Discovery of Novel Phage-Derived Lytic Proteins with Potential Antimicrobial Activity Against *Pseudomonas Aeruginosa*

By

Sibongile Mtimka

Submitted in fulfilment of the academic requirements of

Doctor of Philosophy in Biochemistry

Discipline of Biochemistry

School of Life Sciences

College of Agriculture, Engineering and Science

University of KwaZulu-Natal

Westville

South Africa

Supervisors: Dr Ofentse. J. Pooe

Dr Tsepo L Tsekoa

Dr Lusisizwe Kwezi

Date: June 2024

ABSTRACT

The escalation of bacterial resistance to existing antibiotics has accelerated alarmingly. This trend is further compounded by the misuse of antibiotics, leading to an intensified global challenge. Identifying novel alternatives to antibiotics to curtail the rise in resistance has led to the discovery of several alternatives, including phage-derived lytic proteins. This study investigates a solution to the urgent need for antibiotic alternatives amidst rising antibiotic resistance. It highlights the untapped potential of phage-derived lytic proteins as antibacterial agents. Sequence-based screening was used to mine an online phage database for novel putative sequences encoding lytic proteins focussing on sequences from University of Kwa-Zulu Natal (UKZN). Four open reading frames (ORFs) displaying domains typical of phage-lytic proteins were selected and synthesised. Phage lytic proteins are derivatives of phage and degrade the peptidoglycan component of the bacterial cell wall and have since gained interest as antimicrobials. Their sizes were 396 bp for Holin, 1341 bp LysA, 1374 bp LysB and 1383 bp Lysin, each with no signal peptide for cytosolic expression. Using PCR, they were subsequently sub-cloned into the pEAQ-HT for expression in *Nicotiana benthamiana.* The genes were also sub-cloned into pET-30b(+) vector for expression in *E. coli* bacteria as His_6 tagged recombinants with the tag fused to the N- or C-terminal of the proteins. While challenges were experienced with purification, the bacterial system outperformed the plant-based system in recombinant protein production of all four target proteins. Transient expression in *N. benthamiana* yielded a truncated version of three proteins and no detectable expression of the holin (SM31) protein. Lysin (SM07) was selected for detailed characterisation due to its ancestral lineage to the other proteins and the presence of the potential catalytic domains and catalytic residues for lytic proteins. Lysin (SM07) was purified successfully using IMAC with the His_6 tag on the N-terminal. The enzyme demonstrated promising antimicrobial activity against *P. aeruginosa* with a minimum inhibitory concentration of 0.004 µg/µl with no significant toxicity observed against Vero kidney cells. In a prototype, a surface disinfectant was formulated using SM07 as the sole active ingredient. The spray showed a reduced *P. aeruginosa* bacterial count of approximately 98.7%. Additionally, a 3D structure was modelled using AlphaFold revealing a cell wall binding domain (CBD), and a catalytic/enzymatic domain possessing a typical catalytic dyad (Glu228 and Glu237). Finding a CBD is unusual in Gram-negative active endolysins, raising intriguing possibilities for future antimicrobial research. This discovery will contribute to combatting nosocomial infections resulting from *P. aeruginosa.*

PREFACE

The research contained in this thesis was accomplished by the candidate while situated in the Biomanufacturing Group, Chemicals Cluster at Council of Scientific and Industrial Research, Pretoria and in the School of Life Sciences of the College of Agriculture, Engineering and Science, University of KwaZulu-Natal, Westville Campus, both in South Africa under the supervision of Dr. T.L Tsekoa, Dr. L. Kwezi and Dr. O.J Pooe from 2019 to 2024. The contents of this work have not been submitted in any form to another university and, aside from where the work of others is acknowledged in the text, the outcomes detailed are due to investigations by the candidate. The supervisors and the student have approved the submission of this thesis for examination.

Signed by the candidate:

[redacted]

Signed by the Supervisors:

[redacted] [redacted] [redacted]

Dr. O. J Pooe Dr. T.L Tsekoa Dr. L. Kwezi

Date: 20 june 2024

DECLARATION 1

I, Sibongile Mtimka, declare that:

1. The research reported in this thesis, except where otherwise indicated, and is my original research.

2. This thesis has not been submitted for any degree or examination at any other university.
3. This thesis does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.

4. This thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:

a. Their words have been re-written, but the general information attributed to them has been referenced.

b. Where their exact words have been used, then their writing has been placed in italics and inside quotation marks and referenced.

