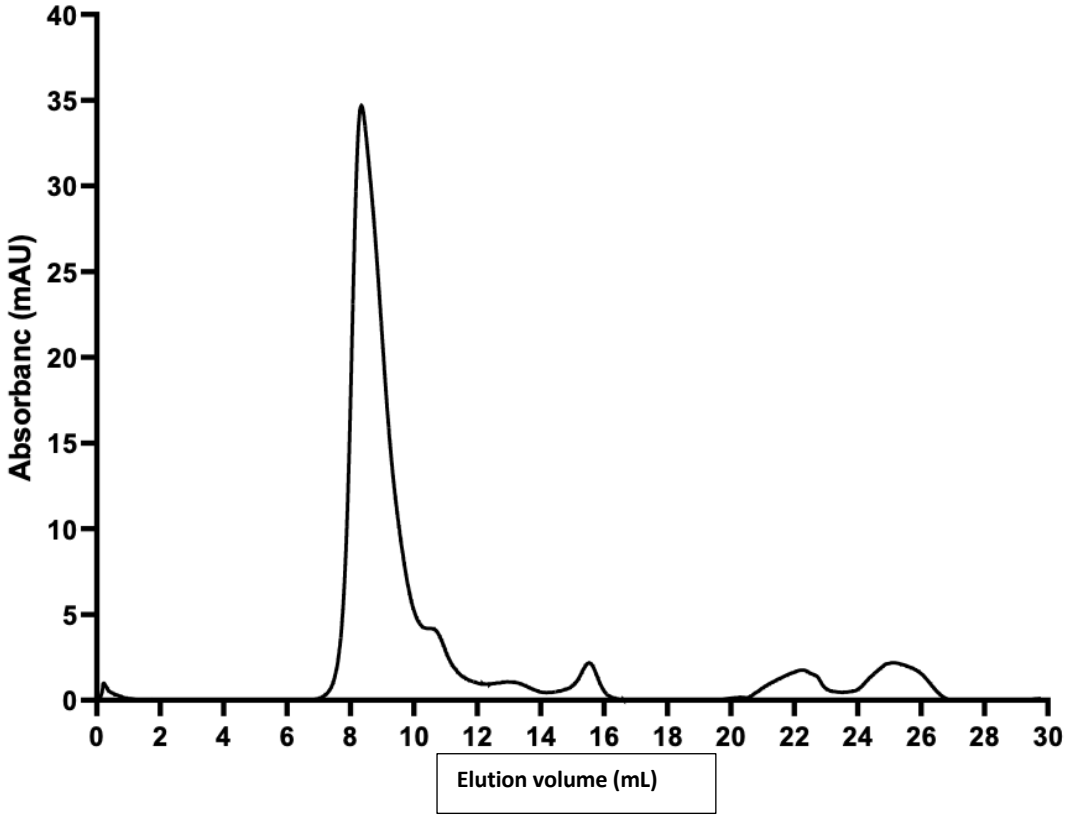
Figure 4.6: Rv3781c/ Rv3783c Ni-NTA Purification. A) SMA solubilized membrane were incubated with Ni-NTA overnight, then transferred to a gravity flow column, and flow through (FT) collected. Resin then washed 4 times with 5 ml solubilization buffer and collected 4 washes (shown the first three washes W1, W2, and W3), then washed with 5 ml of 10 mM, 50 mM, and 100 mM imidazole pH 8 respectively (data not shown). Then His-Tagged proteins were then eluted with 10 ml of 500 mM imidazole buffer pH 8 (E1- E10). Fractions were analysed by electrophoresis on an 4-15 % precast SDS- PAGE and visualized by Coomassie stained. **B)** The SDS-PAGE was also transferred to a nitrocellulose membrane and probed with an anti-His-Tagged antibody**.**

Modifying the Ni-NTA purification process by adjusting the induction time point and incorporating three low-concentration imidazole washes has significantly enhanced the purification. As a result, protein yield has increased. However, there is a notable concern regarding protein purity at this stage. Two faint bands of contaminants were eluted alongside the target protein.

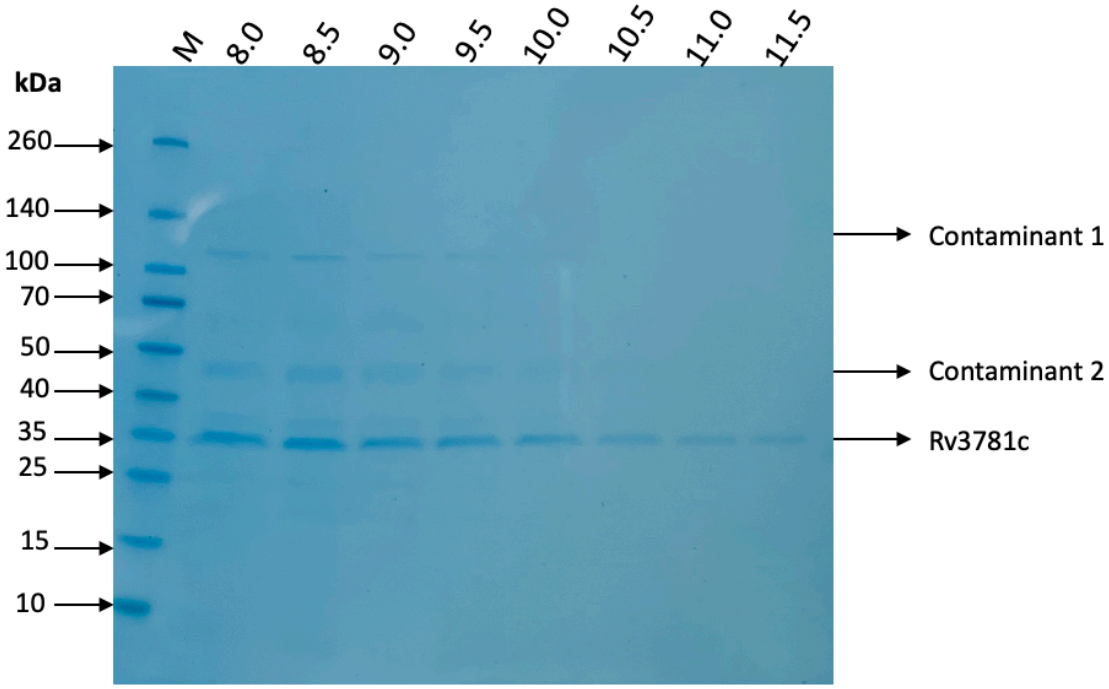
4.2.4.1 Size Exclusion Chromatography (SEC)

In the next step of purification, the combined and concentrated the elution fractions from the Ni-NTA elution were separated using a Superdex 200 Increase 10/300 size exclusion column. Figure 5.7a displays the elution profile of the sample at 280 nm, which reveals the protein elution pattern. The elution of proteins occurred within a moderate range of elution volumes, characterized by one major peak and a small shoulder peak between 8.0 mL and 11.5 mL. Fractions collected from this elution profile underwent testing via SDS-PAGE and western blotting, as depicted in Figure 5.7b and c. Both Coomassie stain gel and western blot analysis confirmed that Rv3781c and Rv3783c were eluted within the range of 8.0 mL to 11.5 mL, as detected by the presence of the His-Tag of Rv3781c across the elution. The Coomassie stain gel revealed the protein fractions along with some contaminants around 120 kDa and 45 kDa. However, these contaminant proteins disappeared between 10.0 mL and 11.5 mL fractions, leaving only the band representing the correct molecular weight of Rv3781c. The western blot analysis exhibited a distinct band for Rv3781c, appearing at approximately 35 kDa, which closely aligns with the known molecular weight of Rv3781c, measured at 29.96 kDa. Additionally, a fainter band around 60 kDa was observed, indicating the potential presence of an oligomer involving Rv3781c. Furthermore, the Coomassie gel displayed a corresponding band at the correct molecular weight for Rv3781c. It's important to note that both Rv3781c and Rv3783c proteins are of similar molecular weights and may not fully separate on the Coomassie gel.

A


Absorbanc (mAU)
40
35
30
25
20
15
10
5
0
0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30
Elution volume (mL)

B


M
8.0
8.5
9.0
9.5
10.0
10.5
11.0
11.5
kDa
260
140
100
70
50
40
35
25
15
10
Contaminant 1
Contaminant 2
Rv3781c

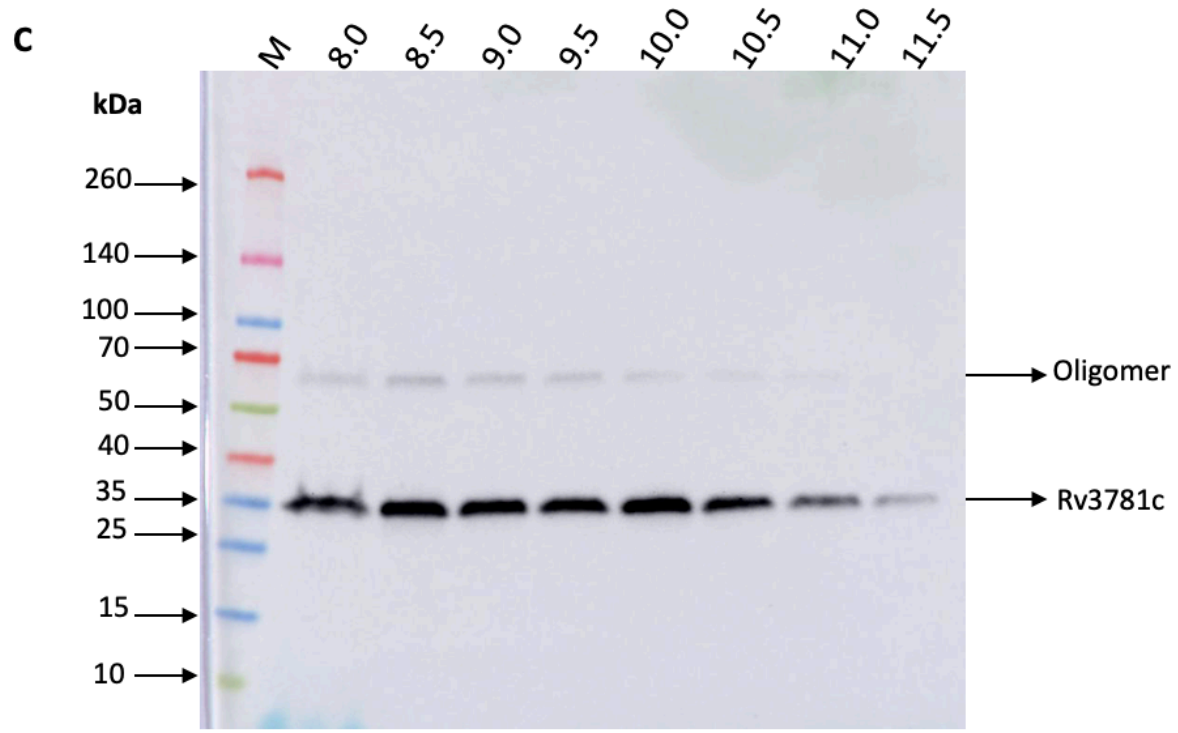


Figure 5.7: SEC purification of Rv3781c/ Rv3783c Elution fractions from Ni- NTA purification were collected, concentrated, and injected onto a Superdex 200 Increase 10/300 GL column. The column was run at 0.5 mL/min collecting 0.5 mL fractions. **A)** SEC elution profile. Elution of protein from the column was monitored by absorbance at 280nm. **B)** SDS-PAGE of elution profile 4 -15 % Precast SDS-PAGE of the SEC elution fractions analysed by electrophoresis and protein bands visualized by Coomassie staining**. C)** The SDS-PAGE was also transferred to a nitrocellulose membrane and probed with an anti-His-Tagged antibody**.**

Size exclusion chromatography has shown a slight improvement in protein purity, which is clearly visible in the 10.5 mL to 11.5 mL fractions, as it has established a single band at the correct molecular weight for Rv3781c. At this point, although protein of high purity has been achieved the yield is too low for downstream analysis. To address these challenges, further modifications where made to the protocol in section 4.2.4. Firstly, the pH of all buffers containing imidazole were altered to 7.5 from 8. This adjustment aims to improve protein stability and assist in maintaining its structural integrity. Secondly, the composition of the elution buffer was adjusted to 20 mL to ensure complete elution of the target protein without leaving any residue on the column. I will also apply the same volume of the low imidazole concentration washes in section 4.2.2 to remove as many contaminant proteins as possible. Finally, if the Ni-NTA purification is successful, the molecular weight cut-off of the

protein concentrator used to concentrate samples prior to SEC will be reduced from 50,000 to 3,000 to address issues of protein loss through the concentrator.

4.2.5 Revised protocol 3: Changing buffer pH and increasing the elution volume:

4.2.5.1 10 mL 10 mM imidazole, 50 mM imidazole, 100 mM imidazole, and 20 mL 500 mM elution buffer pH 7.5:

Rv3781c and Rv3783c were expressed, solubilized, and bound to the resin using the same protocol described in Section 4.2.4. Following an overnight incubation with Ni-NTA resin, the resin was washed four times with a free imidazole buffer (B) pH 8, each time using 5 ml. Subsequently, a series of low imidazole concentration washes (10 mM, 50 mM, 100 mM, pH 7.5) were performed, each once with 10 mL of solution to remove contaminants that were not specifically bound. The His-tagged protein was eluted with 20 mL of a higher imidazole concentration solution (500 mM, pH 7.5). Eluted fractions were collected and analysed using SDS-PAGE and western blot (Figure 4.8a and b), alongside comparisons with the membrane (Mb), solubilized membrane (Sm), flow-through, and the first wash containing buffer-free imidazole. Notably, a significant amount of the solubilized material did not bind to the resin, as demonstrated by the flow-through fraction (FT), and the free imidazole washes (W1), with the low imidazole concentration wash yielding no detectable protein removal from the column. However, the elution with a high concentration of imidazole (500 mM, E1-10) did not result in an improved yield of the proteins of interest. Moreover, two contaminant proteins around 120 kDa and 45 kDa still appeared as faint bands, particularly in the last fractions. The western blot successfully detected the His-tagged proteins at the correct molecular weight in the membrane and the elution fractions.

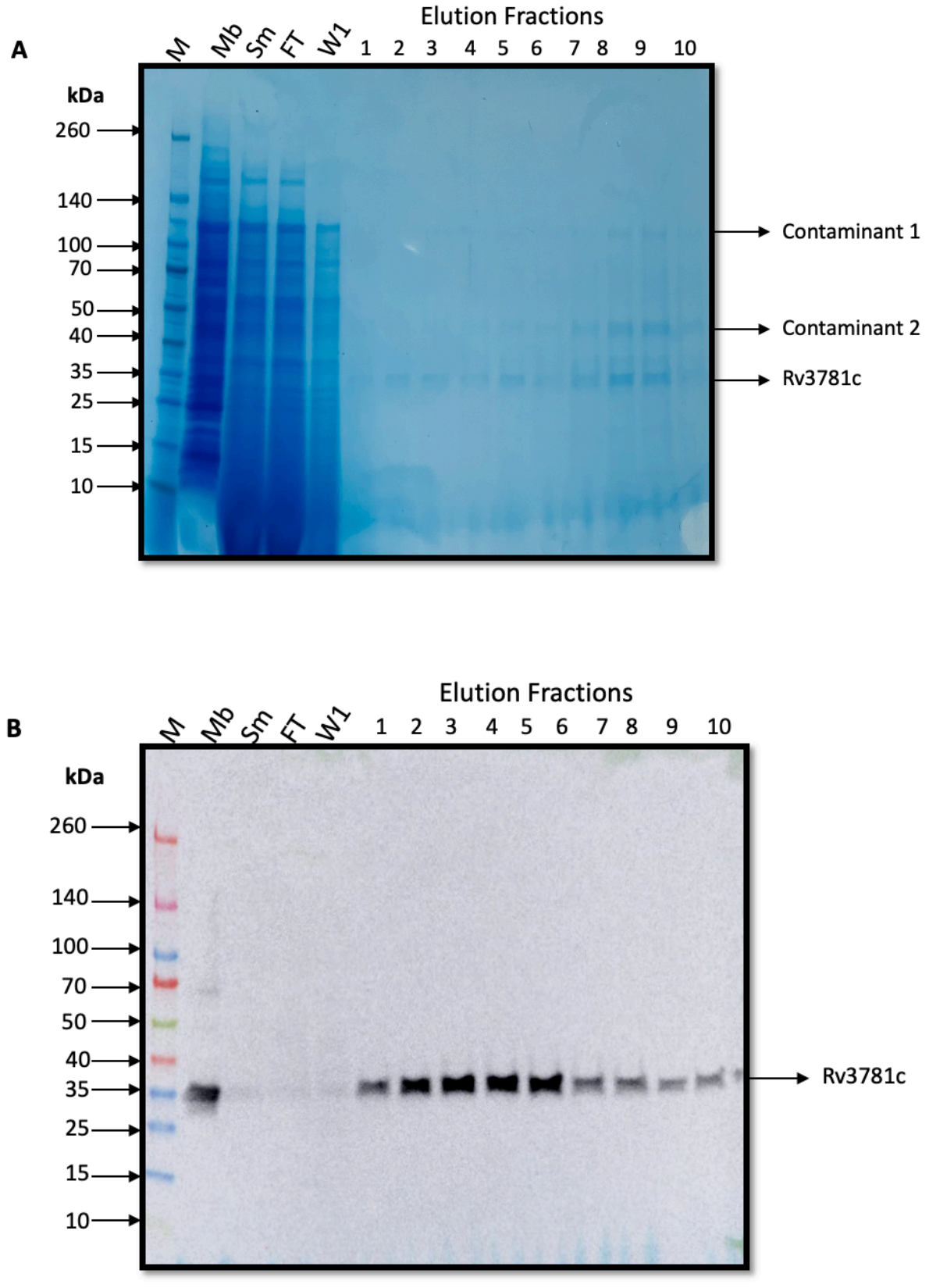


Figure 4.8: Rv3781c/ Rv3783c Ni-NTA Purification. A) SMA solubilized membrane were incubated with Ni-NTA overnight, then transferred to a gravity flow column, and flow through (FT) collected. Resin then washed 4 times with 5 ml solubilization buffer and collected 4 washes (shown the first wash W1), then washed with 10

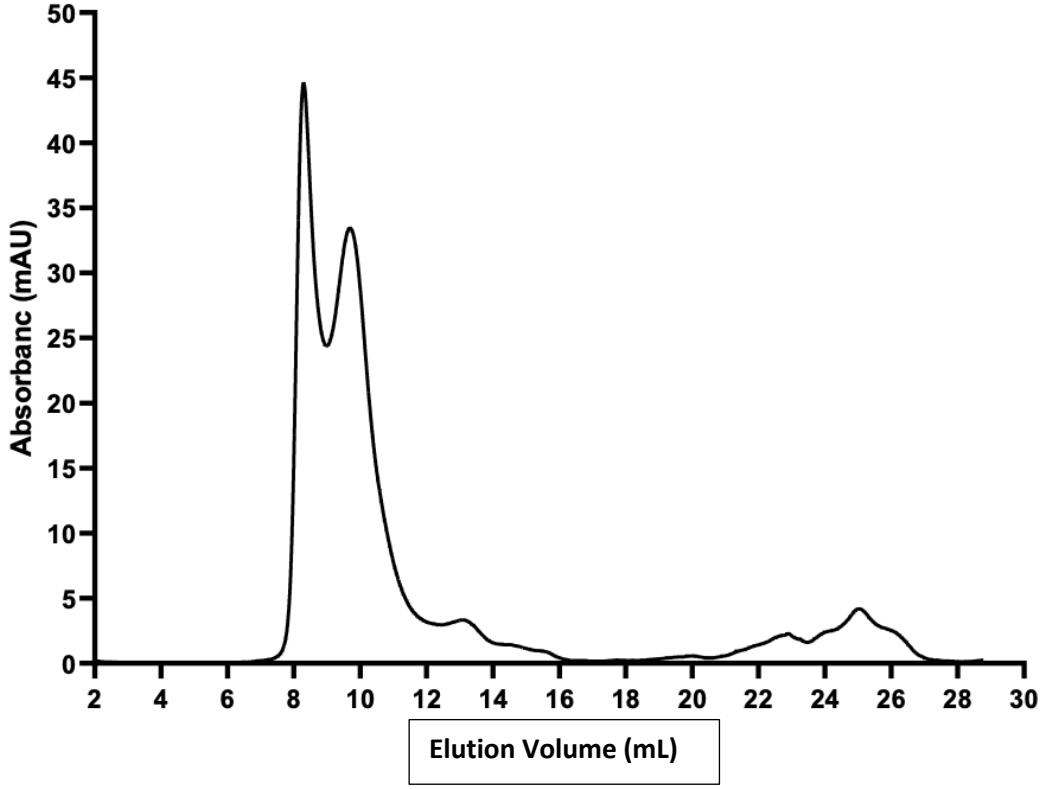
ml of 10 mM, 50 mM, and 100 mM imidazole pH 7.5 respectively (data not shown). Then His-Tagged proteins were then eluted with 20 ml of 500 mM imidazole buffer pH 7.5 (E1- E10). Fractions were analysed by electrophoresis on 4 – 20 % precast SDS- PAGE and visualized by Coomassie stained alongside the membrane (Mb), solubilized membrane (Sm), flow-through (FT), And first wash W1. B) The SDS-PAGE was also transferred to a nitrocellulose membrane and probed with an anti-His-Tagged antibody**.**

The results of the Ni-NTA purification following the recent adjustments did not meet expectations. The protein yield was lower in comparison to the Ni-NTA purification carried out in section 4.2.4, and the protein purity remained at the same level as before, showing no improvement. Upon analysing the western blot results, it became evident that we lost the oligomer that had been present in previous purification attempts. Considering this observation, it was decided to utilize size exclusion chromatography, employing a smaller molecular weight cut-off, to investigate whether the purification could be improved.

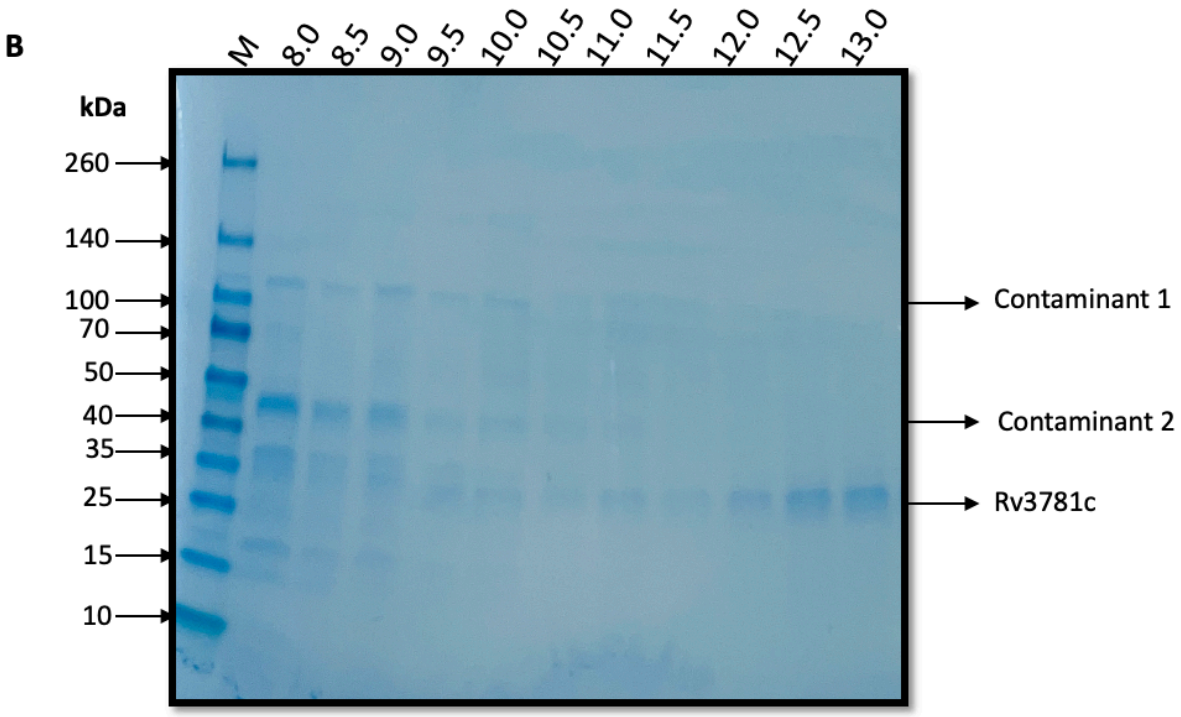
4.2.5.1.1 Size Exclusion Chromatography (SEC)

In the next stage of purification, we combined and concentrated the elution fractions using a 3,000 MWCO concentrator. The sample was subsequently subjected to separation using a Superdex 200 Increase 10/300 size exclusion column. Figure 4.9a illustrates the elution profile of the sample at 280 nm, indicating the pattern of protein elution. Proteins were eluted over a moderate range of elution volumes, with two major peaks occurring between 8.0 mL and 13.0 mL. Fractions collected from this elution profile underwent analysis through SDS-PAGE and western blotting, as shown in figure 4.9b and c. Both Coomassie-stained gel and western blot analysis confirmed that Rv3781c and Rv3783c were eluted between 8.0 mL and 13.0 mL, as indicated by the presence of the His-Tag of Rv3781c across the elution. The Coomassie-stained gel revealed the proteins of interest alongside other contaminants. These contaminant proteins were predominantly found in the initial fractions of the size exclusion chromatography (SEC) profile. The primary contaminants were approximately 120 kDa and 45 kDa in size. The western blot analysis exhibited a distinct band for Rv3781c, which was approximately 35 kDa and spanned the range of 8.0 mL to 13.0 mL. Additionally, it revealed a band at around 60 kDa, suggesting the possible presence of an oligomer that includes Rv3781c.

A


Absorbanc (mAU)
50
45
40
35
30
25
20
15
10
5
0
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
Elution Volume (mL)

B


M
8.0
8.5
9.0
9.5
10.0
10.5
11.0
11.5
12.0
12.5
13.0
kDa
260
140
100
70
50
40
35
25
15
10
Contaminant 1
Contaminant 2
Rv3781c

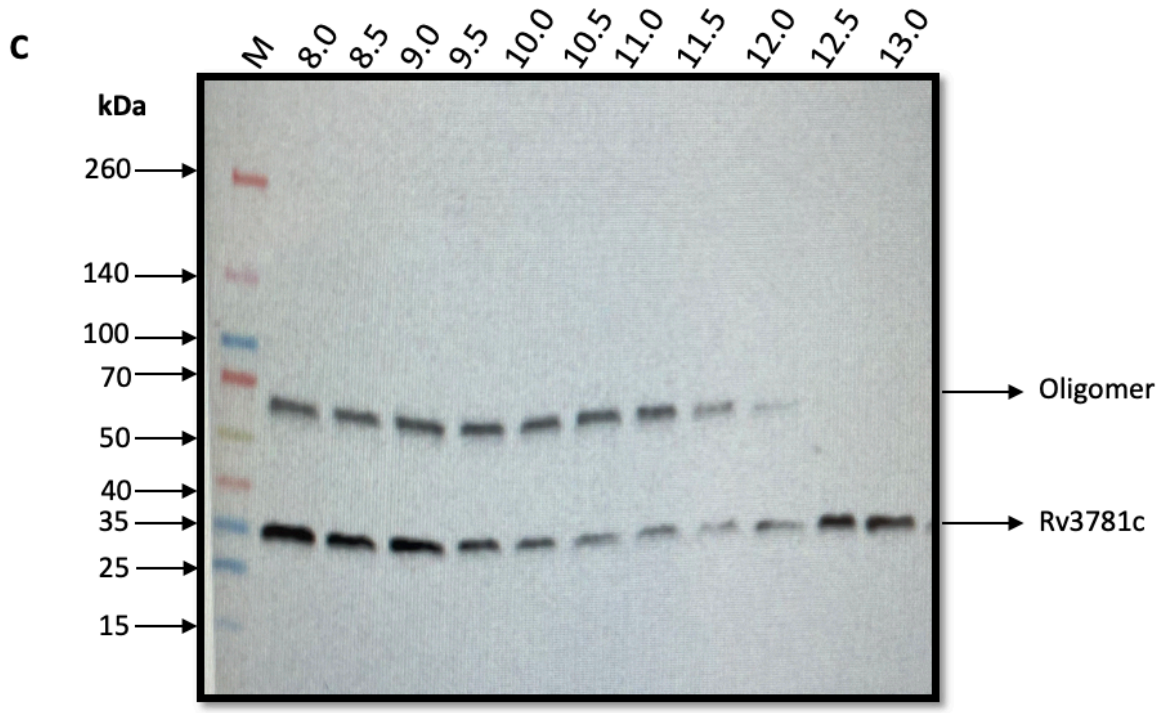


Figure 4.9: SEC purification of Rv3781c/ Rv3783c Elution fraction fractions from Ni- NTA purification were collected, concentrated, and injected onto a Superdex 200 Increase 10/300 GL column. The column was run at 0.5 mL/min collecting 0.5 mL fractions. A) SEC elution profile. Elution of protein from the column was monitored by absorbance at 280 nm. B) SDS-PAGE of elution profile 4 -20 % Precast SDS-PAGE of the SEC elution fractions analysed by electrophoresis and protein bands visualized by Coomassie staining. C) The SDS-PAGE was also transferred to a nitrocellulose membrane and probed with an anti-His-Tagged antibody**.**

The size exclusion chromatography profile displayed three distinct peaks, indicating a successful separation of proteins. However, upon analysing the fractions, there was not a significant change in protein yield and purity, which are the primary objectives of this step. Shifting to a smaller molecular weight cut-off resulted in an increase in the amount of protein added to the SEC, which as expected positively impacted protein yield. In light of these outcomes, further adjustments to the protocol we proposed protocol in an effort to enhance Ni-NTA purification. Th volume of the low imidazole concentration washes was reduced, along with a slight adjustment in the first imidazole wash, increasing it to 20 mM imidazole from the initial 10 mM imidazole, to help minimize the loss of the target protein during this process.

4.2.5.2 Revised Protocol 3: 5 mL 20 mM imidazole, 50 mM imidazole, 100 mM imidazole, and 20 mL 500 mM elution buffer pH 7.5:

Rv3781c and Rv3783c were expressed, solubilized, and immobilized on the resin following the protocol outlined in Section 4.2.5.1. The sole modification to this procedure involved commencing with a 20 mM imidazole wash instead of the previous 10 mM imidazole wash, as well as reducing the washing volume of the low imidazole concentration washes to 5 mL. Eluted fractions were collected and analysed using SDS-PAGE and western blot (Figure 4.10a and b), alongside comparisons with the membrane (Mb), solubilized membrane (Sm), flow-through (FT), and the first wash containing buffer-free imidazole. It is worth noting that a significant amount of the solubilized material did not bind to the resin, as evidenced by the flow-through fraction (FT), and the free imidazole washes (W1), with the low imidazole concentration wash showing no detectable protein removal from the column. However, elution with a high concentration of imidazole (500 mM, E1-10) revealed the presence of three bands on the Coomassie gel, representing eluted proteins around 120 kDa, 45 kDa, and 35 kDa. The western blot successfully detected the His-tagged proteins at the correct molecular weight of Rv3781c, appearing as clear bands at around 35 kDa across all elution fractions, corresponding to the bands around 35 kDa observed on the SDS-PAGE (E1-10). Additionally, the western blot exhibited faint bands around 60 kDa, which likely indicate protein oligomerization. Despite their appearance on the Coomassie gel, the western blot did not detect bands around 120 kDa and 45 kDa.

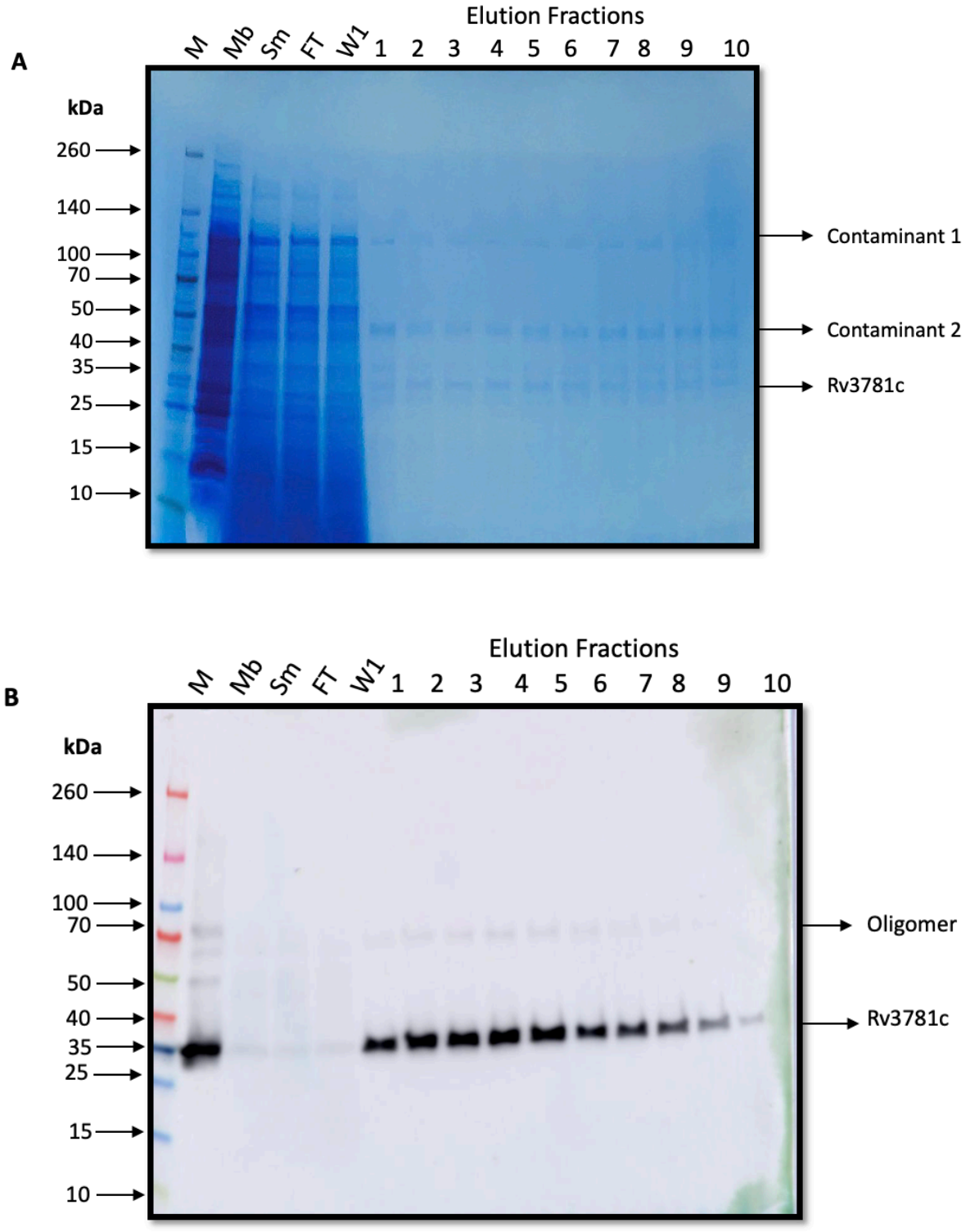


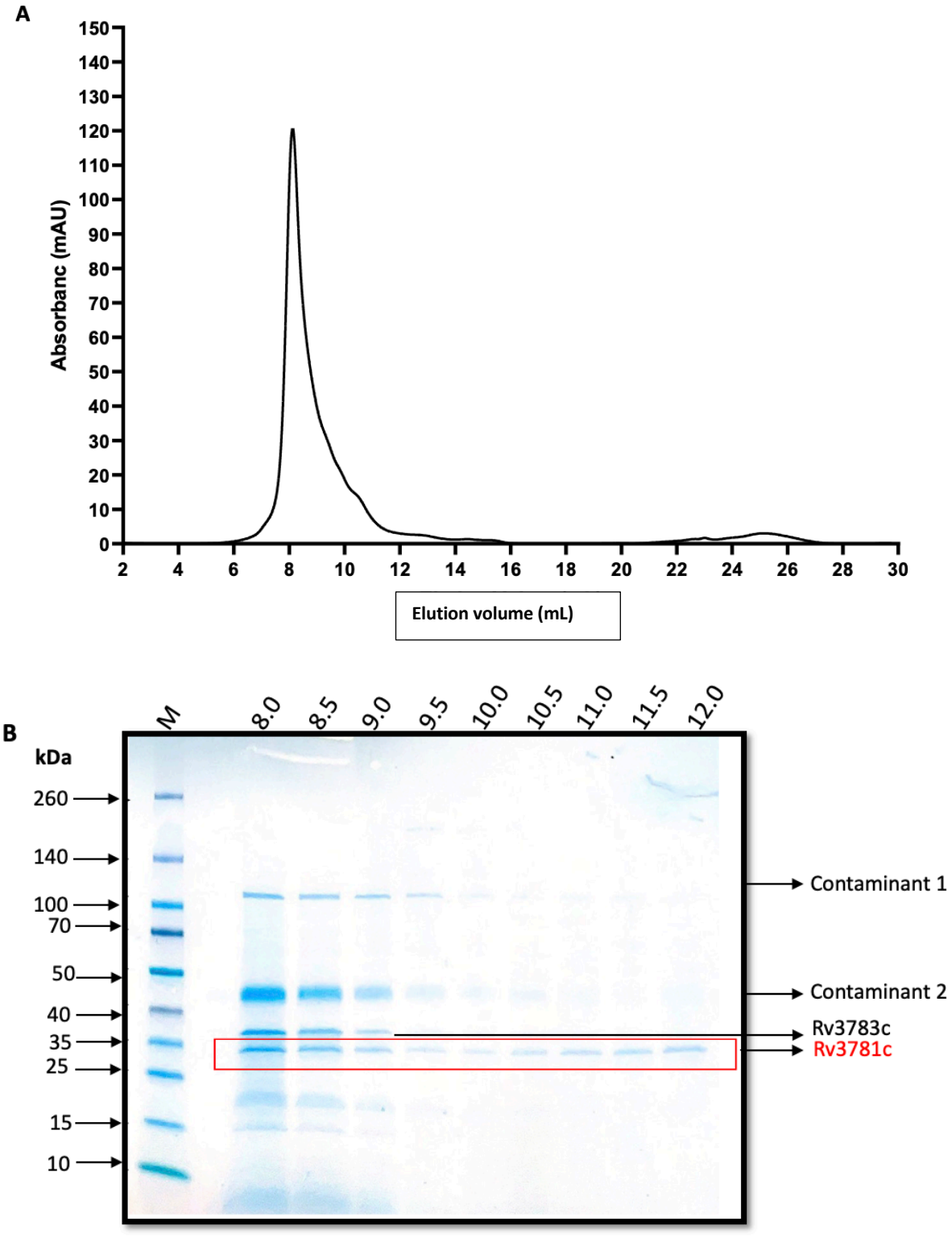
Figure 4.10: Rv3781c/ Rv3783c Ni-NTA Purification. A) SMA solubilized membrane were incubated with Ni-NTA overnight, then transferred to a gravity flow column, and flow through (FT) collected. Resin then washed 4 times with 5 ml solubilization buffer and collected 4 washes (shown the first wash), then washed with 5 ml of 10 mM, 50 mM, and 100 mM imidazole pH 7.5 respectively (data not shown). Then His-Tagged proteins were then eluted with 20 ml of 500 mM imidazole buffer pH 7.5 (E1- E10). Fractions were analysed by electrophoresis on 4 – 20 % precast SDS- PAGE and visualized by Coomassie stained alongside the membrane

(Mb), solubilized membrane (Sm), flow-through (FT), And first wash W1. B) The SDS-PAGE was also transferred to a nitrocellulose membrane and probed with an anti-His-Tagged antibody**.**

These minor adjustments yielded similar results in the Ni-NTA purification, consistent with what was observed in the purification process described in section 4.2.4. The elution fractions displayed the same banding pattern in both purifications, although the protein of interest had a higher yield in the section 4.2.4 purification.

4.2.5.2.1 Size Exclusion Chromatography (SEC)

In the next stage of purification, we combined and concentrated the elution fractions and applied them to a Superdex 200 Increase 10/300 size exclusion column. As shown in Figure 4.11a, this chart illustrates the elution profile of the sample at 280 nm, providing insight into the protein elution process. Notably, the proteins eluted over a moderate range of elution volumes, with one major peak and two smaller shoulder peaks occurring between 8 and 12 mL. Subsequently, fractions collected from this elution profile underwent analysis through SDS-PAGE and western blotting, as depicted in Figure 4.11b and c. Both Coomassie-stained gel and western blot analysis confirmed that Rv3781c and Rv3783c were eluted between 8.0 mL and 12.0 mL, as indicated by the presence of the His-Tag of Rv3781c across the elution. Coomassie-stained gel showed an improvement in purification and protein yield. However, some contaminants persisted, particularly around 120 kDa and 45 kDa, in the early fractions between 8.0 mL and 9.5 mL. These contaminants appeared to be consistent throughout the entire purification process. Additionally, the Coomassie gel displayed a band around 35 kDa, which overlapped with the protein of interest, Rv3781c. This band could potentially correspond to Rv3783c. In contrast, the western blot analysis exhibited a distinct band for Rv3781c, approximately 35 kDa, spanning the range of 8.0 mL to 12.0 mL. Notably, there was no sign of protein oligomerization in these fractions. Despite their presence on the Coomassie gel, the western blot did not detect bands around 120 kDa and 45 kDa.


A
Absorbanc (mAU)
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
Elution volume (mL)
B
M
8.0
8.5
9.0
9.5
10.0
10.5
11.0
11.5
12.0
kDa
260
140
100
70
50
40
35
25
15
10
Contaminant 1
Contaminant 2
Rv3783c
Rv3781c

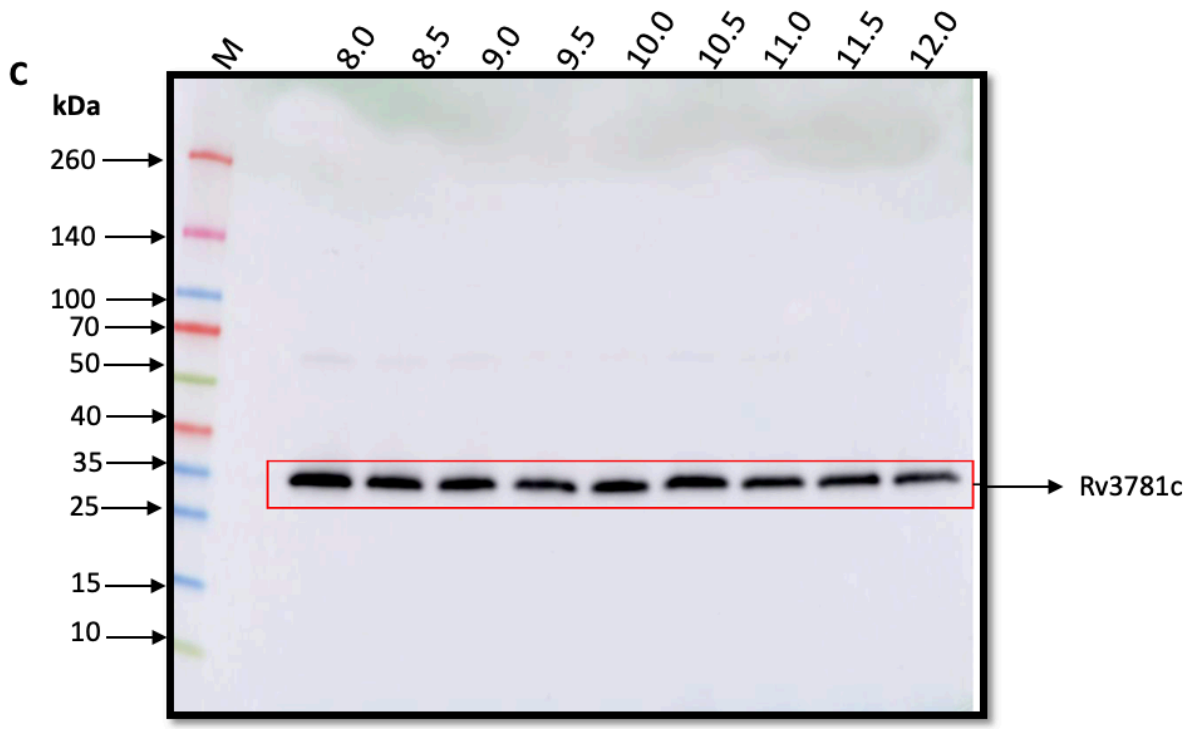


Figure 4.11: SEC purification of Rv3781c/ Rv3783c Elution fraction fractions from Ni- NTA purification were collected, concentrated, and injected onto a Superdex 200 Increase 10/300 GL column. The column was run at 0.5 mL/min collecting 0.5 mL fractions. A) SEC elution profile. Elution of protein from the column was monitored by absorbance at 280 nm. B) SDS-PAGE of elution profile 4 -20 % Precast SDS-PAGE of the SEC elution fractions analysed by electrophoresis and protein bands visualized by Coomassie staining. C) The SDS-PAGE was also transferred to a nitrocellulose membrane and probed with an anti-His-Tagged antibody.

The size exclusion chromatography chromatogram indicates an improvement in protein amount in the elution, suggesting an increase in protein yield. However, the protein separation was suboptimal, resulting in one major peak with two smaller shoulder peaks. This could pose challenges in achieving high protein purity. Upon analysing the fractions, a notable enhancement in the purity of the protein of interest was observed in the later fractions, particularly within the range of 10.0 to 12.0 mL. By doing so, we successfully removed a significant portion of contaminating proteins present in the first four fractions. Additionally, unlike in previous purifications, we did not observe signs of oligomerization in the western blot for this purification.

4.3 CHARACTERISATION OF THE COMPONENTS OF Rv3781c and Rv3783c IN STYRENE MALEIC ACID LIPID PARTICLES

4.3.1 Circular Dichroism (CD) of SMA solubilised Rv3781c and Rv3783c:

Circular dichroism is a technique used to analyse how molecules absorb different types of circularly polarized light. It is particularly valuable in deciphering protein structures due to the distinctive patterns that emerge during specific transitions in the peptide backbone. For instance, observing the circular dichroism spectrum of an alpha-helix reveals two negative points of reduced absorption at 208 nm and 222 nm. Beta-sheets, however, present less straightforward results because their structures exhibit more variation. Typically, they exhibit negative absorption around 216 nm.

CD data was collected for the sample corresponding to fraction in figure 4.12. This fraction contained Rv3781c which meant that the CD data would provide information on the secondary structure content of that protein. The CD data shows a relatively flat negative minimum between 208 and 222 nm and including 216 nm. This indicates that the sample contains both α-helix (208, 222 nm) and β-sheet (216 nm). To further assess these data the spectrum was deconvolved using Betsel. The Betsel analysis predicted that the protein contains 0.1% Alpha, 38.3% Beta and 61.7% loop/unfolded. Figure 4.13 shows the CD data and the result of the fit from Betsel showing that there is a good agreement between the two.

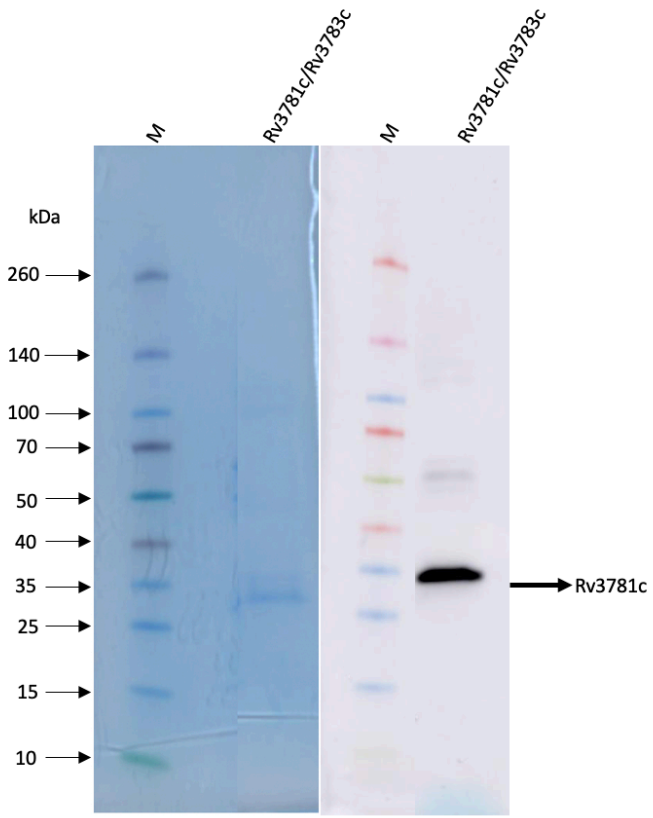


Figure 4.**12**: SDS-PAGE and Western blot of Rv3781c and R3783c fraction after SEC purification

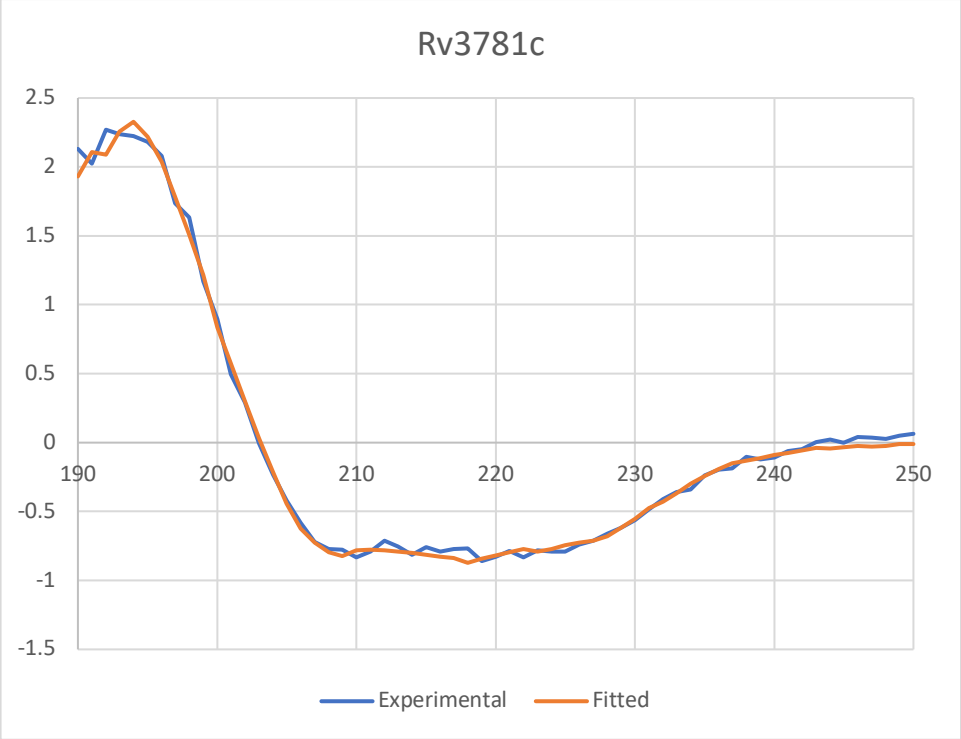


Figure 4.13: CD analysis of Rv3781c on SMALPs. The fractions obtained after the size-exclusion chromatography (SEC) purification of Rv3781c were examined using CD spectroscopy to evaluate its secondary structure. The analysis was performed using a JASCO J-1500 Circular Dichroism Spectrophotometer operated through Spectra Manager software. Measurements were conducted in a quartz glass cuvette with a pathlength of 1 mm, using a protein concentration of approximately 0.1 mg/mL.

4.3.2 Analytical Ultracentrifugation (AUC)

A sample was taken from fraction on figure 4.12 and concentrated before being analysed using Sedimentation Velocity Analytical Ultracentrifugation (svAUC). svAUC provides a measure of the molecular mass of proteins in solution. These data from the svAUC run were analysed using SEDFIT (Schuck, 2000) and a plot of molecular weight verses concentration generated (see figure 4.14).

These data show one dominant peak at 63 KDa with a second lower peak at around 174 KDa. This first peak is consistent with either a homodimer of either protein or a heterodimer containing one of each Rv3781c and Rv3783c. Given the data from the SDS PAGE of the sample which shows only Rv3781c present in significant quantities these data would be consistent with a homodimer of Rv3781c which would be 59 KDa.

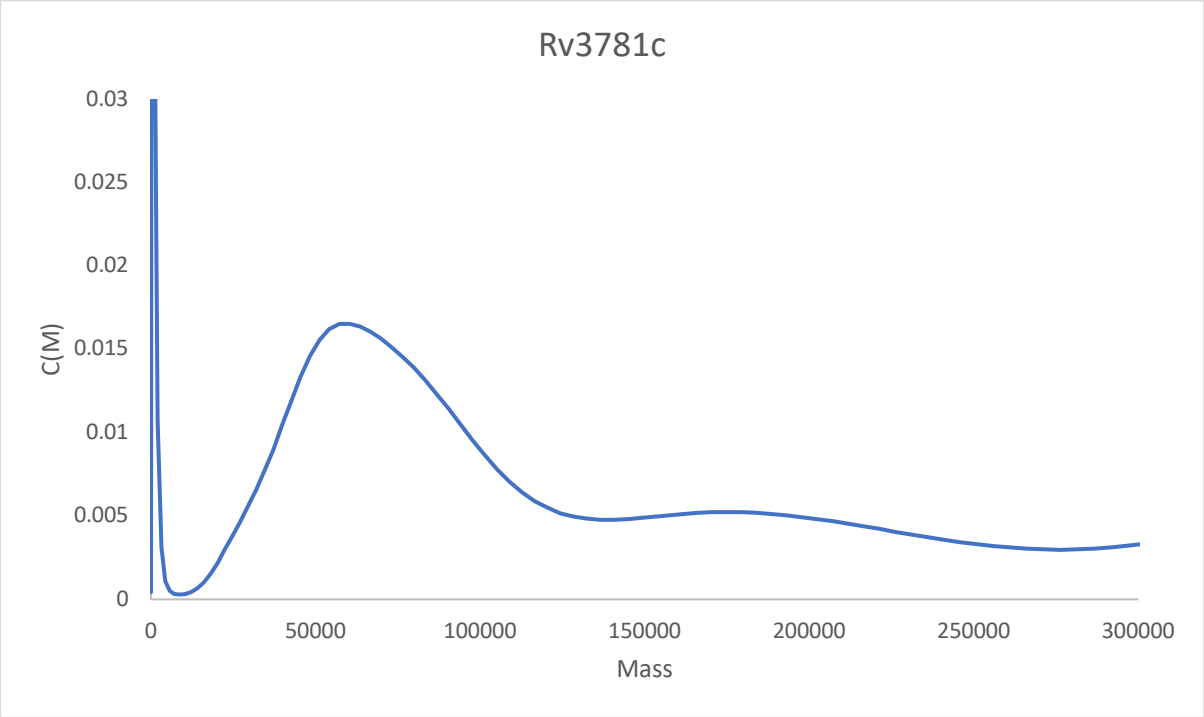


Figure 4.14: AUC Analysis of Rv1273c and Rv1272c SEC fraction

4.4 Discussion

This chapter aimed to examine if the SMALP technique could effectively solubilise the complete *Mycobacterium tuberculosis* ABC-transporter Rv3781c and Rv3783c complex. To purify, Ni-NTA and size exclusion chromatography were used, targeting a protein purity of >80%. Throughout our experimentation, found out that SMA efficiently dissolves proteins in membranes. However, purifying the whole complex with only one tagged protein was difficult, especially ensuring we didn't lose the untagged protein during purification. During the protocol development of Rv1273c and Rv1272c in Chapter 3, similar issues were encountered when aiming to obtain purified protein of Rv3781c and Rv3783c, specifically, the challenge of poor binding to the Ni-NTA, likely stemming from the hindrance caused by SMALP's size in tag binding. Exploring different resins of varying capacities might improve binding and overall purification.