5. This thesis does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the thesis and in the References sections.

Signed: Sibongile Mtimka (219094519)

Date: June 2024

DECLARATION 2: OUTPUTS

Details of contributions of all authors to publications that I contributed to, form part or include the research presented in this thesis are provided below.

Publication List

Publication 1

Title: South Africa's indigenous microbial diversity for industrial applications: A review of the current status and opportunities

Journal: Heliyon

Status: Published

Authors: Varsha Chhiba[#], Priyen Pillay, Sibongile Mtimka, Ghaneshree Moonsamy, Lusisizwe Kwezi, Ofentse Jacob Pooe, Tsepo Lebiletsa Tsekoa[*]

[#]-First author, *Corresponding author

Authors' contributions: All authors contributed to the development and the writing of the manuscript.

Publication 2 (Chapter 2 of this thesis)

Title: An exploratory review of the potential of lytic proteins and bacteriophages for the treatment of tuberculosis

Journal: Microorganisms (MDPI)

Status: Published

Authors: Sibongile Mtimka[#], Priyen Pillay, Lusisizwe Kwezi, Ofentse Jacob Pooe, Tsepo Lebiletsa Tsekoa[*]

[#]-First author, *Corresponding author

Authors' contributions: Sibongile Mtimka wrote most of the manuscript with the help of Priyen Pillay. Lusisizwe Kwezi, Ofentse Jacob Pooe and Tsepo Lebiletsa Tsekoa reviewed. The manuscript has not been submitted by any other author for degree purposes.

Publication 3 (Chapter 4-6 of this thesis)

Title: Cloning, Expression, and Characterisation of a Novel Phage Lytic Protein with promising activity against *P. aeruginosa*

Journal: International Journal of Biotechnology

Status: Pending Submission

Authors: Sibongile Mtimka[#], Kanyane Bridgett Malatji, Patrick Opare Sakyi, Noel David Nogbou, Andrew Munyalo Musyoki, Sipho Mamputha, Vivey Phasha, Lusisizwe Kwezi, Samuel Kojo Kwofie, Ofentse Jacob Pooe, Tsepo Lebiletsa Tsekoa[*]

[#]-First author, *Corresponding author

Authors' contributions: Sibongile Mtimka conducted most of the experiments and drafted the article; Kanyane Bridget Malathi helped with cytotoxicity studies, Sipho Mamputha helped with Mass spectrometry; Vivey Phasha help with prototype development, Noel David Nogbou help with the antimicrobial studies, Patrick Opare Sakyi help with structural biology. The manuscript has not been submitted by any other author for degree purposes.

Conference Oral Presentation

- Council for Scientific and Industrial Research Emerging Researcher Symposium (CSIR-ERS) – 12th - 13th July 2022
    - Discovery of Novel Enzymes with Potential for Antimicrobial Activity
- South African Society of Biochemistry and Molecular Biology (SASBMB) – 23rd - 26th January 2022
    - Bioprospecting of Novel Enzymes with Potential of Antimicrobial Activity

Signed: Sibongile Mtimka (219094519)

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I extend my heartfelt gratitude to my supervisors for their unwavering support and invaluable contributions to my academic journey.

- Dr. Tsepo L. Tsekoa, your generosity in sharing your time, and mentorship have played a pivotal role in molding me into a more skillful researcher.
- Dr. Lusisizwe Kwezi, your guidance and the wealth of knowledge that you shared throughout my research have been instrumental.
- Dr. Ofentse J. Pooe, your support, knowledge sharing, and encouragement have strengthened my determination.

To my mentor, Dr. Konanani J. Rashamuse, your unwavering guidance, encouragement, and constant support have been indispensable in my academic journey. I would like to thank the CSIR-DSI Inter-bursary programme and the Dean of UKZN bursary, without your financial support, this would not have been possible. Special thanks to Albert Mabetha, Taola Shai, and Sharon Kgasago for their technical assistance, which contributed to completing certain tasks. I want to thank Ntate Kgama Mathiba and Gugu Mkhize for their unwavering support and nurturing presence. To the TD group, working with you has been a truly enriching experience. Thank you all for your contributions and teamwork. Let me not forget those aside the group that have helped Vivey, Nothanda, Moloko and Veshara… thank you. To Zimkitha Sigonyela, Mhleli Nenene, Bonang Tselane, Bonang's mother, and Katleho Mosia, thank you for the support, prayers and love throughout this demanding journey. Some of you even stayed with me in the lab during my late nights of running experiments. Your presence and encouragement have been invaluable.