Additionally, considering extra tags like strep or GFP tags at the opposite end could improve specificity. Removing common protein contaminants co-eluting from *E. coli* expression was tricky due to their affinity for nickel, competing with our tagged proteins during binding. Purifying the protein complex in *Mycobacterium smegmatis* might resolve this issue and provide insights into the complex and its surroundings. Furthermore, studying individual proteins apart from the complex could clarify their functions and assist future structural analysis.

Ensuring high-quality, high-yield protein was a significant concern during purification. Low yield hampered detailed structural characterization, limiting advanced methods like small angle X-ray scattering and electron microscopy. Boosting protein yield within SMALP is crucial for obtaining better-resolution data in the future. The data in this chapter came from a 6 L bacterial culture, suggesting scaling up culture volume as a potential solution. However, handling such large volumes of bacterial culture might be time-consuming and challenging for many labs.

CHAPTER 5

SOLUBILIZATION OF *MYCOBACTERUIM TUBERCULOSIS* ABC-TRANSPORTER Rv1218c AND Rv1217c COMPLEX USING STYRENE MALEIC ACID

5.1 Introduction

Rv1218c and Rv1217c are part of an ABC transporter that functions as an exporter, contributing to drug efflux and resistance in *M. tuberculosis* (Cassio Barreto de Oliveira and Balan, 2020). This transporter comprises two nucleotide-binding domains (NBDs) in Rv1218c and a single copy of Rv1217c, which contains two fused membrane-spanning domains (MSDs) (Braibant *et al*., 2000). Studies have shown that Rv1218c facilitates the efflux of various chemicals and antibiotics in *M. smegmatis* (Balganesh *et al*., 2010). Furthermore, research indicates that Rv1217c/Rv1218c are involved in resisting multiple classes of antibiotics such as novobiocins, pyrazolones, biarylpiperazines, bisanilinopyrimidines, pyrroles, and pyridines, contributing to drug resistance in several clinical isolates (Cassio Barreto de Oliveira and Balan, 2020; Day *et al*., 2014). Analysis of these protein using Mycobrowser data displays a molecular weight for the peptide chain of 33.44 kDa for Rv1218c and 56.61 kDa for Rv1217c.

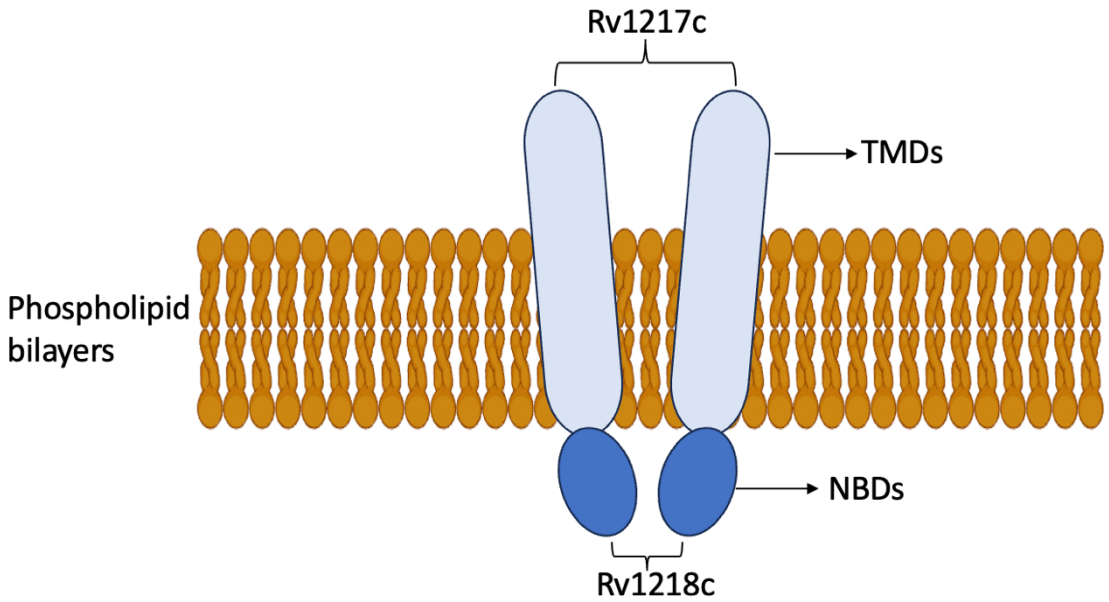


Figure 5.1: Schematic figure of ABC transporters showing the two proteins and the proposed functioning transporter Rv1217c/Rv1218c.

The objective of this project phase was to utilize SMALPs for the extraction of intact complexes involving R1217c and Rv1218c. Plasmids containing both proteins were introduced into *E. coli* BL21 (DE3), and the expression of these proteins was induced using IPTG. In this context, it is proposed that Rv1217c and Rv1218c form a hetero-tetramer, with Rv1218c residing in the membrane and Rv1217c constituting the extracellular domain (Cassio Barreto de Oliveira and Balan, 2020). To facilitate the purification process and verify the successful extraction of the complete complex, a Histidine tag (12 histidine in this case) was added to Rv1218c. The success of this purification is contingent upon the extraction and purification of the entire complex in its intact form. When both the tagged and untagged proteins are present in the final pure sample, it signifies that the SMALP contains the complete complex. It also demonstrates that the complex forms in reality and strongly suggests that the protein acts in nature as a complex. Although it was possible to set up the experiment with both proteins tagged, this approach would not offer conclusive evidence of complex formation. In such a scenario, the presence of both proteins in the final sample might indicate individual proteins in solution rather than a complex.

To obtain highly pure samples, ideally exceeding 80%, for future research, a purification protocol had to be devised. The conventional method for purifying proteins in SMALPs involves a combination of Ni-NTA chromatography followed by SEC (lee *et al*., 2016)

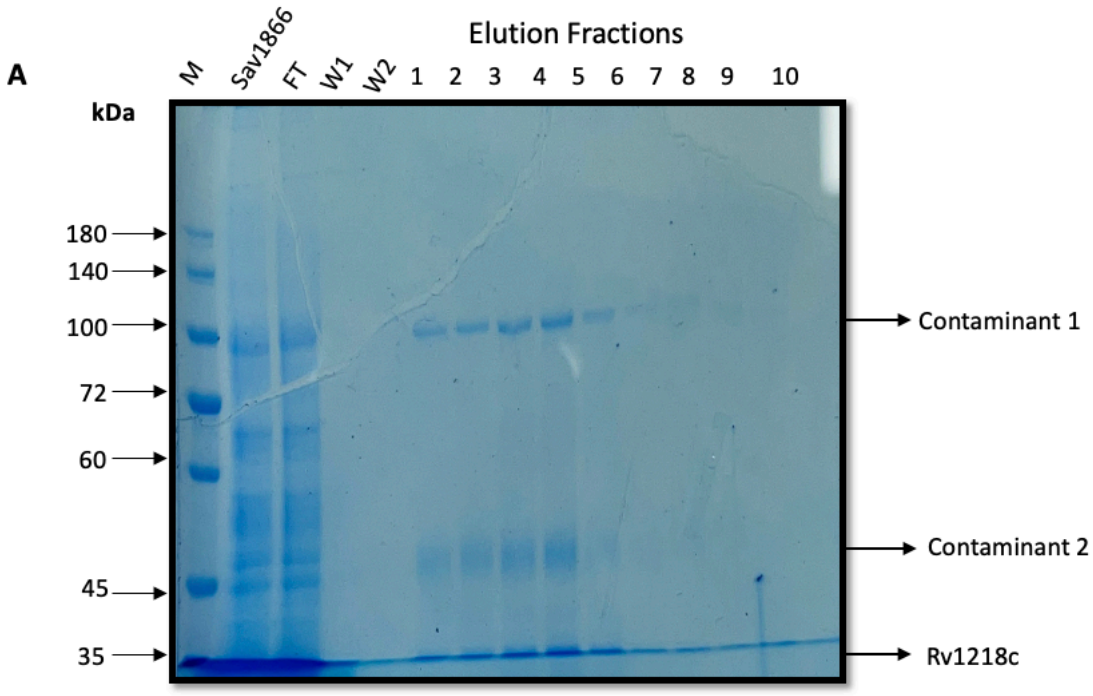
In this chapter, I present a purification trial aimed at achieving a high level of protein purity. This involves the application of both the initial purification protocol and the developed protocols for purifying Rv1272c, Rv1273c, Rv3781c, and Rv3783c.

5.2 Solubilizing Rv1218c/Rv1217c using styrene malic acid co-polymer:

5.2.1 Initial purification conditions (10 mL 20 mM imidazole, 10 mL 500 mM elution buffer pH 8):

Rv1217c and Rv1218c were both overexpressed in the BL21 (DE3) cell strain. Only Rv1218c was tagged with 12 histidine residues. The induction took place overnight when the OD_{600} reached approximately 0.6. Cells were harvested via centrifugation, and the membrane was prepared. During membrane preparation, one tablet of protease inhibitors (EDTA-free) was added. The cells were disrupted using high-pressure homogenization, and membrane fractions were collected after a series of high-speed and ultra-centrifugation steps to remove cell debris and soluble proteins. The membrane fraction was resuspended and combined with SMA, resulting in a final concentration of 2.5%. The insoluble fraction was removed through ultracentrifugation. The SMA-solubilized membrane fraction was bound to prepared Ni-NTA resin overnight on a rotating platform at 4°C to purify Rv1217c and Rv1218c, separating them from other solubilized proteins. The resin was washed four times with a buffer containing free imidazole (Buffer B), 5 ml each time, followed by a single wash with 20 mM low imidazole concentration, using 10 ml to remove non-specifically bound contaminants and collecting the resulting fractions. The His-tagged proteins were eluted with a higher concentration of imidazole (500 mM, pH 8), and the eluted fractions were collected. Subsequently, these eluted fractions, along with the flow-through and the first two washes containing buffer with free imidazole, were analysed using SDS-PAGE (as shown in figure 5.2a). Notably, a significant amount of the solubilized material did not bind to the resin, as demonstrated by the protein bands observed in the flow-through fraction (FT), while the free imidazole washes (W1-2) and the low imidazole concentration wash did not remove or wash away any untagged proteins, as no detectable protein was removed from the column in these stages. In the high-concentration imidazole (500 mM) elution fractions (E1-E10), bands were observed around the 35 kDa mark, which is close to the protein of interest. However, two other proteins with molecular weights around 100 kDa and 45 kDa were also eluted. The SDS-PAGE results were further analysed using a western blot probed using an anti-His tag antibody (as shown in figure 5.2b). The western blot confirmed the

presence of bands around the 35 kDa and the 100 kDa, corresponding to those seen on the Coomassie gel. Additionally, it revealed a band around the 72 kDa mark in the elution fractions (E1-E5) that was not visible on the Coomassie gel. Importantly, the band around 45 kDa was absent on the western blot.



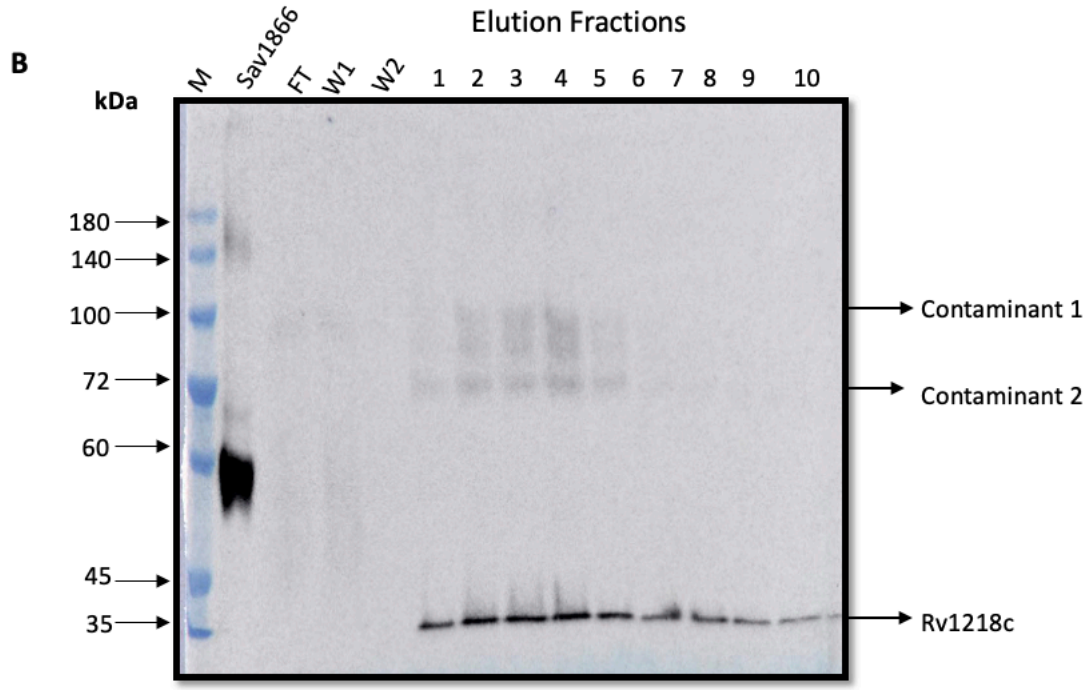


Figure 5.2: R1218c/ Rv1217c Ni-NTA Purification. A) SMA solubilized membrane were incubated with Ni-NTA overnight, then transferred to a gravity flow column, and flow through (FT) collected. Resin then washed 4 times with 5 ml solubilization buffer and collected 4 washes (shown the first two wash W1 and W2), then washed with 10ml of 20 mM imidazole (data not shown). Then His-Tagged proteins were then eluted with 10 ml of 500 mM imidazole buffer (1- 10). Fractions were analysed by electrophoresis on an 10 % handmade SDS-PAGE and visualized by Coomassie stained. B) The SDS-PAGE was also transferred to a nitrocellulose membrane and probed with an anti-His-Tagged antibody**.**

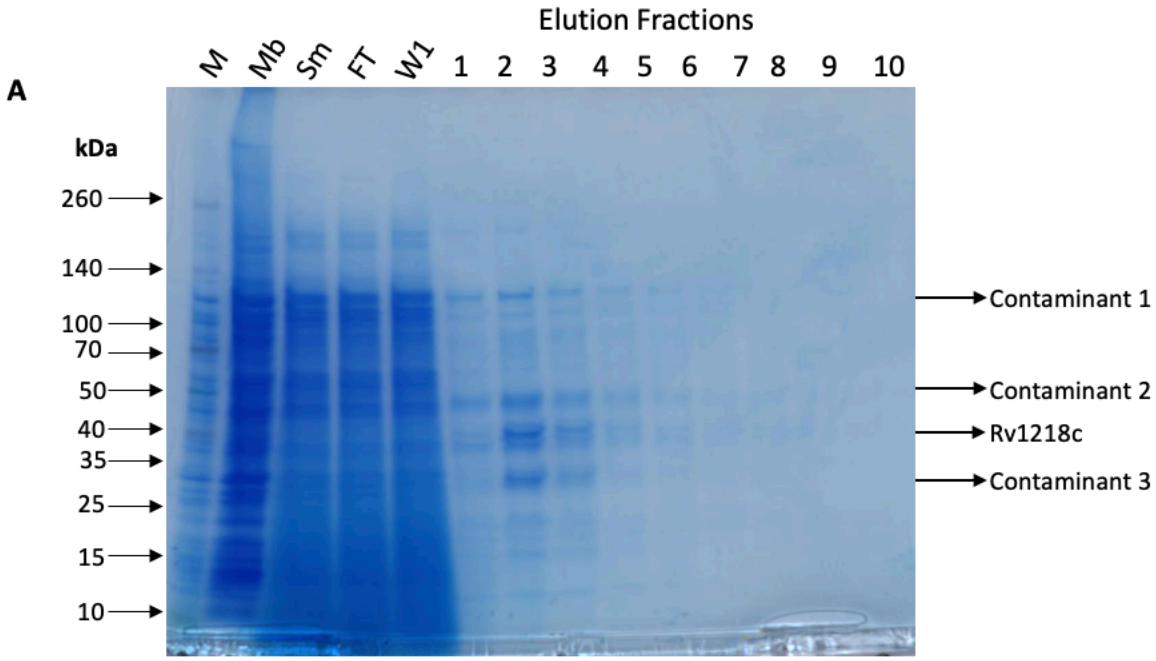
During the initial purification, I observed promising results that demonstrated the ability of SMA to solubilize these proteins. However, the protein yield and level of purity were less than ideal in this early stage. This was evident from the presence of contaminating proteins eluted alongside the protein of interest. Notably, the protein of interest, Rv1218c, migrated very close to the dye front, making it challenging to visualize on a Coomassie gel stain. To address this challenge, I refined the initial protocol by introducing additional wash steps with a low imidazole concentration. This might helped removing more of the contaminating proteins before the elution step with a higher imidazole concentration. Additionally, induced bacterial cells when the OD_{600} falls within the range of 0.6 to 0.8, to boost the growth rate of the cells and, subsequently, enhance the potential production of the protein

of interest. As part of these adjustments, I also optimized the pH level of all buffers containing imidazole to 7.5. This pH adjustment is aimed at improving protein stability and preserving its structural integrity throughout the purification process. Finally, I employed a precast gradient gel to enhance protein separation.

5.2.2 Revised Protocol 1

The purification protocol developed for Rv1273c and Rv1272c was applied to the solubilization and purification of Rv1218c and Rv1217c. Rv1218c and Rv1217c were overexpressed in the BL21 (DE3) cell strain, with only Rv1218c tagged with 12 histidine residues. The induction occurred overnight when the OD_{600} reached a range between 0.6 and 0.8. Cells were harvested via centrifugation, and the membrane was prepared. At the beginning of the membrane preparation, one tablet of EDTA-free protease inhibitors (Mereck) was added. The cells were disrupted using high-pressure homogenization, and membrane fractions were collected after a series of high-speed and ultra-centrifugation steps to remove cell debris and soluble proteins. The membrane fraction was resuspended and combined with SMA, resulting in a final concentration of 2.5%. The SMA-solubilized membrane fraction was bound to prepared Ni-NTA resin overnight on a rotating platform at 4°C as before and the resin was transferred to a gravity flow column, and the flow-through was collected. The resin was washed four times with a buffer free imidazole (Buffer B), 5 ml each time, followed by a single wash with 20 mM and 50 mM low imidazole concentration at pH 7.5, using 10 mL to remove non-specifically bound contaminants, and collecting the resulting fractions. The His-tagged protein was eluted with a higher concentration of imidazole (500 mM, pH 7.5). Eluted fractions were collected and analysed using SDS-PAGE and western blot (as shown in figure 5.3a and b), in addition to examining the membrane (Mb), solubilized membrane (Sm), flow-through (FT), and the first wash with buffer-free imidazole. It is worth noting that again a significant amount of the solubilized material did not bind to the resin, as demonstrated by the flow-through fraction (FT). The free imidazole washes (W1) and the low imidazole concentration wash (not shown) successfully washed away any untagged proteins, with no detectable protein being removed

from the column. Both the flow-through (FT) and the first wash (W1) fractions included a substantial amount of the solubilized material did not bind to the resin. During elution with a high concentration of imidazole (500 mM), four distinct bands were clearly visible in E2 and E3, representing elution proteins at around 120 kDa, 45 kDa, 40 kDa, and 35 kDa. The two bands at around 45 kDa and 40 kDa could still be faintly seen in E4 to E7, albeit slightly less pronounced than in the earlier elution fractions. The western blot detected the His-tagged proteins at the correct molecular weight of Rv1218c, appearing as clear protein bands around 40 kDa across all elution fractions, matching the bands observed around 40 kDa in the SDS-PAGE (E1-E7). Additionally, the western blot revealed faint bands at around 100 kDa in E2 and E3, indicating the likelihood of protein oligomerization.



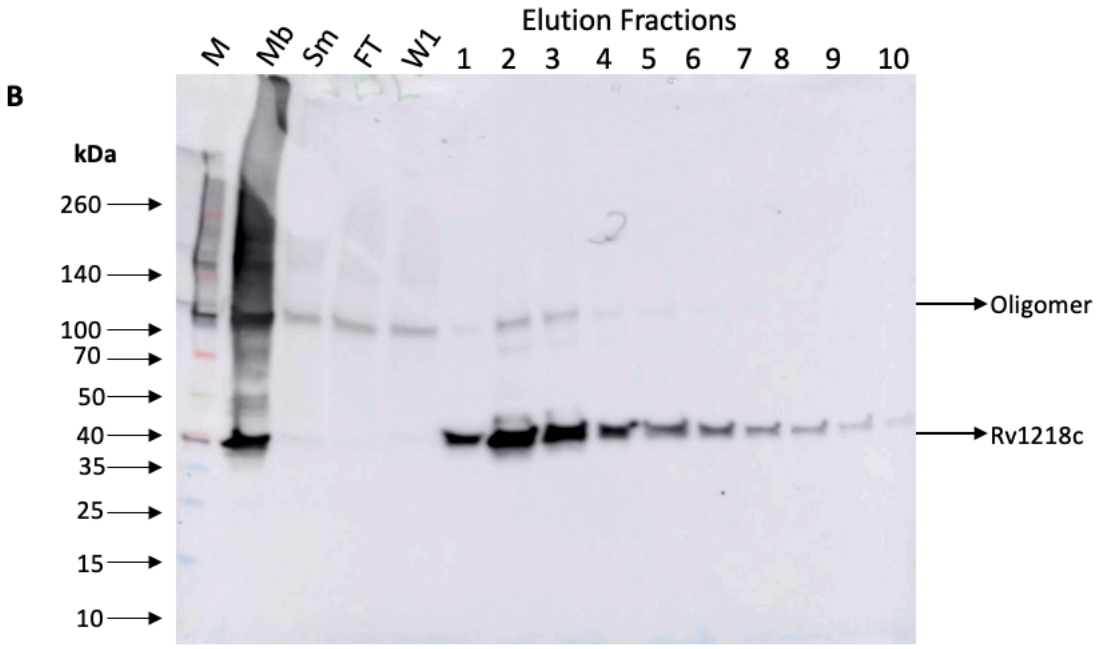
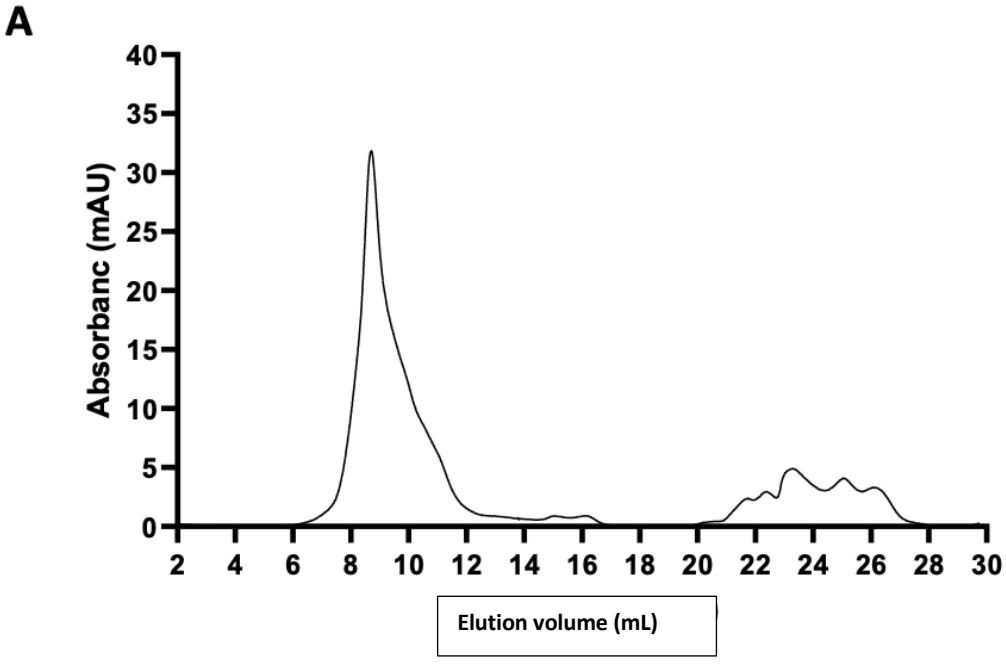


Figure 5.3 : Rv1218c/ Rv1217c Ni-NTA Purification. A) SMA solubilized membrane were incubated with Ni-NTA overnight, then transferred to a gravity flow column, and flow through (FT) collected. Resin then washed 4 times with 5 ml solubilization buffer and collected 4 washes (shown the first wash), then washed with 10 ml of 20 mM and 50 mM imidazole pH 7.5 respectively (data not shown). His-Tagged proteins were then eluted with 20 ml of 500 mM imidazole buffer pH 7.5 (E1- E10). Fractions were analysed by electrophoresis on 4 - 20 % precast SDS- PAGE and visualized by Coomassie stained beside the membrane (Mb), solubilized membrane (Sm). B) The SDS-PAGE was also transferred to a nitrocellulose membrane and probed with an anti-His-Tagged antibody**.**

Implementing the modifications to this protocol has yielded promising results for protein purification. An increase in target protein yield was observed compared to the initial protocol outlined in section 5.2.1. However, the increase in protein yield has had an impact on protein purity, as evidenced by the Coomassie gel staining. Utilizing the Ni-NTA purification protocol developed for Rv1273c and Rv1272c has demonstrated an overall improvement in the purification process. Based on these observations, it was decided that size exclusion chromatography should be included post Ni-NTA chromatography to enhance both the yield and protein purity.

5.2.2.1 Size Exclusion Chromatography

In the next step of purification, the elution fractions were combined and concentrated using a 3,000 MWCO concentrator. Subsequently, the sample was subjected to separation through a Superdex 200 Increase 10/300 size exclusion column, as depicted in figure 5.4a. This figure illustrates the elution profile of the sample at 280 nm, indicating the elution of proteins. The elution occurs within a moderate range of elution volumes, featuring one major peak and two shoulder peaks between 8 and 13 mL. To further assess the fractions within this profile, we conducted SDS-PAGE and Western blotting, as shown in figure 5.4b and 5.4c. Through Coomassie stain gel and, especially, Western blot analysis, confirmed that Rv1217c and Rv1218c eluted between 8.0 mL and 10.5 mL, with the His-Tag of Rv1218c being detected across the elution range. The Coomassie stain gel reveals faint bands between 8.0 – 9.5 mL at the appropriate molecular weight of Rv1218c, along with two faint contaminant bands at approximately 120 kDa and below 50 kDa. The Western blot analysis shows a band for Rv1218c, appearing around 40 kDa within the elution range of 8 to 10.5 mL, and also detects two bands at approximately 140 kDa and 100 kDa. These protein bands are likely indicative of protein oligomerization.



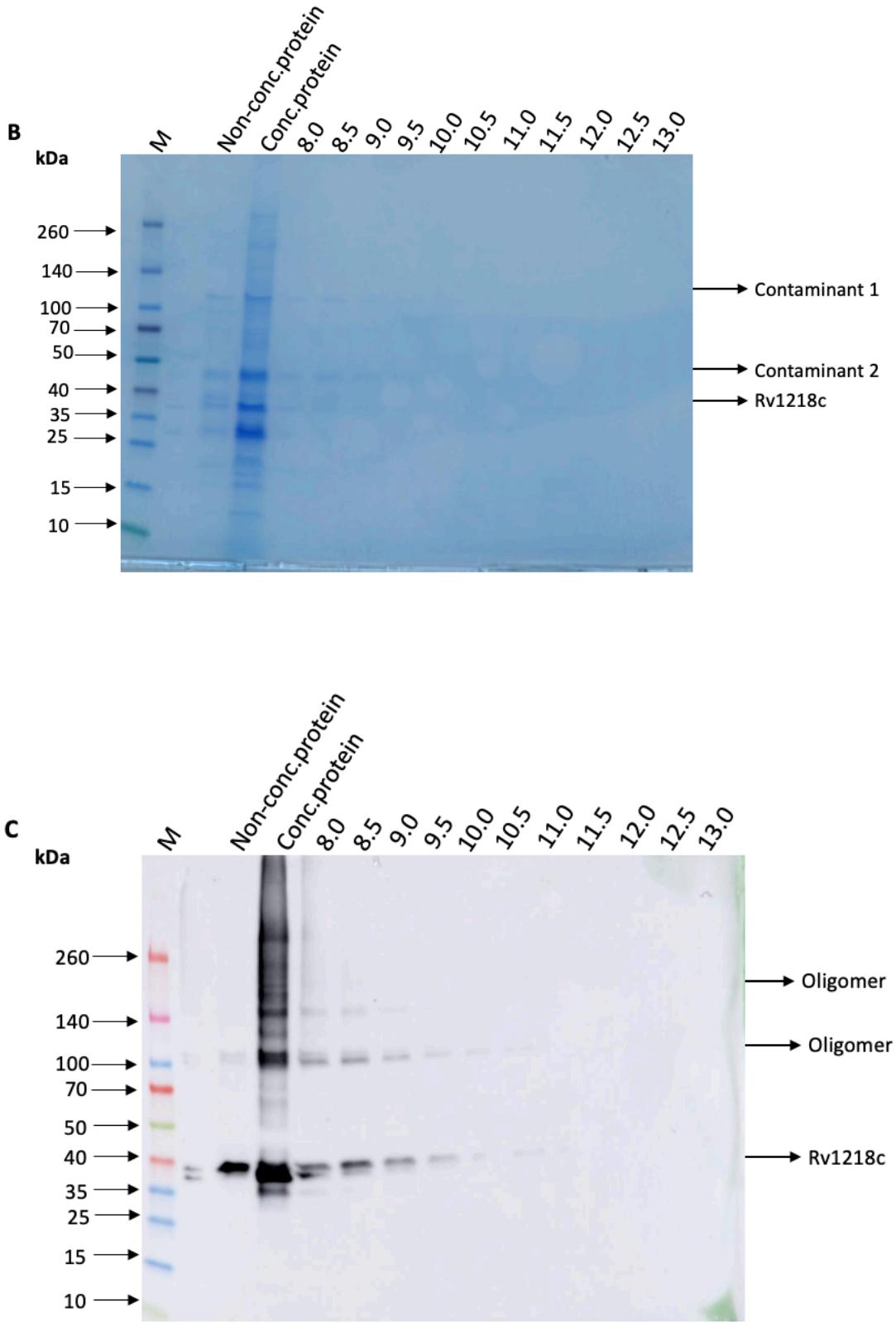


Figure 5.4:SEC purification of Rv1218c/ Rv1217c. Elution fraction fractions from Ni- NTA purification were collected, concentrated, and injected onto a Superdex 200 Increase 10/300 GL column. The column was run at 0.5 mL/min collecting 0.5 mL fractions. A) SEC elution profile. Elution of protein from the column was

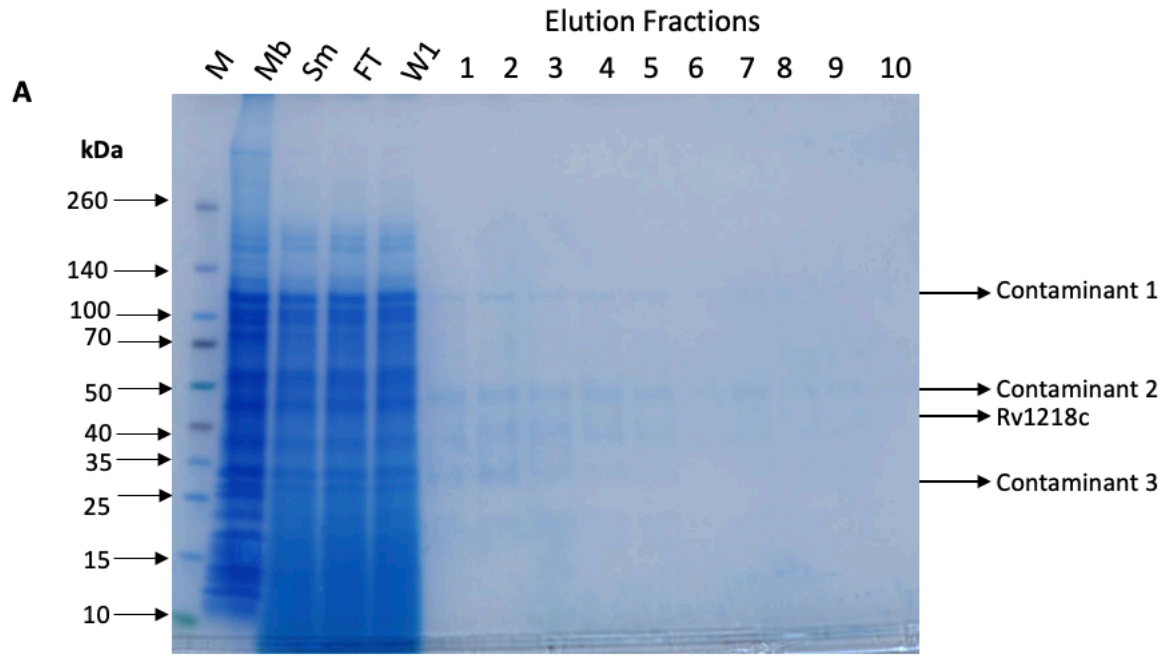
monitored by absorbance at 280 nm. B) SDS-PAGE of elution profile 4 -20 % Precast SDS-PAGE of the SEC elution fractions analysed by electrophoresis and protein bands visualized by Coomassie staining. C) The SDS-PAGE was also transferred to a nitrocellulose membrane and probed with an anti-His-Tagged antibody**.** Concentrated protein (conc. protein), non-concentrated protein (Non-conc. protein).

The size exclusion chromatography profile has shown a single major peak with a low absorbance rate. Upon analysing the fractions along this profile, there has been no significant improvement in protein yield and purity. It appears that a substantial amount of protein may have been lost during this process. To address these issues the purification plan was further improved by adopting aspects of the developed protocol used for Rv3781c and Rv3783c. This protocol includes additional steps with low imidazole concentrations, which may aid in removing as many contaminating proteins as observed in this purification trial.

5.2.3 Revised Protocol 2

The purification protocol originally designed for Rv3781c and Rv3783c to solubilize and purify Rv1218c and Rv1217c was implemented with the aim of enhancing protein yield through a 20:1 resin binding ratio. Rv1218c and Rv1217c were overexpressed in the BL21 (DE3) cell strain, with only Rv1218c tagged with 12 histidine residues. The SMA-solubilized membrane fraction was bound to prepared Ni-NTA resin overnight at a 20:1 resin-to-protein ratio on a rotating platform at 4°C to purify Rv1217c and Rv1218c, segregating them from other solubilized proteins. The resin was washed four times with a buffer containing free imidazole (Buffer B), 5 ml each time, followed by low imidazole concentration washes of 20 mM, 50 mM, and 100 mM at pH 7.5, each conducted once with 5 mL to eliminate non-specifically bound contaminants. Elution of the His-tagged protein occurred at a higher concentration of imidazole (500 mM, pH 7.5). The eluted fractions were collected and analysed using SDS-PAGE and western blot (as shown in figure 5.5a and b), in addition to the examination of the membrane (Mb), solubilized membrane (Sm), flow-through (FT), and the

first wash with buffer-free imidazole. It is evident again that a substantial amount of the solubilized material did not bind to the resin, as demonstrated by the flow-through fraction (FT). The free imidazole washes (W1) and the low imidazole concentration wash (not shown) effectively removed any untagged proteins, with no detectable protein being removed from the column. Both the flow-through (FT) and the first wash (W1) highlight the considerable amount of solubilized material that did not bind to the resin. In the elution with a high concentration of imidazole (500 mM), faint bands appeared in the first five elution fractions around 120 kDa, 50 kDa, 45 kDa, and 30 kDa. Western blot analysis detected the His-tagged proteins at the correct molecular weight of Rv1218c, with clear protein bands around 40 kDa across all elution fractions, matching the bands observed around 40 kDa in the SDS-PAGE (E1-E5). Notably, despite their presence in the Coomassie gel, the western blot did not detect the bands around 120 kDa, 45 kDa, and 30 kDa.



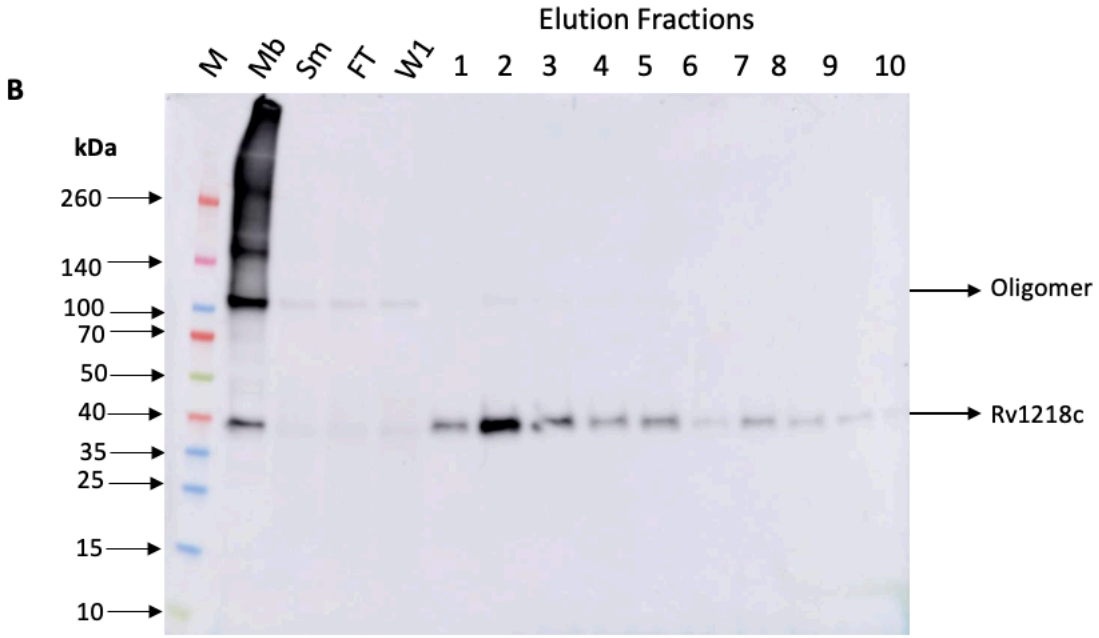
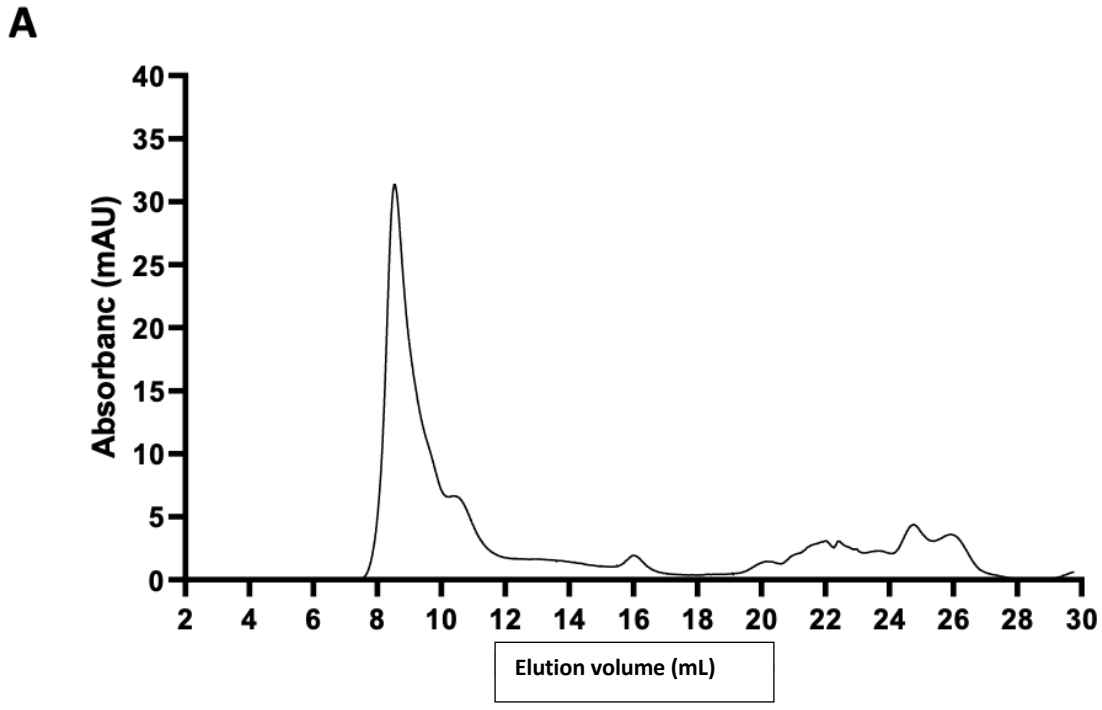


Figure 5.5: Rv1218c/ Rv1217c Ni-NTA Purification. A) SMA solubilized membrane were incubated with Ni-NTA overnight, then transferred to a gravity flow column, and flow through (FT) collected. Resin then washed 4 times with 5 ml solubilization buffer and collected 4 washes (shown the first wash), then washed with 5 ml of 10 mM, 50 mM, and 100 mM imidazole pH 7.5 respectively (data not shown). His-Tagged proteins were then eluted with 20 ml of 500 mM imidazole buffer pH 7.5 (E1- E10). Fractions were analysed by gel electrophoresis on 4 - 20 % precast SDS- PAGE and visualized by Coomassie stained beside the membrane (Mb), solubilized membrane (Sm). B) The SDS-PAGE was also transferred to a nitrocellulose membrane and probed with an anti-His-Tagged antibody**.**

The Ni-NTA purification exhibited a pattern of protein bands similar to the one performed in section (5.2.2.). However, both protein yield and purity fell below the expected levels. The Coomassie gel staining revealed faint bands of contaminants alongside the protein of interest, which was also detected in the western blot. Despite the reduced yield the protein was prepared for Size Exclusion Chromatography.

5.2.3.1 Size Exclusion Chromatography

In the subsequent purification step, the elution fractions were pooled and concentrated using a 3,000 MWCO concentrator. The sample was then subjected to separation through a Superdex 200 Increase 10/300 size exclusion column. As depicted in Figure 5.6a, this figure illustrates the elution profile of the sample at 280 nm, which indicates protein elution. The elution process occurs over a moderate range of elution volumes, featuring a primary peak and two additional shoulder peaks between 8.0 and 12.5 mL. To further evaluate the fractions within this profile, we conducted SDS-PAGE and Western blotting, as illustrated in Figure 5.6b and 5.6c. Through Coomassie stain gel and, particularly, Western blot analysis, we have substantiated that Rv1218c and Rv1217c were eluted between 8.0 mL and 12.5 mL, as evidenced by the detection of the His-Tag of Rv1218c throughout the elution range. The Coomassie stain gel reveals faint bands between 8.0 – 10.5 mL at the appropriate molecular weight of Rv1218c, along with two faint contaminant bands at approximately 120 kDa and below 50 kDa. The Western blot analysis showcases a band for Rv1218c, present between 8.0 and 12.5 mL and detecting two additional bands around 140 kDa and 100 kDa, observed between 8.0 and 10.0 mL. These protein bands are likely indicative of protein oligomerization.



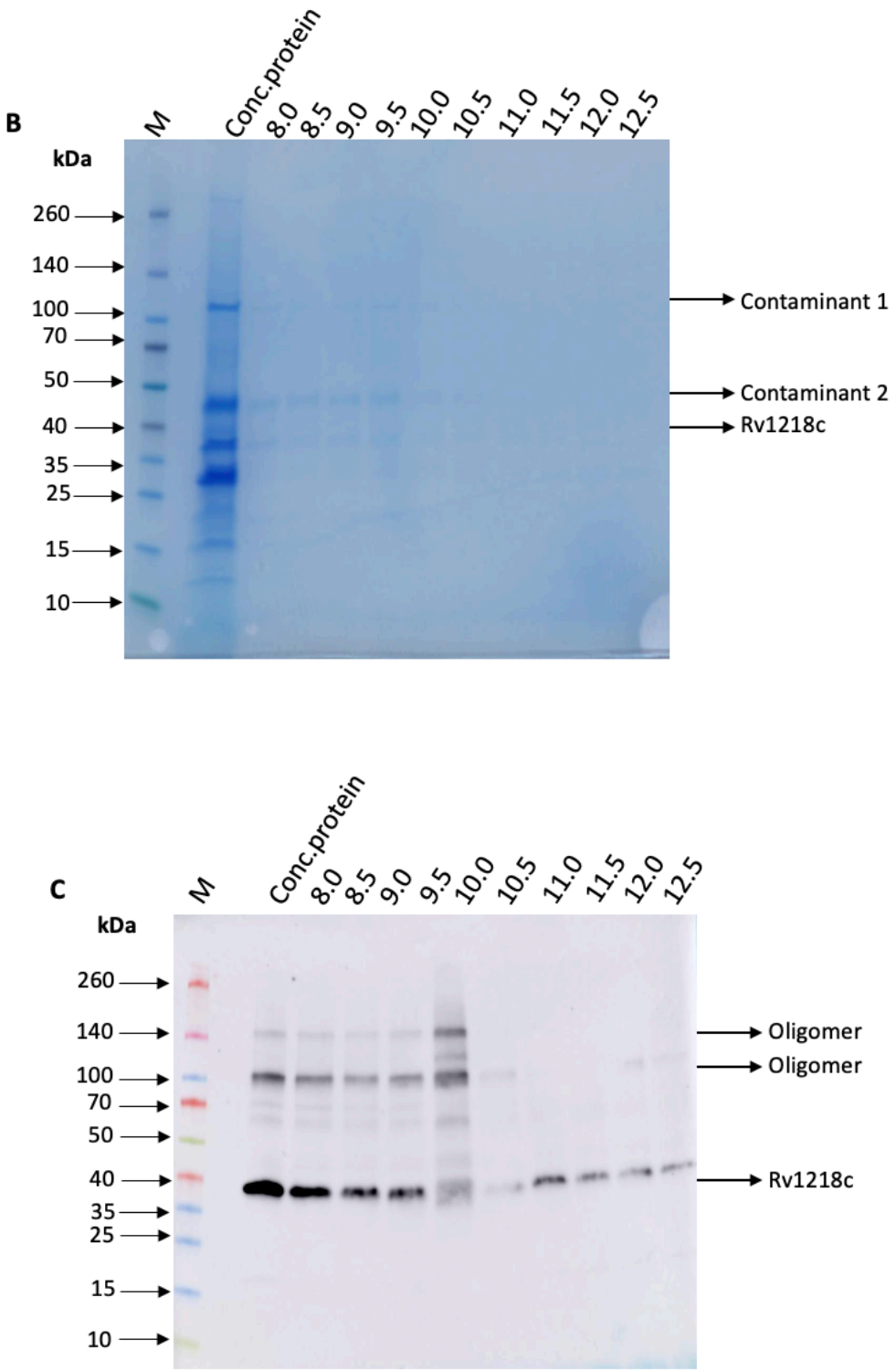


Figure 5.6: SEC purification of Rv1218c/ Rv1217c Elution fraction fractions from Ni- NTA purification were collected, concentrated, and injected onto a Superdex 200 Increase 10/300 GL column. The column was run at 0.5 mL/min collecting 0.5 mL fractions. A) SEC elution profile. Elution of protein from the column was monitored by absorbance at 280 nm. B) SDS-PAGE of elution profile 4 -20 % Precast SDS-PAGE of the SEC elution fractions analysed by electrophoresis and protein bands visualized by Coomassie staining. C) The SDS-

PAGE was also transferred to a nitrocellulose membrane and probed with an anti-His-Tagged antibody. Concentrated protein (conc. protein).

The size exclusion profile revealed one major peak with a small shoulder peak, both displaying low absorbance. However, upon analysing the fractions along the profile, there wasn't a significant improvement in protein yield or purity, which is evident on the Coomassie-stained gel. Conversely, the western blot detected Rv1218c alongside several other proteins, which may indicate the presence of oligomers containing Rv1218c. Unfortunately, at this point, the purification was not successful. In attempt to improve purification performance the same protocol was repeated with a slight adjustment: changing the ratio of Ni-NTA resin binding to the membrane to 1 ml of resin per 10 ml of membrane.

5.2.4 Revised Protocol 3

Rv1218c and Rv1217c were expressed, solubilized, and purified using the same protocol detailed in section (5.2.3). In this variation of the protocol, we adjusted the binding ratio of membrane to resin to 10:1. Eluted fractions were collected and subsequently analysed through SDS-PAGE and Western blot (as shown in figure 5.7a and 5.7b), along with an assessment of the flow-through (FT). Notably, a substantial amount of the solubilized material did not adhere to the resin, as evident in the flow-through fraction (FT). The free imidazole washes, as well as the low imidazole concentration wash (not shown), effectively removed any untagged proteins, with no detectable protein being removed from the column. During the elution step with a high concentration of imidazole (500 mM), faint bands emerged in the first three elution fractions (E1-E3) at approximately 120 kDa, 45 kDa, 40 kDa, and 30 kDa. The Western blot detected the His-Tagged proteins at the correct molecular weight of Rv1218c, manifesting as clear protein bands around 40 kDa across the entire elution fractions, with a notable presence in E1, E2, and E3, aligning with the bands around 40 kDa in the SDS-PAGE (E1-E3). Importantly, despite their appearance in the Coomassie gel, the Western blot did not detect the bands around 120 kDa, 45 kDa, and 30 kDa.