I am forever grateful to my aunts, Khuselwa Malinga-Nopote and Nomangesi Malinga, for their unwavering support, motivation, and encouragement. I cannot thank them enough for making my Ph.D. journey bearable and instilling in me the belief that I could overcome any obstacle.

Lastly, I wish to acknowledge God's divine blessings, who have given me the strength and abilities to embark on and persevere through this challenging journey. Additionally, I thank myself for not giving up.

Thank you to all who have been part of this remarkable journey.

"Ezintweni zonke, bulelani"

ABBREVIATIONS

° C	Degree Celsius
Agro	*Agrobacterium*
BLAST	Basic Local Alignment Search Tool
BSA	Bovine Serum Albumin
CBD	Cell wall Binding Domain
Chlor	Chloramphenicol
CBD	Cell Wall-Binding Domain
CV	Column Volume
DNA	Deoxyribonucleic acid
E. coli	*Escherichia coli*
EAD	Enzymatically Active Domain
GlcNAc	*N*-acetylglucosamine
H	Hour(s)
IPTG	Isopropyl β-D-thiogalactoside
Kan	Kanamycin
Min	Minute(s)
N. Benth	*Nicotiana benthamiana*
OD_{600nm}	Optical Density
OM	Outer Membrane
ORF	Open Reading Frame
o/n	Overnight (16hr)
PG	Peptidoglycan
Rpm	Revolutions per min
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
Sec	Second(s)
Strep	Streptomycin
v/v	Volume per volume
VAPGH	Virion-Associated Peptidoglycan Hydrolase
WHO	World Health Organization

1 CHAPTER ONE: BACKGROUND

1.1 Introduction

During antiquity in Greece and Egypt, moulds and plant extracts were used to treat infections and diseases (Aminov, 2010; Dias *et al.*, 2012; Gould, 2016). This was before the accidental discovery of penicillin by Alexander Fleming in 1928, when the age of antibiotics began. Penicillin came from the mould *Penicillium notatum*, which had contaminated the culture plate and caused the lysis of *Staphylococcal* colonies. Since their discovery, antibiotics have saved millions of lives and transformed modern medicine. In the last 20 years, bacterial resistance to existing antibiotics has grown at an alarming rate. The first case of resistance was detected shortly after the release of penicillin in the 1950s (Ventola, 2015). Antibiotic resistance is becoming increasingly prevalent in bacteria, and fungal pathogens in the general population and hospital settings.

The rise of antibiotic resistance is attributed to overuse, inappropriate prescribing, extensive agricultural use, limited availability of new antibiotics, and regulatory barriers (Ventola, 2015; Aslam *et al.*, 2018a). Multidrug-resistance (MDR) results from various biochemical factors and poses a significant challenge in combatting infectious diseases (Aslam *et al.*, 2018a). Mechanisms of drug resistance include mutation or enzymatic alteration, enzymatic degradation, bypassing inhibited mechanisms, drug target overexpression, and hindrance of drug penetration, often occurring through multiple mechanisms in microbes (Pachori *et al.*, 2019). The bacterial cell wall provides structure and protection and is a common target for antimicrobial treatments. Gram-positive bacteria have a thick peptidoglycan (PG) layer, while Gram-negative bacteria feature an envelope structure with a thinner PG layer between two membranes (Mai-Prochnow *et al.*, 2016a). Peptidoglycan is a major structural component and attachment platform for proteins and polymers (Vollmer and Seligman, 2010a). Exploring new antibiotic strategies includes repurposing existing drugs, developing analogues, and innovating biologics like antibodies, antimicrobial peptides, bacteriophages, and lytic proteins (Hancock and Sahl, 2006; Boyd *et al.*, 2021)

Viruses infect and replicate within cells and are the most abundant entities on Earth. Their origin and evolution history remain uncertain, but they are distinguished by the encapsulation of their DNA/RNA in a protein capsid (Wasik and Turner, 2010). Viral structures can range from simple to complex, classified based on DNA/RNA composition (single or double-stranded, linear or circular). Their diverse structures include tailed, polyhedral, filamentous, and pleomorphic forms (Urayama *et al.*, 2015). Bacteriophages, viruses that infect bacteria,

were first discovered in 1915 and have been isolated from various bacterial communities (Nasser *et al.*, 2002; Kimura *et al.*, 2008; Clokie *et al.*, 2011). Bacteriophages sourced from different environments have been explored for treating antibiotic-resistant bacterial strains, countering the threat posed by pathogens resistant to multiple drugs (Fenton et al., 2010; Nelson et al., 2012; Lin et al., 2017). Lysins, particularly endolysins, have gained attention as potential antimicrobial agents due to their ability to target bacterial PG and thereby eradicate organisms (Burrowes et al., 2011; Lin et al., 2017; Fischetti, 2018). Countries like Georgia have spearheaded bacteriophage research and application for treating various diseases. Finding endolysins to combat drug-resistant infections is crucial, as these proteins can access the PG and kill organisms (Vollmer et al., 2008; Haddad Kashani et al., 2017).