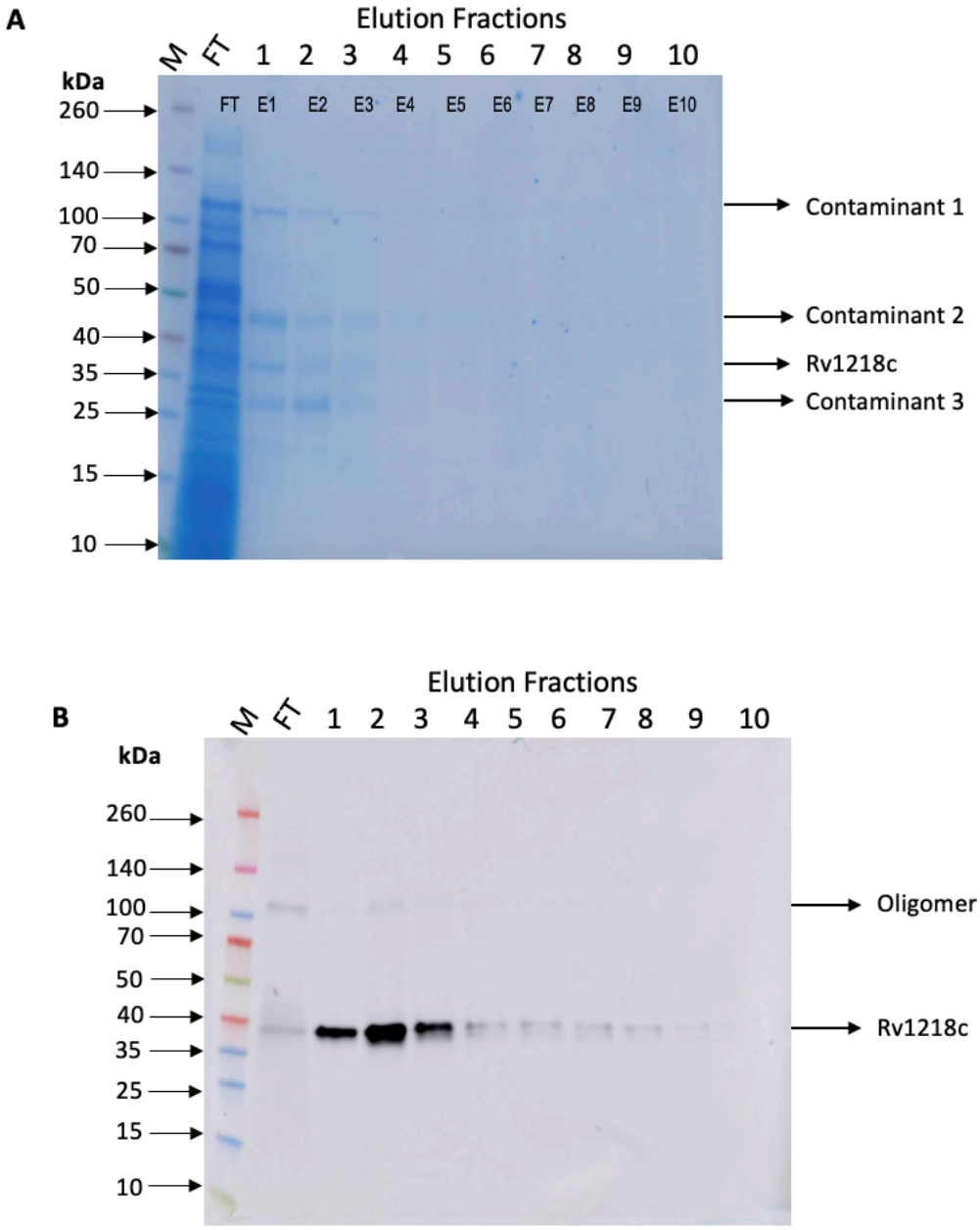
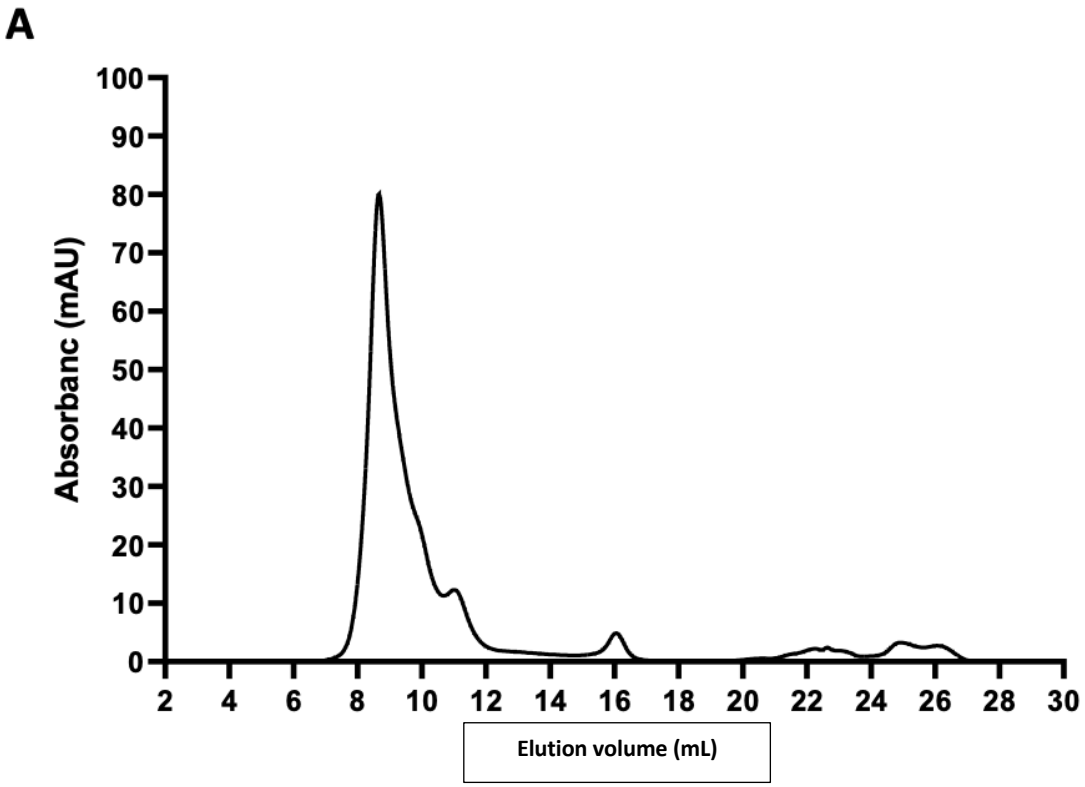


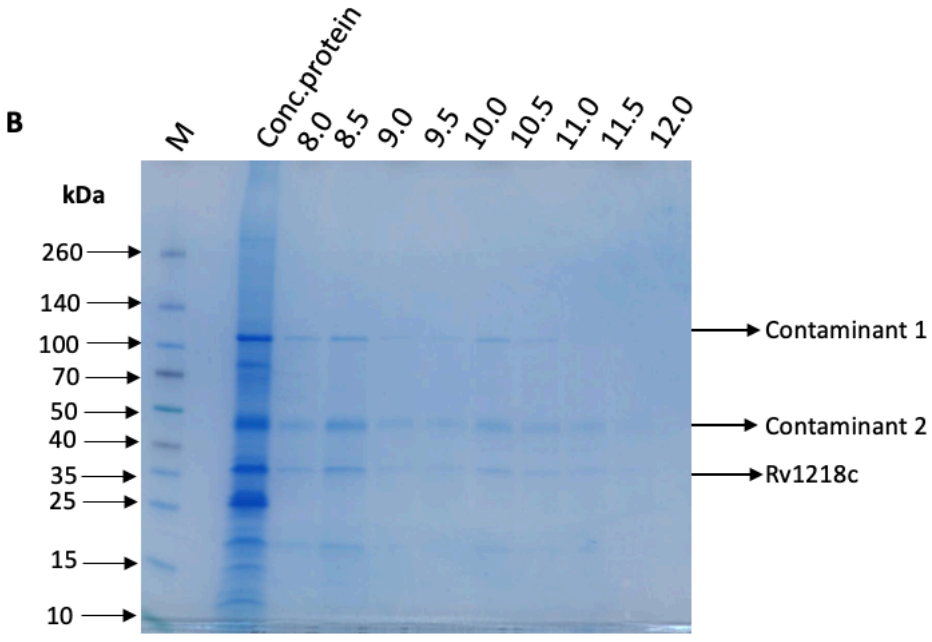
Figure 5.7: Rv1218c/ Rv1217c Ni-NTA Purification. A) SMA solubilized membrane were incubated with Ni-NTA overnight, then transferred to a gravity flow column, and flow through (FT) collected. Resin then washed 4 times with 5 ml solubilization buffer and collected 4 washes (shown the first wash), then washed with 5 ml of 10 mM, 50 mM, and 100 mM imidazole pH 7.5 respectively (data not shown). His-Tagged proteins were then eluted with 20 ml of 500 mM imidazole buffer pH 7.5 (E1- E10). Fractions were analysed by gel electrophoresis on 4 - 20 % precast SDS- PAGE and visualized by Coomassie stained. B) The SDS-PAGE was also transferred to a nitrocellulose membrane and probed with an anti-His-Tagged antibody.

The minor adjustment to the protocol has led to a slight improvement in the purification process. There is an increase in protein yield compared to the one described in section (5.2.3), although it remains unsatisfactory. Unfortunately, achieving a high level of protein purity is still challenging at this stage, even with the adjustments made in the earlier preparations. To enhance sample purity the material was subjected to size exclusion chromatography as per the previous purifications.

5.2.4.1 Size Exclusion Chromatography

After concentration using a 3,000 MWCO concentrator the sample was then subjected to separation through a Superdex 200 Increase 10/300 size exclusion column. As depicted in figure 5.8a, this figure illustrates the elution profile of the sample at 280 nm, indicating protein elution. The elution process occurs over a moderate range of elution volumes, featuring a primary peak and two additional shoulder peaks between 8.0 and 12.0 mL. To further evaluate the fractions within this profile, we conducted SDS-PAGE and Western blotting, as demonstrated in Figure 5.8b and 5.8c. Through Coomassie stain gel and, in particular, Western blot analysis have confirmed that Rv1218c were eluted between 8.0 mL and 12.0 mL. This was achieved by detecting the His-Tag of Rv1218c throughout the elution range. The Coomassie stain gel reveals faint bands within the range of 8.0 – 12.0 mL, corresponding to the correct molecular weight of Rv1218c, along with two faint contaminant bands at approximately 120 kDa and below 50 kDa. The Western blot analysis depicts a band for Rv1218c within the same elution range of 8.0 – 12.0 mL, combined with two additional bands around 140 kDa and 100 kDa, which were not observed in the Coomassie stain gel. These protein bands likely indicate protein oligomerization.


A
Absorbanc (mAU)
100
90
80
70
60
50
40
30
20
10
0
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
Elution volume (mL)


B
M
Conc.protein
8.0
8.5
9.0
9.5
10.0
10.5
11.0
11.5
12.0
kDa
260
140
100
70
50
40
35
25
15
10
Contaminant 1
Contaminant 2
Rv1218c

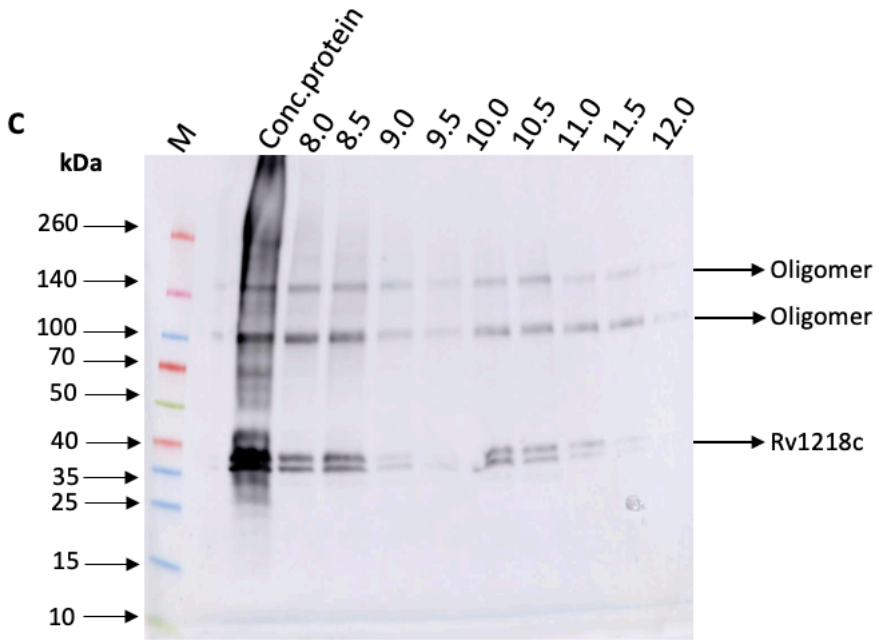


Figure 5.8: SEC purification of Rv1218c/ Rv1217c Elution fraction fractions from Ni- NTA purification were collected, concentrated, and injected onto a Superdex 200 Increase 10/300 GL column. The column was run at 0.5 mL/min collecting 0.5 mL fractions. A) SEC elution profile. Elution of protein from the column was monitored by absorbance at 280 nm. B) SDS-PAGE of elution profile 4 -20 % Precast SDS-PAGE of the SEC elution fractions analysed by electrophoresis and protein bands visualized by Coomassie staining. C) The SDS-PAGE was also transferred to a nitrocellulose membrane and probed with an anti-His-Tagged antibody. Concentrated protein (conc. protein).

The size exclusion profile displayed one major peak with a small shoulder peak, showing a slight increase in absorbance. However, upon analysing the fractions along the profile, there wasn't a significant improvement in protein purity, which is evident from the Coomassie-stained gel. The western blot exhibited a pattern of bands similar to the one in section (5.2.3.1). Overall, the purification of Rv1217c and Rv1218c may not have been successful. However, it does indicate a promising potential for future purification via SMALPs, with the potential to achieve a high percentage of purity and yield.

5.3 CHARACTERISATION OF THE COMPONENTS OF Rv1218c and Rv1217c IN STYRENE MALEIC ACID LIPID PARTICLES

5.3.1 Circular Dichroism (CD) of SMA solubilised Rv1218c and Rv1217c:

Circular dichroism is a technique used to analyse how molecules absorb different types of circularly polarized light. It is particularly valuable in deciphering protein structures due to the distinctive patterns that emerge during specific transitions in the peptide backbone. For instance, observing the circular dichroism spectrum of an α-helix reveals two negative points of reduced absorption at 208 nm and 222 nm. β-sheet, however, present less straightforward results because their structures exhibit more variation. Typically, they exhibit negative absorption around 216 nm.

CD data was collected for the sample corresponding to fraction in figure 5.9. This fraction contained Rv1218c which meant that the CD data would provide information on the secondary structure content of that protein. The CD data shows a clear negative peak at the 222 nm and significant negative peak around 208 nm and including 216 nm. This indicates that the sample contains of β-sheet (216 nm). Since the NBD is the more conserved part of the protein and primarily consists of alpha helices, the CD results suggesting a strong beta-sheet signal may indicate contamination by another protein rich in β-sheet which is dominating the spectrum. To further assess these data the spectrum was deconvolved using Betsel. The Betsel analysis predicted that the protein contains 0.0% Alpha, 45.2% Beta and 54.8% loop/unfolded. Figure 5.10 shows the CD data and the result of the fit from Betsel showing that there is a good agreement between the two.

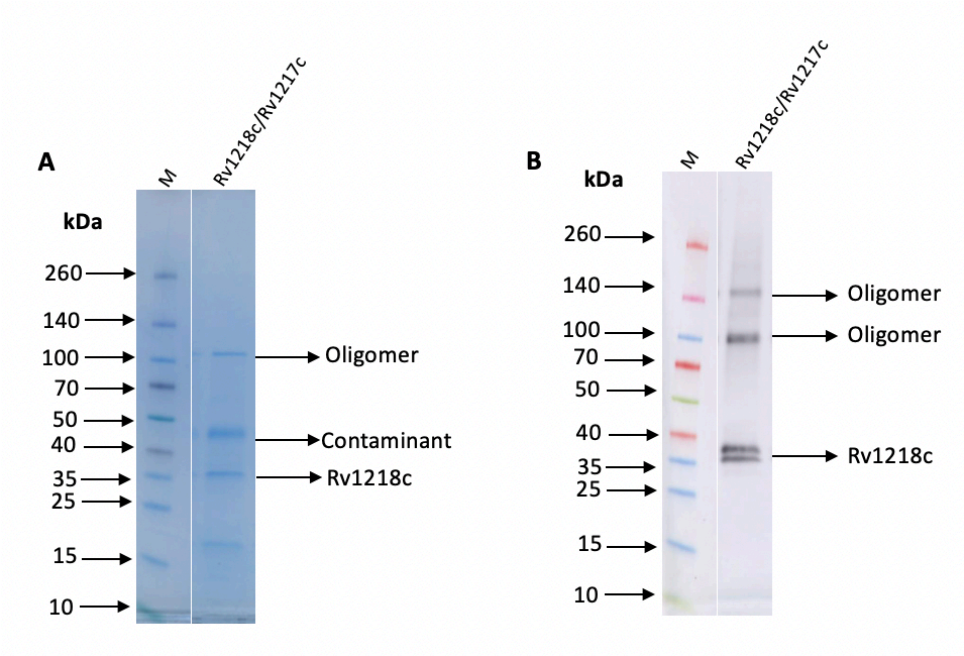


Figure 5.9: SDS-PAGE (A)and Western blot (B)of Rv1218c and R1217c fraction after SEC purification

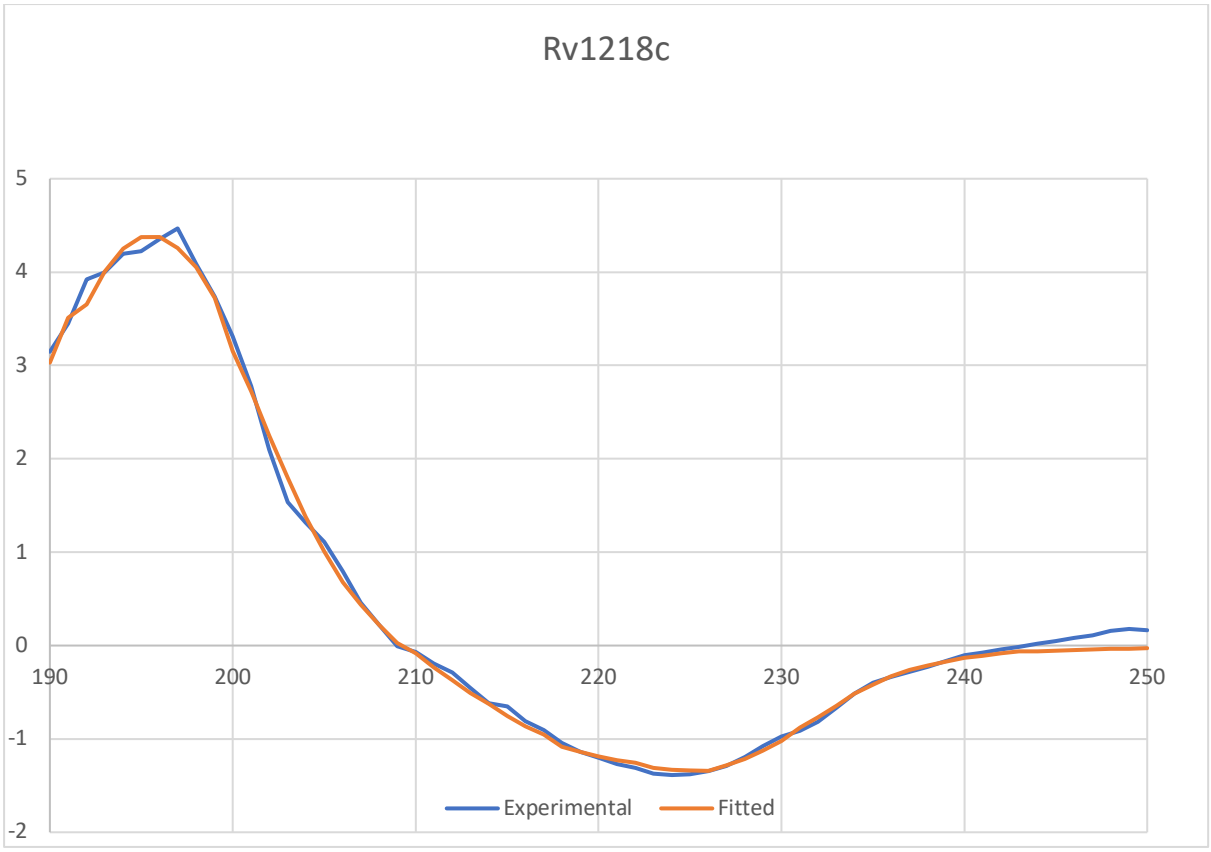


Figure 5.10: CD analysis of Rv1218c on SMALPs. The fractions obtained after the size-exclusion chromatography (SEC) purification of Rv1218c were examined using CD spectroscopy to evaluate its secondary structure. The analysis was performed using a JASCO J-1500 Circular Dichroism Spectrophotometer operated through Spectra Manager software. Measurements were conducted in a quartz glass cuvette with a pathlength of 1 mm, using a protein concentration of approximately 0.1 mg/mL.

5.3.2 Analytical Ultracentrifugation (AUC)

A sample was taken from fraction on figure 5.9 and concentrated before being analysed using Sedimentation Velocity Analytical Ultracentrifugation (svAUC). svAUC provides a measure of the molecular mass of proteins in solution. These data from the svAUC run were analysed using Sedfit (Schuck, 2000) and a plot of molecular weight verses concentration generated (see figure 5.11).

These data show one dominant peak at 290 KDa with a second lower peak at around 32 kDa and third small one around the 102 kDa. This first peak is consistent of an oligomer of Rv1217c and Rv1218c. The second peak around the 32 kDa is a monomer of either protein. Given the data from the SDS PAGE of the sample which shows only Rv1218c present in significant quantities these data would be consistent with a tetramer of Rv1218c which

would be 102 kDa.

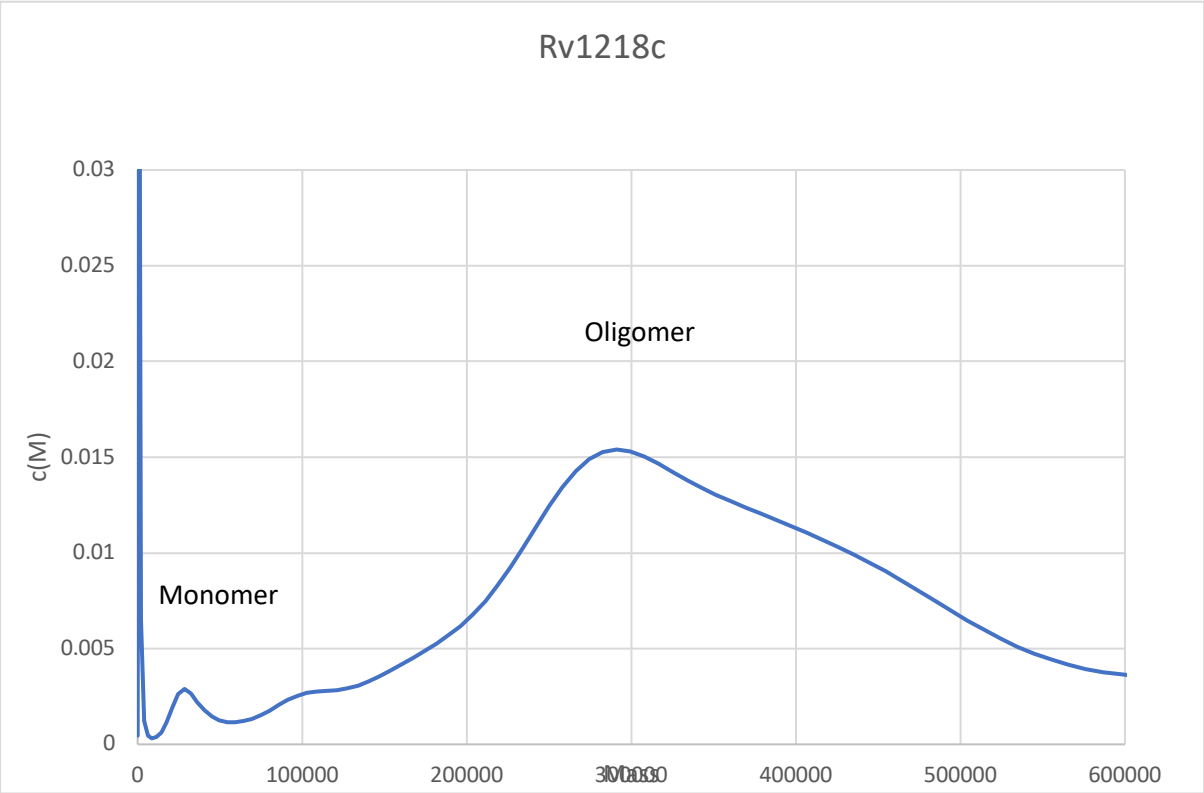


Figure 5.11: AUC Analysis of Rv1218c and Rv1217c SEC fraction

5.4 Discussion

The objective of this chapter was to assess the SMALP method's capability in solubilizing the intact complex of Mycobacterium tuberculosis ABC-transporter Rv1218c and Rv1217c. We employed Ni-NTA and size exclusion chromatography aiming for over 80% purified protein as our benchmark. We followed similar protocols used for RV1273c/72c and Rv3781c/83c.

The outcomes, particularly with Rv3781c and Rv3783c, demonstrated that SMA effectively solubilizes membrane proteins. However, similar to previous purifications, dealing with the complexity of a single tagged protein posed challenges in ensuring we didn't lose the untagged protein during purification. Despite successful solubilization, we encountered issues with inefficient binding to the Ni-NTA, possibly linked to the nature of SMALP.

One potential solution to overcome this issue is to explore different resins with varying capacities, potentially improving binding and overall purification. Additionally, considering additional tags like strep or Green Fluorescent Protein (GFP) tags at the opposite terminus might enhance specificity.

We also faced challenges with common protein contaminants co-eluting with tagged proteins from *E. coli expression*, which had a high affinity for nickel, creating a competitive binding environment. To address these challenges, investigating protein complex purification in *Mycobacterium smegmatis* could provide deeper insights. Furthermore, studying individual proteins separately from the complex might aid in determining their functions and understanding their structure in the future.

Ensuring high-quality, high-yield protein during purification remained a major concern. Low protein yield affected our ability to establish good structural characterization data, limiting the application of higher resolution methods. Scaling up culture volume might be a solution, although it could be time-consuming and challenging for many laboratories. Exploring alternative polymers for protein purification might also provide valuable insights into protein behaviour and surrounding lipids. This could contribute to a better understanding of protein-lipid interactions in similar systems.

CHAPTER 6

DEVELOPING NEW METHOD OF LIPIDS EXTRACTION FROM *MYCOBACTERIA TUBERCULOSIS* VIA STYRENE MALEIC ACID

6.1 Introduction

The cell envelope of *Mycobacterium tuberculosis*, shown in figure 6.1, is a complex structure comprised of glycopolymers, lipids, glycolipids, proteins, and mycolyl-arabinogalactan-peptidoglycan complexes, making it a highly pathogenic organism (Pushkaran et al., 2019; Chiaradia et al., 2017). Polysaccharides like α-D-glucan, trehalose 6,6'-dimycolate (TDM), lipomannan (LM), mannose-capped lipoarabinomannan (Man-LAM), and phosphatidylinositol mannosides (PIMs) are key components of its cell wall, forming a glycopolymer, lipid, and lipoglycan matrix (Vinod et al., 2020). These complex structures play a vital role in modulating the host immune response during infection. For instance, α-D-glucan, TDM, LM, Man-LAM, and PIMs interact with host receptors, such as C type lectin receptor mincle, scavenger receptor A (SRA), MARCO, CD1b receptor on T cells, mannose receptor (MR), producing various immunomodulatory effects (Richardson and Williams, 2014; Bowdish et al., 2009; Vinod et al., 2020).

The intricate structure of the *Mycobacterium tuberculosis* cell envelope presents an ideal target for developing new therapeutics due to its role in pathogenicity, virulence, and inherent drug resistance. However, current analytical methods relying on extraction using harsh organic solvents lacks the ability to selectively extract outer cell envelope components. Amphipathic polymers, like Styrene Maleic Acid (SMA), have been widely used to directly solubilize proteins and lipids from the cytoplasmic membrane via SMA Lipid Particles (SMALPs) and may provide a route to a more effective extraction of *Mtb* membranes.

The aim in this chapter is to employ SMALPs to selectively extract peripheral molecules from mycobacterial cell envelopes. Using [^{14}C]-uniformly labelled lipids, intending to investigate the potential of SMA to extract complex glycolipids (trehalose monomycolate, trehalose dimycolate, and glucose monomycolate), phospholipids (phosphatidyl inositol mannosides), lipoglycans (lipomannan and lipoarabinomannan), and surface-exposed proteins from the envelopes of *Mycobacterium smegmatis*, while avoiding cell rupture.

It should be noted that the work presented in this chapter is an initial investigation which was disrupted by the COVID pandemic meaning that in some cases results are not fully complete. These data are included nonetheless as they show that the method has good potential and further investigations should be continued.