Recombinant DNA technology offers a lot of promising avenues and can be utilised to improve stability of recombinant lytic proteins. Random mutation can be done, and genes expressed in a system that favours stability thus increasing chances of resistance against degradation or aggregation (Wang *et al.*, 2022). To increase folding and stability, site-directed mutagenesis can be used to add specific elements to the protein sequence, such as disulfide bridges or altered surface residues (Arnold and Georgiou, 2003). In addition, with recombinant DNA technology a fusion of lytic protein to a stable protein can produce chimeric proteins with improved properties(Arnold and Georgiou, 2003). By using recombinant DNA technology, recombinant lytic proteins can be optimized for higher efficiency and stability, thereby making them more attractive for therapeutic and industrial applications.

Plant-based expression systems offer a promising approach for producing lytic proteins. They present advantages such as cost-effectiveness and safety compared to bacterial systems, which face challenges in expressing cell-wall-degrading bacteriolytic proteins (Oey *et al.*, 2009). Transient expression techniques using recombinant plant viral vectors, like *Agrobacterium*-mediated delivery and virus expression machinery, have been developed for vaccines and monoclonal antibodies (Komarova et al., 2010; Tsekoa et al., 2020). Transgenic tobacco and its relatives are popular hosts due to their well-established regulatory elements and robust transformation procedures, enabling rapid expression screening, optimization, and scale-up (Fischer *et al.*, 2004; Singh *et al.*, 2021). Plant systems also ensure safety by avoiding common pathogens shared between plants and animals (Mardanova *et al.*, 2017). Expression of phage lytic proteins will not disrupt expression in transient expression systems or be toxic to host organism because the proteins are highly specific, and they target PG found in bacterial cell walls.

Recent advancements in sequencing technologies have enabled the exploration of complex microbial environmental samples, including the study of viral components (Segobola *et al.*,

2018). Unlike prokaryotes, understanding the diversity of viral species is not straightforward (Adriaenssens *et al.*, 2016). In the past, studying virus-associated bacteria relied on culture-dependent methods. However, bioinformatic tools have made it relatively easy to study and analyse biological data without culturing. Bioinformatics is a multidisciplinary field that develops methods and software tools for understanding biological data (Aamer Mehmood, 2014). Over the years, several bioinformatics tools have become available. One of the most frequently used tools of bioinformatics is sequence analysis, but their application spans over various fields. Apart from sequence analysis, their applications include drug design, molecular dynamic simulations, phylogenetic analysis, protein sequence analysis and molecular interaction and modelling (Aamer Mehmood, 2014; Diniz and Canduri, 2017). Their simplicity and capacity to generate a wealth of knowledge about a gene of interest without having to conduct expensive and time-consuming experiments have led to an increase in the use of allowed bioinformatic tools.

Sequencing techniques available have made it possible to sequence hundreds of viral and bacterial genomes at a cost-effective means per year (Klumpp et al., 2012). Sequence-based approaches rely on algorithms and databases to infer gene functions, facilitated by the availability of sequencing technologies that allow cost-effective analysis of numerous genomes (Klumpp et al., 2012; Ngara and Zhang, 2018). However, the accuracy of these techniques is dependent on accurate genome annotation and data completeness. Next-generation sequencing (NGS) technology has, however, changed this field, with platforms like Roche/454, Illumina, SOLID, Helicos, and PacBio RS II offering substantial nucleotide yields in a single run (Voelkerding *et al.*, 2009; Liu *et al.*, 2012).

The challenge posed by antibiotic threatens the efficacy of treatments that have saved countless lives since the discovery of antibiotics. The global rise is fuelled by numerous factors and has resulted in formidable multidrug-resistant bacteria. This crisis has severe global health implications, including increased morbidity, mortality, and economic costs. To address this, alternative therapeutic strategies such as improving existing drugs, developing new analogues, and innovating biologics e.g., antimicrobial peptides, bacteriophages, and lytic proteins is critical.

1.2 Scope of the Research

The global antibiotic resistance crisis has been looming for decades, and concern about the threat this poses to human health and well-being has been growing (Neu, 1992). The global antibiotic resistance crisis developed due to, amongst other things, misuse of antibiotics, evolving bacterial strains and slow progress in developing new antibiotics (Ventola, 2015;

Aslam *et al.*, 2018a). Today, numerous cases of drug-resistant and multidrug-resistant microorganisms exist in hospitals and clinics. The World Health Organisation (WHO) has drawn up a priority list of clinically relevant microorganisms resistant to antibiotics on behalf of member states (Tacconelli *et al.*, 2018). The effect of antibiotic resistance on the economy worldwide is substantial. In the US alone, economic losses are approximately $20 billion, while losses due to loss of productivity are roughly $35 billion annually (Aslam *et al.*, 2018a). The objective of the current study was to explore and develop alternatives.

1.3 Research Aims and Objectives

1.3.1 Aim

The primary aim was to screen an online phage database (https://phagesdb.org/institutions/KRIT/) with special focus on University of KwaZulu-Natal (UZKN) sequences for novel phage lytic proteins using sequence-based techniques and characterise them for their antimicrobial potential. The sequence data is from phage genomes functionally screened from environmental samples.

1.3.2 Objectives

The following specific objectives were pursued:

- To use a sequenced-based approach to identify phage-derived proteins from a public online phage database.
- To recombinantly express these proteins in *Nicotiana benthamiana* and *Escherichia coli* expression systems.
- To characterise the antimicrobial activity of the expressed and purified protein(s).
- To characterise the protein(s) utilising mass spectrometry.
- To characterise the mammalian cell cytotoxicity of the purified protein(s).
- To demonstrate the application of the purified protein in a prototype disinfectant spray.
- To model the 3D structure of the protein using a computational approach.

1.4 Thesis Organisation

The thesis can be divided into seven chapters as below.

Chapter One: Background.

This chapter contains introduction, aims, and objectives as well as the thesis organization.

Chapter Two: Literature review.

Chapter Three: Sequence-based screening for the identification and analysis of phage-derived proteins.

Chapter Four: Recombinant Protein Production of phage-derived proteins for antimicrobial activity.

Chapter Five: Antimicrobial activity screening of SM07 against ESKAPE pathogens and demonstrating its application.

Chapter Six: *In silico* modelling of SM07 to elucidate its function.

Chapter Seven: General Discussion, Conclusion and Future Studies on The Discovery of Novel Phage-Derived Lytic Proteins.

This chapter provides a general overview of the major findings and discusses the possibilities of future studies.

1.5 References

Aamer Mehmood, M. (2014) Use of Bioinformatics Tools in Different Spheres of Life Sciences. J Data Mining Genomics Proteomics 05:

Adriaenssens, E.M., van Zyl, L.J., Cowan, D.A., and Trindade, M.I. (2016) Metaviromics of Namib desert salt pans: A novel lineage of haloarchaeal salterproviruses and a rich source of ssDNA viruses. Viruses 8: 17–19.

Aminov, R.I. (2010) A Brief History of the Antibiotic Era: Lessons Learned and Challenges for the Future. Front Microbiol 1: 134.

Arnold, F.H. and Georgiou, G. (2003) Directed Evolution Library Creation. Directed Evolution Library Creation.

Aslam, B., Wang, W., Arshad, M.I., Khurshid, M., Muzammil, S., Rasool, M.H., *et al*. (2018) Antibiotic resistance: a rundown of a global crisis. Infect Drug Resist 11: 1645–1658.

Boyd, N.K., Teng, C., and Frei, C.R. (2021) Brief Overview of Approaches and Challenges in New Antibiotic Development: A Focus On Drug Repurposing. Front Cell Infect Microbiol 11:.

Burrowes, B., Harper, D.R., Anderson, J., McConville, M., and Enright, M.C. (2011) Bacteriophage therapy: Potential uses in the control of antibiotic-resistant pathogens. Expert Rev Anti Infect Ther 9: 775–785.

Clokie, M.R., Millard, A.D., Letarov, A. V, and Heaphy, S. (2011) Phages in nature. Bacteriophage 1: 31–45.

Dias, D.A., Urban, S., and Roessner, U. (2012) A Historical Overview of Natural Products in Drug Discovery. Metabolites 2: 303–336.

Diniz, W.J.S. and Canduri, F. (2017) Bioinformatics: An overview and its applications. Genetics and Molecular Research 16:.