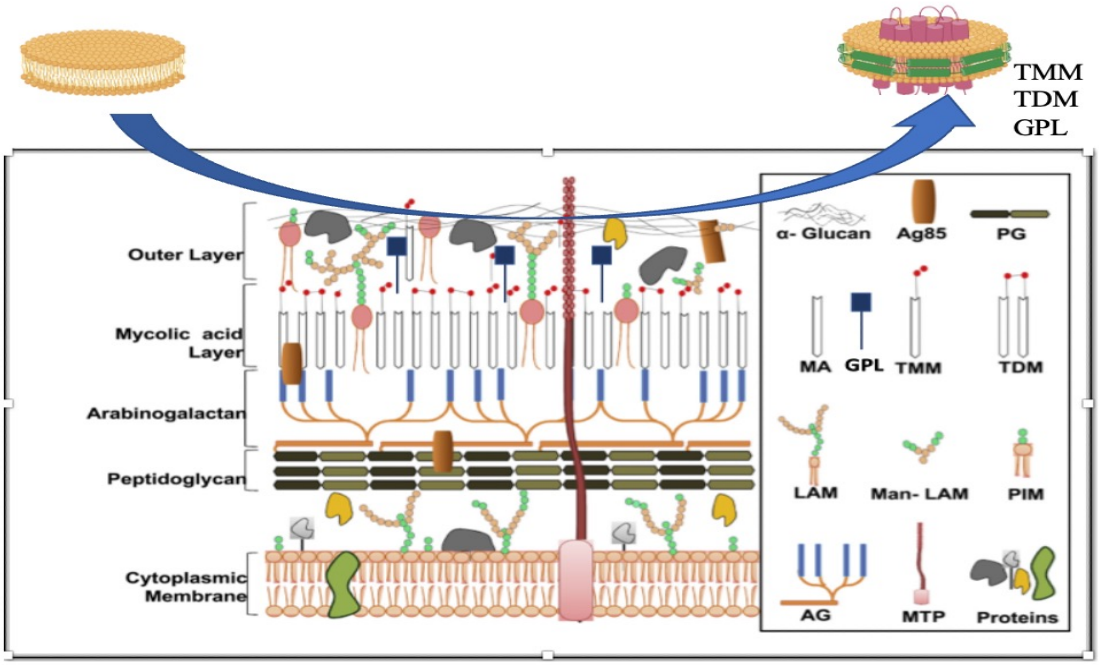


Figure 6.1. **Schematic representation of the Mycobacterium tuberculosis cell envelope and the propose new method of lipid extraction**. Displaying its components such as peptidoglycan, arabinogalactan, and mycolic acid. The surface complex glycolipids shown include trehalose monomycolate (TMM), trehalose dimycolate (TDM), glucose monomycolate (GMM), phospholipids known as phosphatidyl inositol mannosides (PIM), and lipoglycans lipomannan (LM) and lipoarabinomannan (LAM). This figure produced with permission from (Vinod et al., 2020).

6.2 Results

6.2.1 Optimization of [SMA] for extraction of *M. smegmatis* ^{14}C-labelled cell envelope lipids

M. smegmatis wt. was labelled with ^{14}C -acetic acid using the method outlined in section (2.15.1). Lipids were then extracted following the protocol specified in section (2.15.2), and lipid analysis was conducted per the procedure detailed in section (2.15.4). Thin layer chromatography (TLC) was employed to analyse the lipid fractions. The TLC analysis of the extractable lipids (figure 6.2), using the extraction buffers listed in (table 2.5), displayed promising results in lipid extraction. However, the yield of extracted lipids varied across different extraction buffers. Notably, SMA was found to interfere with lipids, impacting their migration on the TLC plate. Using over 2.5% of SMA for extraction led to excessive contamination and poor resolution of lipid components in TLC. Conversely, 1% of SMA did not interfere with peripheral lipids during TLC analysis. Furthermore, through this optimization process, successfully extracted lipids known to be present in the outer layer of mycobacteria, such as trehalose monomycolate (TMM), trehalose dimycolate (TDM), and other lipids like phospholipids (phosphatidylinositol mannosides PIMs, phosphatidylinositol PI).

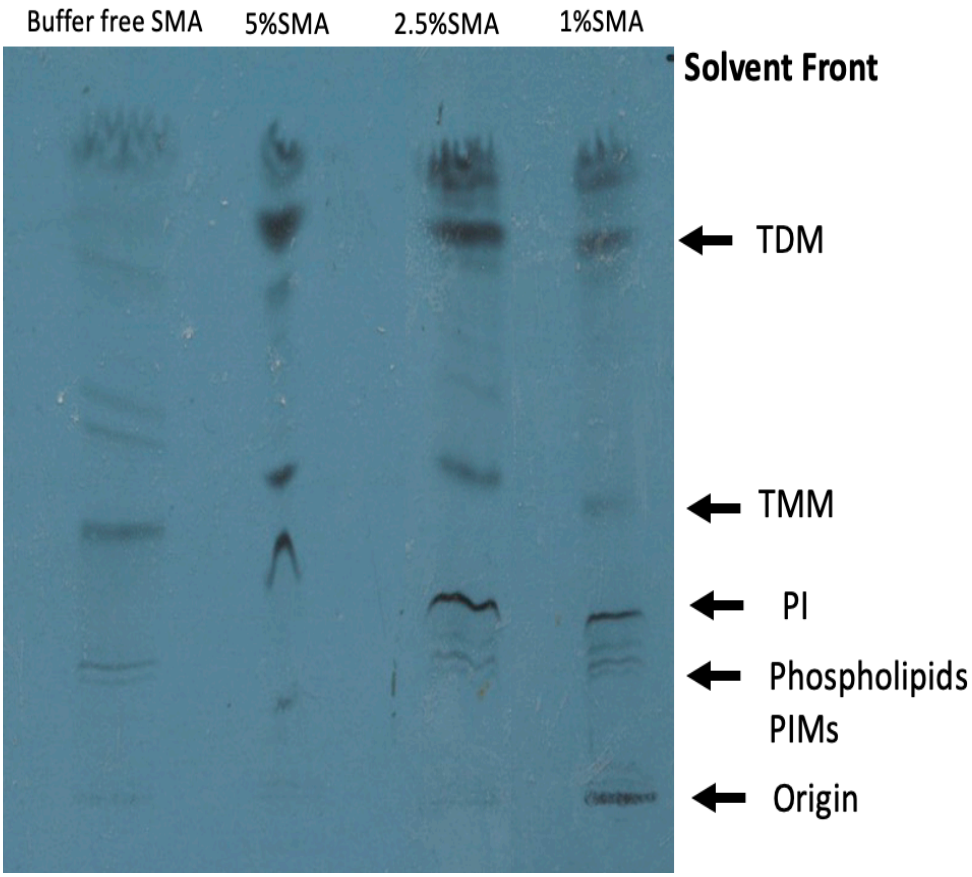


Figure 6.2: Autoradiographs of 1D TLC silica plates of *M. smegmatis* ^{14}C-acetic acid developed in chloroform: methanol: water (80:20:2, v/v/v). *Shows the M. smegmatis* ^{14}C-acetic acid lipids extraction with buffer free SMA (50 mM Tris Hcl, 10% glycerol. 500 mM Nacl), and 3 different concentrations of SMA (5%, 2.5%, and 1%) respectively in the present of glass beads.

6.2.2 1% SMA extraction of *M. smegmatis* ^{14}C-labelled lipids with mechanical agitation using glass beads

To improve the yield of extracted lipids, an additional element into the process were introduced. Due to the distinct structure of the mycobacterial cell envelope, glass beads were incorporated during the extraction process to introduce mechanical agitation, potentially aiding the extraction procedures. Analysis of lipid fractions using TLC with and without mechanical agitation indicated that the presence of mechanical force led to increased lipid extraction yields (Figure 6.3A). Additionally, when employing SMA for extraction, we observed that it extracted different types of lipids compared to buffer extraction. Despite the absence of mechanical force, SMA showed the ability to extract lipids, although the yield was low. To assess the comprehensiveness of lipid extraction, total lipid extraction was performed using organic solvent (2.15.3) on the residual pellet treated with 1% SMA. Figure 6.3B post-SMA extraction via organic solvent indicated that most of the PIMs and PI lipids remained present, suggesting that SMA did not extract all the lipids. Conversely, we noticed a reduction in certain lipids. Using 1% SMA along with mechanical agitation selectively extracted TMM, TDM, and GPL from the mycobacterial cell envelope (A). Subsequent to SMA extraction (B) using organic solvents, there was a reduction in TMM, TDM, and GPL in the mycobacterial cell envelope, likely due to prior extraction via SMA.

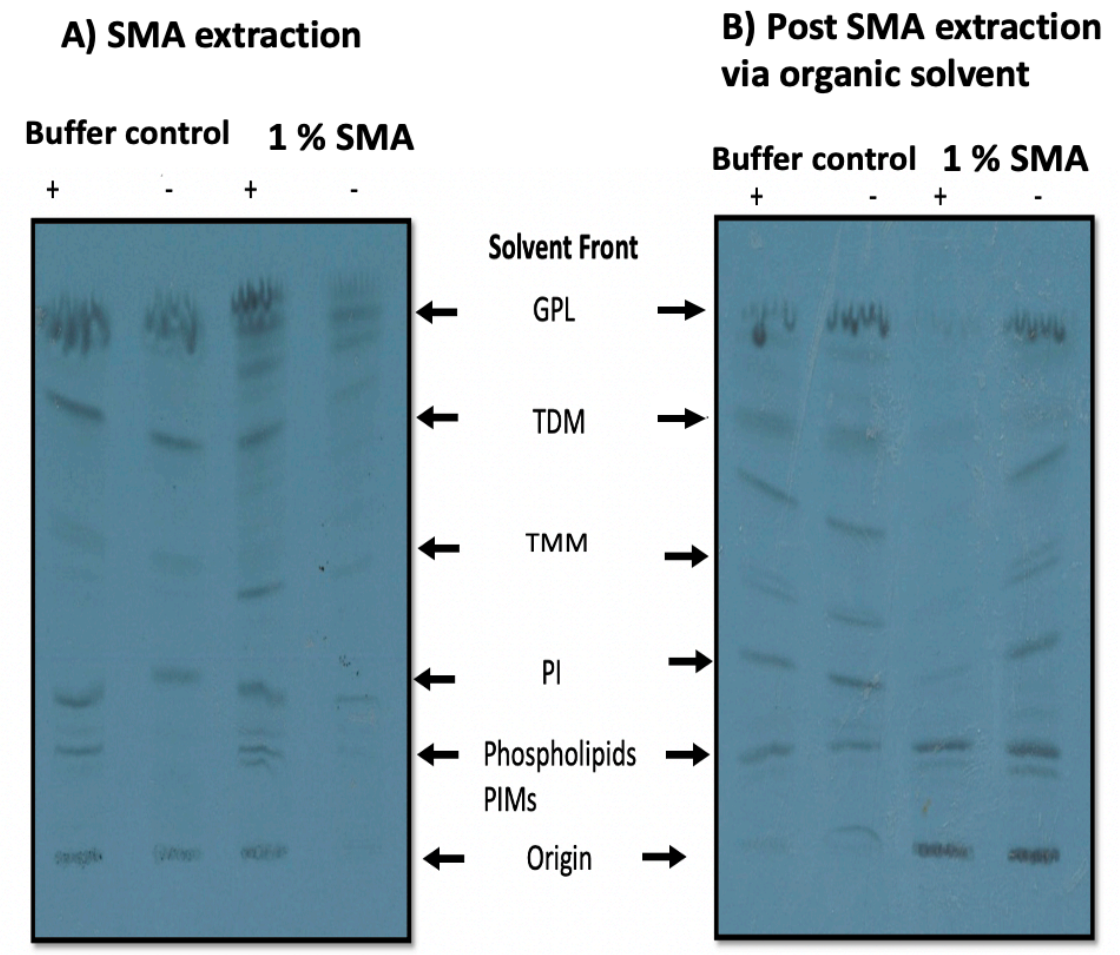


Figure 6.3: Autoradiographs of 1D TLC silica plates of *M. smegmatis* ^{14}C -acetic acid developed in chloroform: methanol: water (80:20:2, v/v/v). A) *Shows the M. smegmatis* ^{14}C -acetic acid lipids extracted via buffer free SMA and 1% SMA with (+) and without (-) glass beads. Also, shows that 1% SMA + and – glass beads have selectively extracted glycopeptidolipids (GPL), and shows sufficient extraction of trehalose monomycolate (TMM), trehalose dimycolate (TDM), phosphatidyl inositol mannosides (PIMs), and phosphatidylinositol (PI) when glass beads are present comparing to the buffer free SMA + and - glass beads and 1% SMA (-) glass beads. B) Total lipids extraction of *M. smegmatis* ^{14}C -acetic acid of leftover pellets that have been treated with 1% SMA and buffer free SMA via 10:10:3 (chloroform: methanol: water).

6.2.3 SMA extraction of *M. smegmatis* ^{14}C-labelled Glycopeptidolipids GPL

While developing the protocol for lipid extraction via SMA, observation from the results of the post-SMA extraction of total lipids using organic solvent in section 6.2.2 that some lipids remained intact within the membrane after the initial SMA extraction attempt. To improve the extraction outcome, we considered conducting two extractions using SMA on the same pellet. First, *M. smegmatis wt* was labelled with ^{14}C -acetic acid following the method outlined in section (2.15.1). Lipids were extracted according to the protocol specified in section (2.15.2) for the initial extraction. The remaining pellets underwent a drying process and then underwent a second round of lipid extraction using the same procedure detailed in section 2.15.2. Analysis of lipids from both extractions was carried out per the procedure outlined in section (2.15.4), with TLC analyses presented in figure 6.4 illustrating the extracted lipids. The lipid profile from the first extraction revealed successful extraction of TDM, TMM, PIMs, and PI. However, during the SMA extraction, we observed interference of SMA with lipids running on the TLC plate, affecting their resolution, potentially due to the use of a SMA extraction buffer that had not been freshly prepared. Conversely, the second lipid extraction via buffer showed no significant extraction of lipids. Notably, the second SMA extraction showed an increased yield of TDM, along with the appearance of a significant band does not present in the first extraction, which identified as glycopeptidolipids (GPL). To confirm our findings regarding GPL, we conducted two-dimensional (2D) TLC analyses (Figure 6.5) to enhance lipid fraction separation and resolution. The 2D TLC results indicated the absence of lipids in the buffer extractions, while the second extraction via SMA corresponded to the 1D TLC results, demonstrating the presence of GPL.

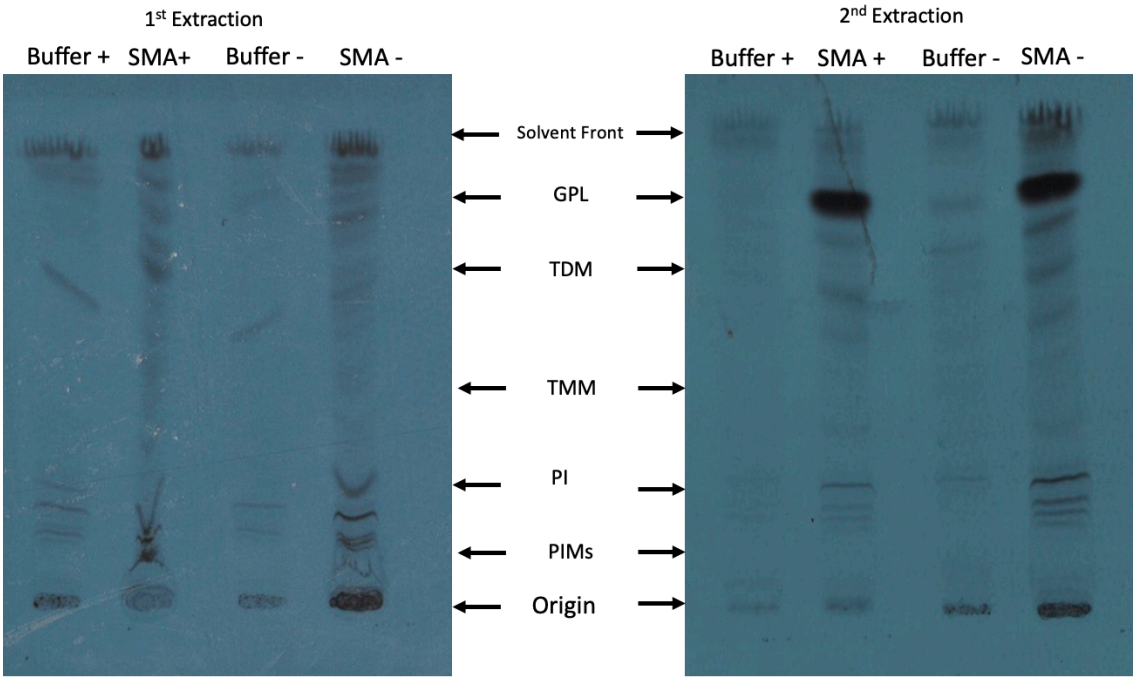


Figure 6.4: Autoradiographs of 1D TLC silica plates of *M. smegmatis* ^{14}C -acetic acid developed in chloroform: methanol: water (80:20:2, v/v/v). *Shows the M. smegmatis* ^{14}C -acetic acid first and second lipids extracted via buffer free SMA and 1% SMA with (+) and without (-) glass beads. Also, shows that 1% SMA + and – glass beads have selectively extracted glycopeptidolipids (GPL) in the second extraction, and shows sufficient extraction of trehalose monomycolate (TMM), trehalose dimycolate (TDM), phosphatidyl inositol mannosides (PIMs), and phosphatidylinositol (PI) when glass beads are present comparing to the buffer free SMA + and - glass beads and 1% SMA (-) glass beads.

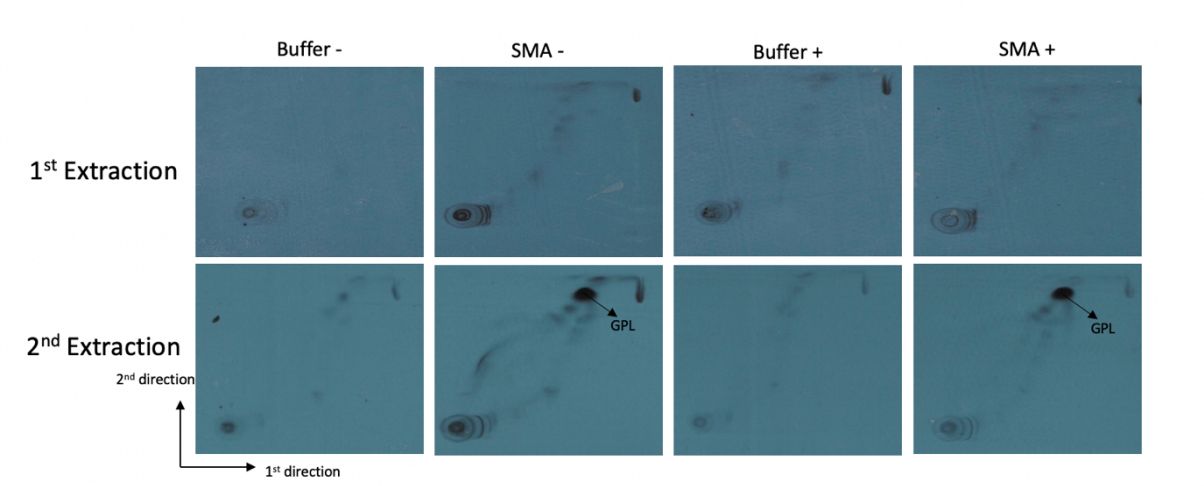


Figure 6.5: Autoradiographs of 2D TLC silica plates of *M. smegmatis* ^{14}C -acetic acid developed in system D. *Shows the M. smegmatis* ^{14}C -acetic acid first and second lipids extracted via buffer free SMA and 1% SMA with (+) and without (-) glass beads developed in system D 1st direction (chloroform: methanol: water) (250:75:4), 2nd direction (chloroform: acetone: methanol: water) (100: 120: 5: 6). Also, shows that 1% SMA + and – glass beads have selectively extracted glycopeptidolipids (GPL) in the second extraction.

6.3 Discussion

The aim of this chapter was to devise a novel method for extracting mycobacterial lipids from the cell envelope. Traditional analytical methods rely on harsh organic solvent extraction conditions that fail to selectively extract outer cell envelope components. This chapter introduces the use of amphipathic polymers like Styrene Maleic Acid (SMA) as a potential new approach for investigating the *Mycobacterium tuberculosis* cell envelope.

Our findings demonstrate that a 1% concentration of SMA is optimal for lipid extraction. This concentration allows for lipid extraction without interfering with peripheral lipids during TLC analysis, minimizes contamination, and preserves lipid resolution. Furthermore, SMA exhibits the capability to extract complex glycopeptidolipids (GPL) and adequately extract trehalose monomycolate (TMM), trehalose dimycolate (TDM), phosphatidyl inositol mannosides (PIMs), and phosphatidylinositol (PI) from the mycobacterial cell envelope.

However, further exploration and refinement of this approach are necessary to establish a comprehensive protocol for lipid extraction via SMA.

Unfortunately, certain circumstances hindered the completion of the assessment of this new method. Notably, the departure of my direct supervisor who was an expert in cell wall biology during the third year of this project which reduced communication during the protocol development significantly impacted the progress. However, it is clear that the method shows utility with improves on and complements current methods.

CHAPTER 7

GENERAL DISCUSSION

Discussion:

Studying *Mycobacterium tuberculosis* is crucial for addressing one of the most pressing global health challenges: tuberculosis (TB). As the second leading infectious cause of death worldwide, TB continues to claim millions of lives, with 10 million new cases and 1.3 million deaths reported in 2022 alone (World Health Organization, 2023). Current methods for TB detection, treatment, and prevention fall short, highlighting the urgent need for improved control measures. A key focus is understanding latent TB, where bacteria persist after treatment, as this knowledge is vital for developing better strategies to prevent and manage the disease (World Health Organization, 2023). The rise of drug-resistant strains adds to the complexity, underscoring the need for new antibiotics and treatments (Smith, 2003). Despite the long-standing use of the BCG vaccine, TB remains a leading killer, pushing researchers to explore more effective vaccines. Studying how *M. tuberculosis* evades the immune system and causes disease offers valuable insights for creating innovative therapies. This pathogen has evolved alongside humans for millennia, making it a prime model for understanding host-pathogen interactions (World Health Organization, 2023). Additionally, discovering biomarkers to identify individuals at varying risk levels could pave the way for personalized treatment approaches. Ultimately, research on *M. tuberculosis* holds the key to developing the next generation of vaccines, drugs, and diagnostic tools to combat this enduring global health threat.

The cell envelope of *Mycobacterium tuberculosis* is an excellent target for novel therapeutics due to its complex architecture, vital for pathogenicity, virulence, and drug resistance. Precise knowledge of molecular interactions on the *M. tuberculosis* envelope surface is key to identifying crucial components. The aim of this study has been to develop our understanding of elements of cell envelope, both its protein and lipid components.

Polymer nano-encapsulation techniques offer promise in studying membrane proteins. Amphipathic polymers like Styrene Maleic Acid (SMA) have been extensively used to directly solubilize proteins and lipids from the cytoplasmic membrane via SMA Lipid Particles

(SMALPs) (Knowles et al., 2009). Using styrene maleic acid (SMA) polymers for membrane protein extraction offers several advantages over traditional detergent-based methods. One of the key benefits is that SMA preserves the native lipid bilayer around the membrane proteins, ensuring greater stability compared to detergent-solubilized proteins, which often suffer from destabilizing effects due to lipid loss (Dörr et al., 2016). This stability is crucial for functional and structural studies, including high-resolution imaging through cryo-electron microscopy. Additionally, SMA-extracted proteins provide valuable structural information, enabling detailed analysis of protein-lipid interactions, surrounding lipids, and protein-protein interactions within complexes (Unger et al., 2021). The method also allows for real-time binding assays and functional studies of proteins that are typically unstable in detergent environments (Wheatley et al., 2016). SMA can help identify the lipids present in the cell envelope and reveal protein compositions, providing valuable insights into the behaviour of *Mycobacterium tuberculosis* during infection and drug resistance. Furthermore, since no studies have yet been conducted to solubilize and extract the selected ABC transporter complexes using SMA, this research could be the first step toward developing new drugs for TB treatment.

Targeting *Mycobacterium tuberculosis* ABC transporters offers a promising avenue for discovering new drug targets to combat tuberculosis. These transporters are critical for the bacterium's survival and virulence, making them ideal candidates for disrupting key bacterial functions (1). ABC transporters play diverse roles, from nutrient uptake to drug resistance and cell wall biosynthesis and inhibiting them could simultaneously affect multiple vital processes in the bacteria (01). Many of these transporters also have unique structural features not found in human cells, allowing for selective targeting without harming the host (Oh et al., 2021). Importantly, ABC transporters contribute to drug resistance by pumping antibiotics out of bacterial cells, so inhibiting them could boost the effectiveness of current treatments (Shetye et al., 2020). Novel transporters like MmpL3 and EccB3 have already been identified as promising drug targets. Recent advances, including whole-cell screening combined with genomic analysis, have shown success in identifying ABC transporters as potential targets for new anti-TB compounds (Ioerger et al., 2013). Moreover, combining ABC transporter inhibitors with existing antibiotics could enhance

their efficacy and help overcome drug resistance (Shetye et al., 2020). However, drug development is complex, and further research is needed to translate these findings into effective treatments (Mi et al., 2022).

Based on the available information, three ABC transporters complexes have been chosen to investigate in this study Rv1272c/73c, Rv3781c/83c, and Rv1217c/18c. Those three complexes thought to be involved in active transport of drugs across the membrane multidrug resistance by an export mechanism. Responsible for energy coupling to the transport system and for the translocation of the substrate across the membrane, form an ATP-driven O-antigen/lipopolysaccharide export apparatus, and involved in active transport of tetronasin across the membrane.

This study explores whether the SMALP method can successfully solubilize the intact complexes of three *Mycobacterium tuberculosis* ABC transporters: Rv1272c/73c, Rv3781c/83c, and Rv1217c/18c. Purification was carried out using Ni-NTA affinity chromatography followed by size exclusion chromatography (SEC), with a goal of achieving over 80% protein purity. The initial objective was to develop a single protocol applicable to all three complexes. However, the varying molecular weights of these complexes made this a challenging task. Following the protocol described by Lee et al. (2016), the study managed to optimize conditions and successfully solubilize and purify the Rv1273c/72c and Rv3781c/83c complexes. In contrast, obtaining a pure protein sample for the Rv1218c/17c complex proved more difficult.

Purifying the entire complex with only one tagged protein posed a challenge, particularly in ensuring that the untagged protein was not lost during the purification process. During the protocol development of Rv1273c and Rv1272c in Chapter 3, similar issues were encountered when aiming to obtain purified protein of Rv3781c and Rv3783c Chapter 4, and Rv1218c/17c Chapter 5. Specifically, the challenge of poor binding to the Ni-NTA, possibly due to the nature of SMALP. SMALP's 7-10 nm size might obstruct tag binding to the resin.

Exploring different resins with varying capacities could enhance binding and overall purification. Common protein contaminants co-eluting with tagged proteins from *E. coli* expression were challenging to remove due to their nickel affinity, creating competition during binding. Using a double-tagged purification technique could potentially resolve this issue by tagging one protein with a combination of tags, such as a hexahistidine tag paired with a FLAG tag, Strep II tag, streptavidin-binding peptide (SBP) tag, calmodulin-binding peptide (CBP), glutathione S-transferase (GST), maltose-binding protein (MBP), S-tag, HA tag, or c-Myc tag. However, it is important to note that larger affinity tags like GST or MBP can sometimes alter the structure and biological activity of the fusion protein (Zhao et al., 2013).

Alternatively, expressing the protein complexes in *Mycobacterium smegmatis* could address this issue, offering insights into the complex and its surrounding lipids. Furthermore, studying individual proteins apart from the complex could elucidate their functions and aid future structural determination.

Ensuring high-quality, high-yield protein was a major concern during purification. Low yield hindered robust structural characterization, limiting the application of higher-resolution methods like small angle X-ray scattering and electron microscopy. Increasing protein yield within SMALP is crucial for achieving higher-resolution data in the future. Low expression levels were a significant issue affecting these extractions. This is a common challenge in membrane protein research and can occur for various reasons, including toxicity to host cells, protein misfolding, and aggregation (Jensen et al., 2017). Changing the expression system might improve the expression like using different host system or used different strain for membrane expression (Pandey et al., 2016). Also, adjusting expression conditions can significantly improve protein yield. Lowering the expression temperature helps reduce stress on host cells, while using weaker promoters or inducible systems allows better control over protein production. Additionally, optimizing the media composition and growth conditions can further enhance expression efficiency (Pandey et al., 2016, Jensen et al., 2017). Another way of improving the expression to use GFP fusion screening involves

attaching a GFP tag to the protein, which allows for rapid assessment of its expression levels and folding quality (Zhao et al., 2013).

This study describes a novel method for extracting mycobacterial lipids from the cell envelope. Traditional analytical methods rely on harsh organic solvent extraction conditions that fail to selectively extract outer cell envelope components. This study introduces the use of amphipathic polymers like Styrene Maleic Acid (SMA) as a potential new approach for investigating the *Mycobacterium tuberculosis* cell envelope.

Our findings demonstrate that a 1% concentration of SMA is optimal for lipid extraction. This concentration allows for lipid extraction without interfering with peripheral lipids during TLC analysis, minimizes contamination, and preserves lipid resolution. Furthermore, SMA exhibits the capability to extract complex glycopeptidolipids (GPL) and adequately extract trehalose monomycolate (TMM), trehalose dimycolate (TDM), phosphatidyl inositol mannosides (PIMs), and phosphatidylinositol (PI) from the mycobacterial cell envelope.

However, further exploration and refinement of this approach are necessary to establish a comprehensive protocol for lipid extraction via SMA. Currently, there are no tuberculosis (TB) drugs specifically designed to target the lipids trehalose dimycolate (TDM) and trehalose monomycolate (TM) directly. However, research and drug development are increasingly focused on targeting lipid metabolism (Kim and Shin, 2023). For example, drugs like isoniazid (INH) and ethionamide (ETH) inhibit mycolic acid synthesis, a crucial component of the mycobacterial cell wall that includes TDM and TM. As a result, these drugs indirectly impact the synthesis and stability of these lipids (Kim and Shin, 2023, Jackson, 2014). Moreover, Advanced drug delivery systems, such as liposomes and solid lipid nanoparticles, are being investigated to improve the delivery and effectiveness of current TB medications. These systems can enhance drug targeting to *Mycobacterium tuberculosis*, potentially affecting lipid-related processes, including those involving TDM and TM (Buya et al., 2021).

In conclusion, this study successfully developed a protocol for solubilizing and purifying three complexes of an ATP-binding protein ABC transporter from *Mycobacterium tuberculosis*, which is likely involved in drug transport as an efflux system. This achievement provides a crucial foundation for further research into the role of these transporters in drug resistance and could pave the way for new therapeutic strategies against tuberculosis. Additionally, this study demonstrates that SMALPs selectively extract peripheral molecules from the cell envelope of mycobacteria. Using [^{14} C]-uniformly labelled lipids, showing that SMA can extract complex molecules and surface-exposed proteins from the envelopes of *Mycobacterium smegmatis* without causing cell rupture. Each part of the project offers promising avenues for future research.

CHAPTER 8

REFERENCES

ABRAHAMS, K. A. & BESRA, G. S. 2018. Mycobacterial cell wall biosynthesis: a multifaceted antibiotic target. *Parasitology,* 145**,** 116-133.

ALDERWICK, L. J., HARRISON, J., LLOYD, G. S. & BIRCH, H. L. 2015. The Mycobacterial Cell Wall--Peptidoglycan and Arabinogalactan. *Cold Spring Harb Perspect Med,* 5**,** a021113.

ALDERWICK, L. J., RADMACHER, E., SEIDEL, M., GANDE, R., HITCHEN, P. G., MORRIS, H. R., DELL, A., SAHM, H., EGGELING, L. & BESRA, G. S. 2005. Deletion of Cg-emb in corynebacterianeae leads to a novel truncated cell wall arabinogalactan, whereas inactivation of Cg-ubiA results in an arabinan-deficient mutant with a cell wall galactan core. *J Biol Chem,* 280**,** 32362-71.

ALTERI, C. J., XICOHTÉNCATL-CORTES, J., HESS, S., CABALLERO-OLÍN, G., GIRÓN, J. A. & FRIEDMAN, R. L. 2007. Mycobacterium tuberculosis produces pili during human infection. *Proc Natl Acad Sci U S A,* 104**,** 5145-50.

AMBUDKAR, S. V., KIM, I. W., XIA, D. & SAUNA, Z. E. 2006. The A-loop, a novel conserved aromatic acid subdomain upstream of the Walker A motif in ABC transporters, is critical for ATP binding. *FEBS Lett,* 580**,** 1049-55.

ANDERSON, R. J. 1929. CHEMICAL INVESTIGATION OF BIOLOGICALLY ACTIVE LIPOIDS OF TUBERCLE BACILLI. *Proc Natl Acad Sci U S A,* 15**,** 628-33.

ARBUES, A., LUGO-VILLARINO, G., NEYROLLES, O., GUILHOT, C. & ASTARIE-DEQUEKER, C. 2014. Playing hide-and-seek with host macrophages through the use of mycobacterial cell envelope phthiocerol dimycocerosates and phenolic glycolipids. *Front Cell Infect Microbiol,* 4**,** 173.

ARBUÉS, A., MALAGA, W., CONSTANT, P., GUILHOT, C., PRANDI, J. & ASTARIE-DEQUEKER, C. 2016. Trisaccharides of Phenolic Glycolipids Confer Advantages to Pathogenic Mycobacteria through Manipulation of Host-Cell Pattern-Recognition Receptors. *ACS Chem Biol,* 11**,** 2865-2875.

ARMSTRONG, J. A. & HART, P. D. 1975. Phagosome-lysosome interactions in cultured macrophages infected with virulent tubercle bacilli. Reversal of the usual nonfusion pattern and observations on bacterial survival. *J Exp Med,* 142**,** 1-16.

BAILO, R., BHATT, A. & AÍNSA, J. A. 2015. Lipid transport in Mycobacterium tuberculosis and its implications in virulence and drug development. *Biochem Pharmacol,* 96**,** 159-67.

BANSAL-MUTALIK, R. & NIKAIDO, H. 2014. Mycobacterial outer membrane is a lipid bilayer and the inner membrane is unusually rich in diacyl phosphatidylinositol dimannosides. *Proc Natl Acad Sci U S A,* 111**,** 4958-63.

BE, N. A., LAMICHHANE, G., GROSSET, J., TYAGI, S., CHENG, Q. J., KIM, K. S., BISHAI, W. R. & JAIN, S. K. 2008. Murine model to study the invasion and survival of Mycobacterium tuberculosis in the central nervous system. *J Infect Dis,* 198**,** 1520-8.

BEIS, K. 2015. Structural basis for the mechanism of ABC transporters. *Biochem Soc Trans,* 43**,** 889-93.

BEKIERKUNST, A. 1968. Acute granulomatous response produced in mice by trehalose-6,6-dimycolate. *J Bacteriol,* 96**,** 958-61.

BEKIERKUNST, A., YARKONI, E., FLECHNER, I., MORECKI, S., VILKAS, E. & LEDERER, E. 1971. Immune response to sheep red blood cells in mice pretreated with mycobacterial fractions. *Infect Immun,* 4**,** 256-63.

BERGMANN, S., ROHDE, M. & HAMMERSCHMIDT, S. 2004. Glyceraldehyde-3-phosphate dehydrogenase of Streptococcus pneumoniae is a surface-displayed plasminogen-binding protein. *Infect Immun,* 72**,** 2416-9.

BERNTSSON, R. P., SMITS, S. H., SCHMITT, L., SLOTBOOM, D. J. & POOLMAN, B. 2010. A structural classification of substrate-binding proteins. *FEBS Lett,* 584**,** 2606-17.

BESRA, G. S., KHOO, K. H., MCNEIL, M. R., DELL, A., MORRIS, H. R. & BRENNAN, P. J. 1995. A new interpretation of the structure of the mycolyl-arabinogalactan complex of Mycobacterium tuberculosis as revealed through characterization of oligoglycosylalditol fragments by fast-atom bombardment mass spectrometry and 1H nuclear magnetic resonance spectroscopy. *Biochemistry,* 34**,** 4257-66.

BHATT, A., BROWN, A. K., SINGH, A., MINNIKIN, D. E. & BESRA, G. S. 2008. Loss of a mycobacterial gene encoding a reductase leads to an altered cell wall containing beta-oxo-mycolic acid analogs and accumulation of ketones. *Chem Biol,* 15**,** 930-9.

BHATT, A., MOLLE, V., BESRA, G. S., JACOBS, W. R., JR. & KREMER, L. 2007. The Mycobacterium tuberculosis FAS-II condensing enzymes: their role in mycolic acid biosynthesis, acid-fastness, pathogenesis and in future drug development. *Mol Microbiol,* 64**,** 1442-54.

BLOCH, H. 1950. Studies on the virulence of tubercle bacilli; the relationship of the physiological state of the organisms to their pathogenicity. *J Exp Med,* 92**,** 507-26.

BORST, P. & ELFERINK, R. O. 2002. Mammalian ABC transporters in health and disease. *Annu Rev Biochem,* 71**,** 537-92.

BOWDISH, D. M., SAKAMOTO, K., KIM, M. J., KROOS, M., MUKHOPADHYAY, S., LEIFER, C. A., TRYGGVASON, K., GORDON, S. & RUSSELL, D. G. 2009. MARCO, TLR2, and CD14 are required for macrophage cytokine responses to mycobacterial trehalose dimycolate and Mycobacterium tuberculosis. *PLoS Pathog,* 5**,** e1000474.

BRAIBANT, M., GILOT, P. & CONTENT, J. 2000. The ATP binding cassette (ABC) transport systems of Mycobacterium tuberculosis. *FEMS Microbiol Rev,* 24**,** 449-67.

BRENNAN, P. J. 2003. Structure, function, and biogenesis of the cell wall of Mycobacterium tuberculosis. *Tuberculosis (Edinb),* 83**,** 91-7.

BRENNAN, P. J. & NIKAIDO, H. 1995. The envelope of mycobacteria. *Annu Rev Biochem,* 64**,** 29-63.

BUSSI, C. & GUTIERREZ, M. G. 2019. Mycobacterium tuberculosis infection of host cells in space and time. *FEMS Microbiol Rev,* 43**,** 341-361.

BUTLER, W. R., AHEARN, D. G. & KILBURN, J. O. 1986. High-performance liquid chromatography of mycolic acids as a tool in the identification of Corynebacterium, Nocardia, Rhodococcus, and Mycobacterium species. *J Clin Microbiol,* 23**,** 182-5.

BUYA, A. B., WITIKA, B. A., BAPOLISI, A. M., MWILA, C., MUKUBWA, G. K., MEMVANGA, P. B., MAKONI, P. A. & NKANGA, C. I. 2021. Application of Lipid-Based Nanocarriers for Antitubercular Drug Delivery: A Review. *Pharmaceutics,* 13.

CARROLL, M. V., SIM, R. B., BIGI, F., JÄKEL, A., ANTROBUS, R. & MITCHELL, D. A. 2010. Identification of four novel DC-SIGN ligands on Mycobacterium bovis BCG. *Protein Cell,* 1**,** 859-70.

CASHMORE, T. J., KLATT, S., YAMARYO-BOTTE, Y., BRAMMANANTH, R., RAINCZUK, A. K., MCCONVILLE, M. J., CRELLIN, P. K. & COPPEL, R. L. 2017. Identification of a Membrane Protein Required for Lipomannan Maturation and Lipoarabinomannan Synthesis in Corynebacterineae. *J Biol Chem,* 292**,** 4976-4986.

CASSIO BARRETO DE OLIVEIRA, M. & BALAN, A. 2020. The ATP-Binding Cassette (ABC) Transport Systems in Mycobacterium tuberculosis: Structure, Function, and Possible Targets for Therapeutics. *Biology (Basel),* 9.

CONVERSE, S. E., MOUGOUS, J. D., LEAVELL, M. D., LEARY, J. A., BERTOZZI, C. R. & COX, J. S. 2003. MmpL8 is required for sulfolipid-1 biosynthesis and Mycobacterium tuberculosis virulence. *Proc Natl Acad Sci U S A,* 100**,** 6121-6.

CUTHBERTSON, L., KOS, V. & WHITFIELD, C. 2010. ABC transporters involved in export of cell surface glycoconjugates. *Microbiol Mol Biol Rev,* 74**,** 341-62.

DAFFÉ, M., LACAVE, C., LANÉELLE, M. A. & LANÉELLE, G. 1987. Structure of the major triglycosyl phenol-phthiocerol of Mycobacterium tuberculosis (strain Canetti). *Eur J Biochem,* 167**,** 155-60.

DAFFE, M., MCNEIL, M. & BRENNAN, P. J. 1991. Novel type-specific lipooligosaccharides from Mycobacterium tuberculosis. *Biochemistry,* 30**,** 378-88.

DAO, D. N., KREMER, L., GUÉRARDEL, Y., MOLANO, A., JACOBS, W. R., JR., PORCELLI, S. A. & BRIKEN, V. 2004. Mycobacterium tuberculosis lipomannan induces apoptosis and interleukin-12 production in macrophages. *Infect Immun,* 72**,** 2067-74.

DAVIDSON, A. L., DASSA, E., ORELLE, C. & CHEN, J. 2008. Structure, function, and evolution of bacterial ATP-binding cassette systems. *Microbiol Mol Biol Rev,* 72**,** 317-64, table of contents.

DAWSON, R. J. & LOCHER, K. P. 2006. Structure of a bacterial multidrug ABC transporter. *Nature,* 443**,** 180-5.

DILWORTH, M. V., PIEL, M. S., BETTANEY, K. E., MA, P., LUO, J., SHARPLES, D., POYNER, D. R., GROSS, S. R., MONCOQ, K., HENDERSON, P. J. F., MIROUX, B. & BILL, R. M. 2018. Microbial expression systems for membrane proteins. *Methods,* 147**,** 3-39.

DOBOS, K. M., KHOO, K. H., SWIDEREK, K. M., BRENNAN, P. J. & BELISLE, J. T. 1996. Definition of the full extent of glycosylation of the 45-kilodalton glycoprotein of Mycobacterium tuberculosis. *J Bacteriol,* 178**,** 2498-506.

DOMINGO-GONZALEZ, R., DAS, S., GRIFFITHS, K. L., AHMED, M., BAMBOUSKOVA, M., GOPAL, R., GONDI, S., MUÑOZ-TORRICO, M., SALAZAR-LEZAMA, M. A., CRUZ-LAGUNAS, A., JIMÉNEZ-ÁLVAREZ, L., RAMIREZ-MARTINEZ, G., ESPINOSA-SOTO, R., SULTANA, T., LYONS-WEILER, J., REINHART, T. A., ARCOS, J., DE LA LUZ GARCIA-HERNANDEZ, M., MASTRANGELO, M. A., AL-HAMMADI, N., TOWNSEND, R., BALADA-LLASAT, J. M., TORRELLES, J. B., KAPLAN, G., HORNE, W., KOLLS, J. K., ARTYOMOV, M. N., RANGEL-MORENO, J., ZÚÑIGA, J. & KHADER, S. A. 2017. Interleukin-17 limits hypoxia-inducible factor 1α and development of hypoxic granulomas during tuberculosis. *JCI Insight,* 2.

DONG, H., ZHANG, Z., TANG, X., PATERSON, N. G. & DONG, C. 2017. Structural and functional insights into the lipopolysaccharide ABC transporter LptB(2)FG. *Nat Commun,* 8**,** 222.

DÖRR, J. M., SCHEIDELAAR, S., KOORENGEVEL, M. C., DOMINGUEZ, J. J., SCHÄFER, M., VAN WALREE, C. A. & KILLIAN, J. A. 2016. The styrene-maleic acid copolymer: a versatile tool in membrane research. *Eur Biophys J,* 45**,** 3-21.

DRIESSEN, N. N., UMMELS, R., MAASKANT, J. J., GURCHA, S. S., BESRA, G. S., AINGE, G. D., LARSEN, D. S., PAINTER, G. F., VANDENBROUCKE-GRAULS, C. M., GEURTSEN, J. & APPELMELK, B. J. 2009. Role of phosphatidylinositol mannosides in the interaction between mycobacteria and DC-SIGN. *Infect Immun,* 77**,** 4538-47.

DU, D., WANG-KAN, X., NEUBERGER, A., VAN VEEN, H. W., POS, K. M., PIDDOCK, L. J. V. & LUISI, B. F. 2018. Multidrug efflux pumps: structure, function and regulation. *Nat Rev Microbiol,* 16**,** 523-539.

DUMKE, R., HAUSNER, M. & JACOBS, E. 2011. Role of Mycoplasma pneumoniae glyceraldehyde-3-phosphate dehydrogenase (GAPDH) in mediating interactions with the human extracellular matrix. *Microbiology (Reading),* 157**,** 2328-2338.

ERNST, W. A., MAHER, J., CHO, S., NIAZI, K. R., CHATTERJEE, D., MOODY, D. B., BESRA, G. S., WATANABE, Y., JENSEN, P. E., PORCELLI, S. A., KRONENBERG, M. & MODLIN, R. L. 1998. Molecular interaction of CD1b with lipoglycan antigens. *Immunity,* 8**,** 331-40.

ESPARZA, M., PALOMARES, B., GARCÍA, T., ESPINOSA, P., ZENTENO, E. & MANCILLA, R. 2015. PstS-1, the 38-kDa Mycobacterium tuberculosis glycoprotein, is an adhesin, which binds the macrophage mannose receptor and promotes phagocytosis. *Scand J Immunol,* 81**,** 46-55.

FERRARI, G., LANGEN, H., NAITO, M. & PIETERS, J. 1999. A coat protein on phagosomes involved in the intracellular survival of mycobacteria. *Cell,* 97**,** 435-47.

FINK, A. L. 1999. Chaperone-mediated protein folding. *Physiol Rev,* 79**,** 425-49.

FISCHER, K., CHATTERJEE, D., TORRELLES, J., BRENNAN, P. J., KAUFMANN, S. H. & SCHAIBLE, U. E. 2001. Mycobacterial lysocardiolipin is exported from phagosomes upon cleavage of cardiolipin by a macrophage-derived lysosomal phospholipase A2. *J Immunol,* 167**,** 2187-92.

FRANKFATER, C., ABRAMOVITCH, R. B., PURDY, G. E., TURK, J., LEGENTIL, L., LEMIÈGRE, L. & FONG-FU, H. 2019. Multiple-stage Precursor Ion Separation and High Resolution Mass Spectrometry toward Structural Characterization of 2,3-Diacyltrehalose Family from Mycobacterium tuberculosis. *Separations,* 6**,** 4.

FRATTI, R. A., CHUA, J., VERGNE, I. & DERETIC, V. 2003. Mycobacterium tuberculosis glycosylated phosphatidylinositol causes phagosome maturation arrest. *Proc Natl Acad Sci U S A,* 100**,** 5437-42.

GARCIA-VILANOVA, A., CHAN, J. & TORRELLES, J. B. 2019. Underestimated Manipulative Roles of Mycobacterium tuberculosis Cell Envelope Glycolipids During Infection. *Front Immunol,* 10**,** 2909.

GATFIELD, J., ALBRECHT, I., ZANOLARI, B., STEINMETZ, M. O. & PIETERS, J. 2005. Association of the leukocyte plasma membrane with the actin cytoskeleton through coiled coil-mediated trimeric coronin 1 molecules. *Mol Biol Cell,* 16**,** 2786-98.

GATFIELD, J. & PIETERS, J. 2000. Essential role for cholesterol in entry of mycobacteria into macrophages. *Science,* 288**,** 1647-50.

GAUTIER, N., LÓPEZ MARÍN, L. M., LANÉELLE, M. A. & DAFFÉ, M. 1992. Structure of mycoside F, a family of trehalose-containing glycolipids of Mycobacterium fortuitum. *FEMS Microbiol Lett,* 77**,** 81-7.

GILLERON, M., QUESNIAUX, V. F. & PUZO, G. 2003. Acylation state of the phosphatidylinositol hexamannosides from Mycobacterium bovis bacillus Calmette Guerin and mycobacterium tuberculosis H37Rv and its implication in Toll-like receptor response. *J Biol Chem,* 278**,** 29880-9.

GILLERON, M., STENGER, S., MAZORRA, Z., WITTKE, F., MARIOTTI, S., BÖHMER, G., PRANDI, J., MORI, L., PUZO, G. & DE LIBERO, G. 2004. Diacylated sulfoglycolipids are novel mycobacterial antigens stimulating CD1-restricted T cells during infection with Mycobacterium tuberculosis. *J Exp Med,* 199**,** 649-59.

GORDON, S. V. & PARISH, T. 2018. Microbe Profile: Mycobacterium tuberculosis: Humanity's deadly microbial foe. *Microbiology (Reading),* 164**,** 437-439.

GOREN, M. B., D'ARCY HART, P., YOUNG, M. R. & ARMSTRONG, J. A. 1976. Prevention of phagosome-lysosome fusion in cultured macrophages by sulfatides of Mycobacterium tuberculosis. *Proc Natl Acad Sci U S A,* 73**,** 2510-4.

GOREN, M. B. & MOR, N. 1990. Influence of phagosomal contents on the apparent inhibition of phagosome-lysosome fusion mediated by polyanionic substances in mouse peritoneal macrophages. *Biochem Cell Biol,* 68**,** 24-32.

GOVENDER, V. S., RAMSUGIT, S. & PILLAY, M. 2014. Mycobacterium tuberculosis adhesins: potential biomarkers as anti-tuberculosis therapeutic and diagnostic targets. *Microbiology (Reading),* 160**,** 1821-1831.

GRIME, R. L., LOGAN, R. T., NESTOROW, S. A., SRIDHAR, P., EDWARDS, P. C., TATE, C. G., KLUMPERMAN, B., DAFFORN, T. R., POYNER, D. R., REEVES, P. J. & WHEATLEY, M. 2021. Differences in SMA-like polymer architecture dictate the conformational changes exhibited by the membrane protein rhodopsin encapsulated in lipid nanoparticles. *Nanoscale,* 13**,** 13519-13528.

GU, S., CHEN, J., DOBOS, K. M., BRADBURY, E. M., BELISLE, J. T. & CHEN, X. 2003. Comprehensive proteomic profiling of the membrane constituents of a Mycobacterium tuberculosis strain. *Mol Cell Proteomics,* 2**,** 1284-96.

GULATI, S., JAMSHAD, M., KNOWLES, T. J., MORRISON, K. A., DOWNING, R., CANT, N., COLLINS, R., KOENDERINK, J. B., FORD, R. C., OVERDUIN, M., KERR, I. D., DAFFORN, T. R. & ROTHNIE, A. J. 2014. Detergent-free purification of ABC (ATP-binding-cassette) transporters. *Biochem J,* 461**,** 269-78.

HASAN, Z., SCHLAX, C., KUHN, L., LEFKOVITS, I., YOUNG, D., THOLE, J. & PIETERS, J. 1997. Isolation and characterization of the mycobacterial phagosome: segregation from the endosomal/lysosomal pathway. *Mol Microbiol,* 24**,** 545-53.

HATZIOS, S. K., SCHELLE, M. W., HOLSCLAW, C. M., BEHRENS, C. R., BOTYANSZKI, Z., LIN, F. L., CARLSON, B. L., KUMAR, P., LEARY, J. A. & BERTOZZI, C. R. 2009. PapA3 is an acyltransferase required for polyacyltrehalose biosynthesis in Mycobacterium tuberculosis. *J Biol Chem,* 284**,** 12745-51.

HENDERSON, B. & MARTIN, A. 2011. Bacterial virulence in the moonlight: multitasking bacterial moonlighting proteins are virulence determinants in infectious disease. *Infect Immun,* 79**,** 3476-91.

HUNTER, R. L., OLSEN, M. R., JAGANNATH, C. & ACTOR, J. K. 2006. Multiple roles of cord factor in the pathogenesis of primary, secondary, and cavitary tuberculosis, including a revised description of the pathology of secondary disease. *Ann Clin Lab Sci,* 36**,** 371-86.

INDRIGO, J., HUNTER, R. L. & ACTOR, J. K. 2002. Influence of trehalose 6,6'-dimycolate (TDM) during mycobacterial infection of bone marrow macrophages. *Microbiology (Reading),* 148**,** 1991-1998.

IOERGER, T. R., O'MALLEY, T., LIAO, R., GUINN, K. M., HICKEY, M. J., MOHAIDEEN, N., MURPHY, K. C., BOSHOFF, H. I., MIZRAHI, V., RUBIN, E. J., SASSETTI, C. M., BARRY, C. E., 3RD, SHERMAN, D. R., PARISH, T. & SACCHETTINI, J. C. 2013. Identification of new drug targets and resistance mechanisms in Mycobacterium tuberculosis. *PLoS One,* 8**,** e75245.

ISHIKAWA, E., ISHIKAWA, T., MORITA, Y. S., TOYONAGA, K., YAMADA, H., TAKEUCHI, O., KINOSHITA, T., AKIRA, S., YOSHIKAI, Y. & YAMASAKI, S. 2009. Direct recognition of the mycobacterial glycolipid, trehalose dimycolate, by C-type lectin Mincle. *J Exp Med,* 206**,** 2879-88.

JACKSON, M. 2014. The mycobacterial cell envelope-lipids. *Cold Spring Harb Perspect Med,* 4.

JANKUTE, M., COX, J. A., HARRISON, J. & BESRA, G. S. 2015. Assembly of the Mycobacterial Cell Wall. *Annu Rev Microbiol,* 69**,** 405-23.

JAYACHANDRAN, R., SUNDARAMURTHY, V., COMBALUZIER, B., MUELLER, P., KORF, H., HUYGEN, K., MIYAZAKI, T., ALBRECHT, I., MASSNER, J. & PIETERS, J. 2007. Survival of mycobacteria in macrophages is mediated by coronin 1-dependent activation of calcineurin. *Cell,* 130**,** 37-50.

JENSEN, H. M., ENG, T., CHUBUKOV, V., HERBERT, R. A. & MUKHOPADHYAY, A. 2017. Improving membrane protein expression and function using genomic edits. *Sci Rep,* 7**,** 13030.

JONES, P. M. & GEORGE, A. M. 2002. Mechanism of ABC transporters: a molecular dynamics simulation of a well characterized nucleotide-binding subunit. *Proc Natl Acad Sci U S A,* 99**,** 12639-44.

JÓZEFOWSKI, S., SOBOTA, A., PAWŁOWSKI, A. & KWIATKOWSKA, K. 2011. Mycobacterium tuberculosis lipoarabinomannan enhances LPS-induced TNF-α production and inhibits NO secretion by engaging scavenger receptors. *Microb Pathog,* 50**,** 350-9.

JUNG, S. B., YANG, C. S., LEE, J. S., SHIN, A. R., JUNG, S. S., SON, J. W., HARDING, C. V., KIM, H. J., PARK, J. K., PAIK, T. H., SONG, C. H. & JO, E. K. 2006. The mycobacterial 38-kilodalton glycolipoprotein antigen activates the mitogen-activated protein kinase pathway and release of proinflammatory cytokines through Toll-like receptors 2 and 4 in human monocytes. *Infect Immun,* 74**,** 2686-96.

KALSCHEUER, R., PALACIOS, A., ANSO, I., CIFUENTE, J., ANGUITA, J., JACOBS, W. R., JR., GUERIN, M. E. & PRADOS-ROSALES, R. 2019. The Mycobacterium tuberculosis capsule: a cell structure with key implications in pathogenesis. *Biochem J,* 476**,** 1995-2016.

KARAKOUSIS, P. C., YOSHIMATSU, T., LAMICHHANE, G., WOOLWINE, S. C., NUERMBERGER, E. L., GROSSET, J. & BISHAI, W. R. 2004. Dormancy phenotype displayed by extracellular Mycobacterium tuberculosis within artificial granulomas in mice. *J Exp Med,* 200**,** 647-57.

KAUFMANN, S. H. 2001. How can immunology contribute to the control of tuberculosis? *Nat Rev Immunol,* 1**,** 20-30.

KHOO, K. H., DELL, A., MORRIS, H. R., BRENNAN, P. J. & CHATTERJEE, D. 1995. Structural definition of acylated phosphatidylinositol mannosides from Mycobacterium tuberculosis: definition of a common anchor for lipomannan and lipoarabinomannan. *Glycobiology,* 5**,** 117-27.

KILLICK, K. E., C, N. C., O'FARRELLY, C., HOKAMP, K., MACHUGH, D. E. & HARRIS, J. 2013. Receptor-mediated recognition of mycobacterial pathogens. *Cell Microbiol,* 15**,** 1484-95.

KIM, H. & SHIN, S. J. 2023. Revolutionizing control strategies against Mycobacterium tuberculosis infection through selected targeting of lipid metabolism. *Cell Mol Life Sci,* 80**,** 291.

KINHIKAR, A. G., VARGAS, D., LI, H., MAHAFFEY, S. B., HINDS, L., BELISLE, J. T. & LAAL, S. 2006. Mycobacterium tuberculosis malate synthase is a laminin-binding adhesin. *Mol Microbiol,* 60**,** 999-1013.

KINHIKAR, A. G., VERMA, I., CHANDRA, D., SINGH, K. K., WELDINGH, K., ANDERSEN, P., HSU, T., JACOBS, W. R., JR. & LAAL, S. 2010. Potential role for ESAT6 in dissemination of M. tuberculosis via human lung epithelial cells. *Mol Microbiol,* 75**,** 92-106.

KNOWLES, T. J., FINKA, R., SMITH, C., LIN, Y. P., DAFFORN, T. & OVERDUIN, M. 2009. Membrane proteins solubilized intact in lipid containing nanoparticles bounded by styrene maleic acid copolymer. *J Am Chem Soc,* 131**,** 7484-5.

KOLIWER-BRANDL, H., SYSON, K., VAN DE WEERD, R., CHANDRA, G., APPELMELK, B., ALBER, M., IOERGER, T. R., JACOBS, W. R., JR., GEURTSEN, J., BORNEMANN, S. & KALSCHEUER, R. 2016. Metabolic Network for the Biosynthesis of Intra- and Extracellular α-Glucans Required for Virulence of Mycobacterium tuberculosis. *PLoS Pathog,* 12**,** e1005768.

KONG, T. H., COATES, A. R., BUTCHER, P. D., HICKMAN, C. J. & SHINNICK, T. M. 1993. Mycobacterium tuberculosis expresses two chaperonin-60 homologs. *Proc Natl Acad Sci U S A,* 90**,** 2608-12.

KORKHOV, V. M., MIREKU, S. A. & LOCHER, K. P. 2012. Structure of AMP-PNP-bound vitamin B12 transporter BtuCD-F. *Nature,* 490**,** 367-72.

KREMER, L., MAUGHAN, W. N., WILSON, R. A., DOVER, L. G. & BESRA, G. S. 2002. The M. tuberculosis antigen 85 complex and mycolyltransferase activity. *Lett Appl Microbiol,* 34**,** 233-7.

KUMAR, P., SCHELLE, M. W., JAIN, M., LIN, F. L., PETZOLD, C. J., LEAVELL, M. D., LEARY, J. A., COX, J. S. & BERTOZZI, C. R. 2007. PapA1 and PapA2 are acyltransferases essential for the biosynthesis of the Mycobacterium tuberculosis virulence factor sulfolipid-1. *Proc Natl Acad Sci U S A,* 104**,** 11221-6.

LEDERER, E., ADAM, A., CIORBARU, R., PETIT, J. F. & WIETZERBIN, J. 1975. Cell walls of Mycobacteria and related organisms; chemistry and immunostimulant properties. *Mol Cell Biochem,* 7**,** 87-104.

LEE, H. J., LEE, H. S., YOUN, T., BYRNE, B. & CHAE, P. S. 2022. Impact of novel detergents on membrane protein studies. *Chem,* 8**,** 980-1013.

LEE, J. Y., YANG, J. G., ZHITNITSKY, D., LEWINSON, O. & REES, D. C. 2014. Structural basis for heavy metal detoxification by an Atm1-type ABC exporter. *Science,* 343**,** 1133-6.

LEE, K. S., DUBEY, V. S., KOLATTUKUDY, P. E., SONG, C. H., SHIN, A. R., JUNG, S. B., YANG, C. S., KIM, S. Y., JO, E. K., PARK, J. K. & KIM, H. J. 2007. Diacyltrehalose of Mycobacterium tuberculosis inhibits lipopolysaccharide- and mycobacteria-induced proinflammatory cytokine production in human monocytic cells. *FEMS Microbiol Lett,* 267**,** 121-8.

LEE, S. C., KNOWLES, T. J., POSTIS, V. L., JAMSHAD, M., PARSLOW, R. A., LIN, Y. P., GOLDMAN, A., SRIDHAR, P., OVERDUIN, M., MUENCH, S. P. & DAFFORN, T. R. 2016. A method for detergent-free isolation of membrane proteins in their local lipid environment. *Nat Protoc,* 11**,** 1149-62.

LEE, S. C. & POLLOCK, N. L. 2016. Membrane proteins: is the future disc shaped? *Biochem Soc Trans,* 44**,** 1011-8.

LEIGH, C. D. & BERTOZZI, C. R. 2008. Synthetic studies toward Mycobacterium tuberculosis sulfolipid-I. *J Org Chem,* 73**,** 1008-17.

LEWTHWAITE, J. C., COATES, A. R., TORMAY, P., SINGH, M., MASCAGNI, P., POOLE, S., ROBERTS, M., SHARP, L. & HENDERSON, B. 2001. Mycobacterium tuberculosis chaperonin 60.1 is a more potent cytokine stimulator than chaperonin 60.2 (Hsp 65) and contains a CD14-binding domain. *Infect Immun,* 69**,** 7349-55.

LI, Y., ORLANDO, B. J. & LIAO, M. 2019. Structural basis of lipopolysaccharide extraction by the LptB(2)FGC complex. *Nature,* 567**,** 486-490.
LIN, D. Y., HUANG, S. & CHEN, J. 2015. Crystal structures of a polypeptide processing and secretion transporter. *Nature,* 523**,** 425-30.
LIU, J., BARRY, C. E., 3RD, BESRA, G. S. & NIKAIDO, H. 1996. Mycolic acid structure determines the fluidity of the mycobacterial cell wall. *J Biol Chem,* 271**,** 29545-51.
LOCHER, K. P. 2016. Mechanistic diversity in ATP-binding cassette (ABC) transporters. *Nat Struct Mol Biol,* 23**,** 487-93.
LOCHER, K. P., LEE, A. T. & REES, D. C. 2002. The E. coli BtuCD structure: a framework for ABC transporter architecture and mechanism. *Science,* 296**,** 1091-8.
LUO, Q., YANG, X., YU, S., SHI, H., WANG, K., XIAO, L., ZHU, G., SUN, C., LI, T., LI, D., ZHANG, X., ZHOU, M. & HUANG, Y. 2017. Structural basis for lipopolysaccharide extraction by ABC transporter LptB(2)FG. *Nat Struct Mol Biol,* 24**,** 469-474.
MAEDA, N., NIGOU, J., HERRMANN, J. L., JACKSON, M., AMARA, A., LAGRANGE, P. H., PUZO, G., GICQUEL, B. & NEYROLLES, O. 2003. The cell surface receptor DC-SIGN discriminates between Mycobacterium species through selective recognition of the mannose caps on lipoarabinomannan. *J Biol Chem,* 278**,** 5513-6.
MAHAPATRA, S., YAGI, T., BELISLE, J. T., ESPINOSA, B. J., HILL, P. J., MCNEIL, M. R., BRENNAN, P. J. & CRICK, D. C. 2005. Mycobacterial lipid II is composed of a complex mixture of modified muramyl and peptide moieties linked to decaprenyl phosphate. *J Bacteriol,* 187**,** 2747-57.
MAWUENYEGA, K. G., FORST, C. V., DOBOS, K. M., BELISLE, J. T., CHEN, J., BRADBURY, E. M., BRADBURY, A. R. & CHEN, X. 2005. Mycobacterium tuberculosis functional network analysis by global subcellular protein profiling. *Mol Biol Cell,* 16**,** 396-404.
MI, J., GONG, W. & WU, X. 2022. Advances in Key Drug Target Identification and New Drug Development for Tuberculosis. *Biomed Res Int,* 2022**,** 5099312.
MIGLIORI, G. B., ONG, C. W. M., PETRONE, L., D'AMBROSIO, L., CENTIS, R. & GOLETTI, D. 2021. The definition of tuberculosis infection based on the spectrum of tuberculosis disease. *Breathe (Sheff),* 17**,** 210079.
MINNIKIN, D. E., DOBSON, G., SESARDIC, D. & RIDELL, M. 1985. Mycolipenates and mycolipanolates of trehalose from mycobacterium tuberculosis. *J Gen Microbiol,* 131**,** 1369-74.
MINNIKIN, D. E., LEE, O. Y., WU, H. H., BESRA, G. S., BHATT, A., NATARAJ, V., ROTHSCHILD, B. M., SPIGELMAN, M. & DONOGHUE, H. D. 2015a. Ancient mycobacterial lipids: Key reference biomarkers in charting the evolution of tuberculosis. *Tuberculosis (Edinb),* 95 Suppl 1**,** S133-9.
MINNIKIN, D. E., LEE, O. Y. C., WU, H. H. T., NATARAJ, V., DONOGHUE, H. D., RIDELL, M., WATANABE, M., ALDERWICK, L., BHATT, A. & BESRA, G. S. 2015b. Pathophysiological Implications of Cell Envelope Structure in Mycobacterium tuberculosis and Related Taxa. *Tuberculosis - Expanding Knowledge.*
MISHRA, A. K., ALDERWICK, L. J., RITTMANN, D., TATITURI, R. V., NIGOU, J., GILLERON, M., EGGELING, L. & BESRA, G. S. 2007. Identification of an alpha(1-->6) mannopyranosyltransferase (MptA), involved in Corynebacterium glutamicum lipomanann biosynthesis, and identification of its orthologue in Mycobacterium tuberculosis. *Mol Microbiol,* 65**,** 1503-17.
MOUGOUS, J. D., PETZOLD, C. J., SENARATNE, R. H., LEE, D. H., AKEY, D. L., LIN, F. L., MUNCHEL, S. E., PRATT, M. R., RILEY, L. W., LEARY, J. A., BERGER, J. M. & BERTOZZI, C.

R. 2004. Identification, function and structure of the mycobacterial sulfotransferase that initiates sulfolipid-1 biosynthesis. *Nat Struct Mol Biol,* 11**,** 721-9.
MOYNIHAN, P. J., CADBY, I. T., VEERAPEN, N., JANKUTE, M., CROSATTI, M., MUKAMOLOVA, G. V., LOVERING, A. L. & BESRA, G. S. 2019. The hydrolase LpqI primes mycobacterial peptidoglycan recycling. *Nat Commun,* 10**,** 2647.
NAIDOO, N., RAMSUGIT, S. & PILLAY, M. 2014. Mycobacterium tuberculosis pili (MTP), a putative biomarker for a tuberculosis diagnostic test. *Tuberculosis (Edinb),* 94**,** 338-45.
NATARAJ, V., VARELA, C., JAVID, A., SINGH, A., BESRA, G. S. & BHATT, A. 2015. Mycolic acids: deciphering and targeting the Achilles' heel of the tubercle bacillus. *Mol Microbiol,* 98**,** 7-16.
OH, S., TRIFONOV, L., YADAV, V. D., BARRY, C. E., 3RD & BOSHOFF, H. I. 2021. Tuberculosis Drug Discovery: A Decade of Hit Assessment for Defined Targets. *Front Cell Infect Microbiol,* 11**,** 611304.
OLDHAM, M. L., KHARE, D., QUIOCHO, F. A., DAVIDSON, A. L. & CHEN, J. 2007. Crystal structure of a catalytic intermediate of the maltose transporter. *Nature,* 450**,** 515-21.
OWENS, T. W., TAYLOR, R. J., PAHIL, K. S., BERTANI, B. R., RUIZ, N., KRUSE, A. C. & KAHNE, D. 2019. Structural basis of unidirectional export of lipopolysaccharide to the cell surface. *Nature,* 567**,** 550-553.
PALANIVEL, J., SOUNDERRAJAN, V., THANGAM, T., RAO, S. S., HARSHAVARDHAN, S. & PARTHASARATHY, K. 2023. Latent Tuberculosis: Challenges in Diagnosis and Treatment, Perspectives, and the Crucial Role of Biomarkers. *Curr Microbiol,* 80**,** 392.
PANDEY, A., SHIN, K., PATTERSON, R. E., LIU, X. Q. & RAINEY, J. K. 2016. Current strategies for protein production and purification enabling membrane protein structural biology. *Biochem Cell Biol,* 94**,** 507-527.
PAPA, F., CRUAUD, P. & DAVID, H. L. 1989. Antigenicity and specificity of selected glycolipid fractions from Mycobacterium tuberculosis. *Res Microbiol,* 140**,** 569-78.
PARANT, M., PARANT, F., CHEDID, L., DRAPIER, J. C., PETIT, J. F., WIETZERBIN, J. & LEDERER 1977. Enhancement of nonspecific immunity to bacterial infection by cord factor (6,6'-trehalose dimycolate). *J Infect Dis,* 135**,** 771-7.
PETHE, K., ALONSO, S., BIET, F., DELOGU, G., BRENNAN, M. J., LOCHT, C. & MENOZZI, F. D. 2001. The heparin-binding haemagglutinin of M. tuberculosis is required for extrapulmonary dissemination. *Nature,* 412**,** 190-4.
PIETERS, J. 2008. Mycobacterium tuberculosis and the macrophage: maintaining a balance. *Cell Host Microbe,* 3**,** 399-407.
PITARQUE, S., HERRMANN, J. L., DUTEYRAT, J. L., JACKSON, M., STEWART, G. R., LECOINTE, F., PAYRE, B., SCHWARTZ, O., YOUNG, D. B., MARCHAL, G., LAGRANGE, P. H., PUZO, G., GICQUEL, B., NIGOU, J. & NEYROLLES, O. 2005. Deciphering the molecular bases of Mycobacterium tuberculosis binding to the lectin DC-SIGN reveals an underestimated complexity. *Biochem J,* 392**,** 615-24.
POLLOCK, N. L., LLOYD, J., MONTINARO, C., RAI, M. & DAFFORN, T. R. 2022. Conformational trapping of an ABC transporter in polymer lipid nanoparticles. *Biochem J,* 479**,** 145-159.
POSTIS, V., RAWSON, S., MITCHELL, J. K., LEE, S. C., PARSLOW, R. A., DAFFORN, T. R., BALDWIN, S. A. & MUENCH, S. P. 2015. The use of SMALPs as a novel membrane protein scaffold for structure study by negative stain electron microscopy. *Biochim Biophys Acta,* 1848**,** 496-501.

PUECH, V., CHAMI, M., LEMASSU, A., LANÉELLE, M. A., SCHIFFLER, B., GOUNON, P., BAYAN, N., BENZ, R. & DAFFÉ, M. 2001. Structure of the cell envelope of corynebacteria: importance of the non-covalently bound lipids in the formation of the cell wall permeability barrier and fracture plane. *Microbiology (Reading),* 147**,** 1365-1382.

REES, D. C., JOHNSON, E. & LEWINSON, O. 2009. ABC transporters: the power to change. *Nat Rev Mol Cell Biol,* 10**,** 218-27.

REMPEL, S., STANEK, W. K. & SLOTBOOM, D. J. 2019. ECF-Type ATP-Binding Cassette Transporters. *Annu Rev Biochem,* 88**,** 551-576.

REN, H., DOVER, L. G., ISLAM, S. T., ALEXANDER, D. C., CHEN, J. M., BESRA, G. S. & LIU, J. 2007. Identification of the lipooligosaccharide biosynthetic gene cluster from Mycobacterium marinum. *Mol Microbiol,* 63**,** 1345-59.

RHOADES, E. R., STREETER, C., TURK, J. & HSU, F. F. 2011. Characterization of sulfolipids of Mycobacterium tuberculosis H37Rv by multiple-stage linear ion-trap high-resolution mass spectrometry with electrospray ionization reveals that the family of sulfolipid II predominates. *Biochemistry,* 50**,** 9135-47.

RICHARDSON, M. B. & WILLIAMS, S. J. 2014. MCL and Mincle: C-Type Lectin Receptors That Sense Damaged Self and Pathogen-Associated Molecular Patterns. *Front Immunol,* 5**,** 288.

ROBEY, R. W., PLUCHINO, K. M., HALL, M. D., FOJO, A. T., BATES, S. E. & GOTTESMAN, M. M. 2018. Revisiting the role of ABC transporters in multidrug-resistant cancer. *Nat Rev Cancer,* 18**,** 452-464.

ROTH, M. G. 2004. Phosphoinositides in constitutive membrane traffic. *Physiol Rev,* 84**,** 699-730.

ROUSSEAU, C., NEYROLLES, O., BORDAT, Y., GIROUX, S., SIRAKOVA, T. D., PREVOST, M. C., KOLATTUKUDY, P. E., GICQUEL, B. & JACKSON, M. 2003. Deficiency in mycolipenate- and mycosanoate-derived acyltrehaloses enhances early interactions of Mycobacterium tuberculosis with host cells. *Cell Microbiol,* 5**,** 405-15.

RUSSELL, D. G. 2001. Mycobacterium tuberculosis: here today, and here tomorrow. *Nat Rev Mol Cell Biol,* 2**,** 569-77.

SAAVEDRA, R., SEGURA, E., LEYVA, R., ESPARZA, L. A. & LÓPEZ-MARÍN, L. M. 2001. Mycobacterial di-O-acyl-trehalose inhibits mitogen- and antigen-induced proliferation of murine T cells in vitro. *Clin Diagn Lab Immunol,* 8**,** 1081-8.

SALEH, M. T. & BELISLE, J. T. 2000. Secretion of an acid phosphatase (SapM) by Mycobacterium tuberculosis that is similar to eukaryotic acid phosphatases. *J Bacteriol,* 182**,** 6850-3.

SAURIN, W., HOFNUNG, M. & DASSA, E. 1999. Getting in or out: early segregation between importers and exporters in the evolution of ATP-binding cassette (ABC) transporters. *J Mol Evol,* 48**,** 22-41.

SCHLEIFER, K. H. & KANDLER, O. 1972. Peptidoglycan types of bacterial cell walls and their taxonomic implications. *Bacteriol Rev,* 36**,** 407-77.

SCHUCK, P. 2000. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys J,* 78**,** 1606-19.

SHETYE, G. S., FRANZBLAU, S. G. & CHO, S. 2020. New tuberculosis drug targets, their inhibitors, and potential therapeutic impact. *Transl Res,* 220**,** 68-97.

SILVA, C. L., EKIZLERIAN, S. M. & FAZIOLI, R. A. 1985. Role of cord factor in the modulation of infection caused by mycobacteria. *Am J Pathol,* 118**,** 238-47.

SMITH, I. 2003. Mycobacterium tuberculosis pathogenesis and molecular determinants of virulence. *Clin Microbiol Rev,* 16**,** 463-96.
STROUD, Z., HALL, S. C. L. & DAFFORN, T. R. 2018. Purification of membrane proteins free from conventional detergents: SMA, new polymers, new opportunities and new insights. *Methods,* 147**,** 106-117.
SU, C. C., KLENOTIC, P. A., BOLLA, J. R., PURDY, G. E., ROBINSON, C. V. & YU, E. W. 2019. MmpL3 is a lipid transporter that binds trehalose monomycolate and phosphatidylethanolamine. *Proc Natl Acad Sci U S A,* 116**,** 11241-11246.
SUNDARARAJAN, S. & MUNIYAN, R. 2021. Latent tuberculosis: interaction of virulence factors in Mycobacterium tuberculosis. *Mol Biol Rep,* 48**,** 6181-6196.
TAILLEUX, L., SCHWARTZ, O., HERRMANN, J. L., PIVERT, E., JACKSON, M., AMARA, A., LEGRES, L., DREHER, D., NICOD, L. P., GLUCKMAN, J. C., LAGRANGE, P. H., GICQUEL, B. & NEYROLLES, O. 2003. DC-SIGN is the major Mycobacterium tuberculosis receptor on human dendritic cells. *J Exp Med,* 197**,** 121-7.
TAKAYAMA, K., WANG, C. & BESRA, G. S. 2005. Pathway to synthesis and processing of mycolic acids in Mycobacterium tuberculosis. *Clin Microbiol Rev,* 18**,** 81-101.
TAKEYA, K. & HISATSUNE, K. 1963. Mycobacterial cell walls. I. Methods of preparation and treatment with various chemicals. *J Bacteriol,* 85**,** 16-23.
TANG, P. & JOHNSTON, J. 2017. Treatment of Latent Tuberculosis Infection. *Curr Treat Options Infect Dis,* 9**,** 371-379.
TELFORD, J. L., BAROCCHI, M. A., MARGARIT, I., RAPPUOLI, R. & GRANDI, G. 2006. Pili in gram-positive pathogens. *Nat Rev Microbiol,* 4**,** 509-19.
TER BEEK, J., GUSKOV, A. & SLOTBOOM, D. J. 2014. Structural diversity of ABC transporters. *J Gen Physiol,* 143**,** 419-35.
THOMAS, C. & TAMPÉ, R. 2020. Structural and Mechanistic Principles of ABC Transporters. *Annu Rev Biochem,* 89**,** 605-636.
TORRELLES, J. B. & SCHLESINGER, L. S. 2010. Diversity in Mycobacterium tuberculosis mannosylated cell wall determinants impacts adaptation to the host. *Tuberculosis (Edinb),* 90**,** 84-93.
TOUCHETTE, M. H. & SEELIGER, J. C. 2017. Transport of outer membrane lipids in mycobacteria. *Biochim Biophys Acta Mol Cell Biol Lipids,* 1862**,** 1340-1354.
UNGER, L., RONCO-CAMPAÑA, A., KITCHEN, P., BILL, R. M. & ROTHNIE, A. J. 2021. Biological insights from SMA-extracted proteins. *Biochem Soc Trans,* 49**,** 1349-1359.
VAN HEIJENOORT, J. 2001. Formation of the glycan chains in the synthesis of bacterial peptidoglycan. *Glycobiology,* 11**,** 25r-36r.
VERGNE, I., CHUA, J. & DERETIC, V. 2003. Tuberculosis toxin blocking phagosome maturation inhibits a novel Ca2+/calmodulin-PI3K hVPS34 cascade. *J Exp Med,* 198**,** 653-9.
VILLENEUVE, C., GILLERON, M., MARIDONNEAU-PARINI, I., DAFFÉ, M., ASTARIE-DEQUEKER, C. & ETIENNE, G. 2005. Mycobacteria use their surface-exposed glycolipids to infect human macrophages through a receptor-dependent process. *J Lipid Res,* 46**,** 475-83.
VINOD, V., VIJAYRAJRATNAM, S., VASUDEVAN, A. K. & BISWAS, R. 2020. The cell surface adhesins of Mycobacterium tuberculosis. *Microbiol Res,* 232**,** 126392.
VOLLMER, W., BLANOT, D. & DE PEDRO, M. A. 2008. Peptidoglycan structure and architecture. *FEMS Microbiol Rev,* 32**,** 149-67.
WALBURGER, A., KOUL, A., FERRARI, G., NGUYEN, L., PRESCIANOTTO-BASCHONG, C., HUYGEN, K., KLEBL, B., THOMPSON, C., BACHER, G. & PIETERS, J. 2004. Protein kinase

G from pathogenic mycobacteria promotes survival within macrophages. *Science,* 304**,** 1800-4.
WARD, A., REYES, C. L., YU, J., ROTH, C. B. & CHANG, G. 2007. Flexibility in the ABC transporter MsbA: Alternating access with a twist. *Proc Natl Acad Sci U S A,* 104**,** 19005-10.
WELSH, K. J., ABBOTT, A. N., HWANG, S. A., INDRIGO, J., ARMITIGE, L. Y., BLACKBURN, M. R., HUNTER, R. L. & ACTOR, J. K. 2008. A role for tumour necrosis factor-alpha, complement C5 and interleukin-6 in the initiation and development of the mycobacterial cord factor trehalose 6,6'-dimycolate induced granulomatous response. *Microbiology (Reading),* 154**,** 1813-1824.
WHEATLEY, M., CHARLTON, J., JAMSHAD, M., ROUTLEDGE, S. J., BAILEY, S., LA-BORDE, P. J., AZAM, M. T., LOGAN, R. T., BILL, R. M., DAFFORN, T. R. & POYNER, D. R. 2016. GPCR-styrene maleic acid lipid particles (GPCR-SMALPs): their nature and potential. *Biochem Soc Trans,* 44**,** 619-23.
WILSON, R. A., MAUGHAN, W. N., KREMER, L., BESRA, G. S. & FÜTTERER, K. 2004. The structure of Mycobacterium tuberculosis MPT51 (FbpC1) defines a new family of non-catalytic alpha/beta hydrolases. *J Mol Biol,* 335**,** 519-30.
WOLFE, L. M., MAHAFFEY, S. B., KRUH, N. A. & DOBOS, K. M. 2010. Proteomic definition of the cell wall of Mycobacterium tuberculosis. *J Proteome Res,* 9**,** 5816-26.
WONG, K., MA, J., ROTHNIE, A., BIGGIN, P. C. & KERR, I. D. 2014. Towards understanding promiscuity in multidrug efflux pumps. *Trends Biochem Sci,* 39**,** 8-16.
WU, Y., XIONG, D. C., CHEN, S. C., WANG, Y. S. & YE, X. S. 2017. Total synthesis of mycobacterial arabinogalactan containing 92 monosaccharide units. *Nat Commun,* 8**,** 14851.
YARKONI, E. & BEKIERKUNST, A. 1976. Nonspecific resistance against infection with Salmonella typhi and Salmonella typhimurium induced in mice by cord factor (trehalose-6,6'-dimycolate) and its analogues. *Infect Immun,* 14**,** 1125-9.
ZHAO, X., LI, G. & LIANG, S. 2013. Several affinity tags commonly used in chromatographic purification. *J Anal Methods Chem,* 2013**,** 581093.
ZUMLA, A., NAHID, P. & COLE, S. T. 2013. Advances in the development of new tuberculosis drugs and treatment regimens. *Nat Rev Drug Discov,* 12**,** 388-404.