Fenton, M., Mcauliffe, O., O'mahony, J., and Coffey, A. (2010) Recombinant bacteriophage lysins as antibacterials.

Fischer, R., Stoger, E., Schillberg, S., Christou, P., and Twyman, R.M. (2004) Plant-based production of biopharmaceuticals. Curr Opin Plant Biol 7: 152–158.

Fischetti, V.A. (2018) Development of phage lysins as novel therapeutics: A historical perspective. Viruses 10:.

Gould, K. (2016) Antibiotics: from prehistory to the present day. Journal of Antimicrobial Chemotherapy 71: 572–575.

Haddad Kashani, H., Schmelcher, M., Sabzalipoor, H., Seyed Hosseini, E., and Moniri, R. (2017) Recombinant Endolysins as Potential Therapeutics against Antibiotic-Resistant Staphylococcus aureus : Current Status of Research and Novel Delivery Strategies. Clin Microbiol Rev 31: e00071-17.

Hancock, R.E.W. and Sahl, H.G. (2006) Antimicrobial and host-defense peptides as new anti-infective therapeutic strategies. Nat Biotechnol 24: 1551–1557.

Kimura, M., Jia, Z.J., Nakayama, N., and Asakawa, S. (2008) Ecology of viruses in soils: Past, present and future perspectives. Soil Sci Plant Nutr 54: 1–32.

Klumpp, J., Fouts, D.E., and Sozhamannan, S. (2012a) Next generation sequencing technologies and the changing landscape of phage genomics. Bacteriophage 2: 190–199.

Klumpp, J., Fouts, D.E., and Sozhamannan, S. (2012b) Next generation sequencing technologies and the changing landscape of phage genomics. Bacteriophage 2: 190–199.

Komarova, T. V, Baschieri, S., Donini, M., Marusic, C., Benvenuto, E., and Dorokhov, Y.L. (2010) Transient expression systems for plant-derived biopharmaceuticals.

Lin, D.M., Koskella, B., and Lin, H.C. (2017a) Phage therapy: An alternative to antibiotics in the age of multi-drug resistance. World J Gastrointest Pharmacol Ther 8: 162–173.

Lin, D.M., Koskella, B., and Lin, H.C. (2017b) Phage therapy: An alternative to antibiotics in the age of multi-drug resistance. World J Gastrointest Pharmacol Ther 8: 162.

Liu, L., Li, Y., Li, S., Hu, N., He, Y., Pong, R., *et al*. (2012) Comparison of Next-Generation Sequencing Systems. J Biomed Biotechnol 2012: 1–11.

Mai-Prochnow, A., Clauson, M., Hong, J., and Murphy, A.B. (2016) Gram positive and Gram negative bacteria differ in their sensitivity to cold plasma. Sci Rep 6: 38610.

Mardanova, E.S., Blokhina, E.A., Tsybalova, L.M., Peyret, H., Lomonossoff, G.P., and Ravin, N. V (2017) Efficient Transient Expression of Recombinant Proteins in Plants by the Novel pEff Vector Based on the Genome of Potato Virus X. Front Plant Sci 8: 247.

Nasser, A.M., Glozman, R., and Nitzan, Y. (2002) Contribution of microbial activity to virus reduction in saturated soil.

Nelson, D.C., Schmelcher, M., Rodriguez-Rubio, L., Klumpp, J., Pritchard, D.G., Dong, S., and Donovan, D.M. (2012) Endolysins as Antimicrobials. Adv Virus Res 83: 299–365.

Ngara, T.R. and Zhang, H. (2018) Recent Advances in Function-based Metagenomic Screening. Genomics Proteomics Bioinformatics 16: 405–415.

Oey, M., Lohse, M., Kreikemeyer, B., and Bock, R. (2009) Exhaustion of the chloroplast protein synthesis capacity by massive expression of a highly stable protein antibiotic. Plant Journal 57: 436–445.

Pachori, P., Gothalwal, R., and Gandhi, P. (2019) Emergence of antibiotic resistance Pseudomonas aeruginosa in intensive care unit; a critical review. Genes Dis 6: 109–119.

Segobola, J., Adriaenssens, E., Tsekoa, T., Rashamuse, K., and Cowan, D. (2018) Exploring Viral Diversity in a Unique South African Soil Habitat. Sci Rep 8: 111.

Singh, A.A., Pillay, P., and Tsekoa, T.L. (2021) Engineering approaches in plant molecular farming for global health. Vaccines (Basel) 9:.

Tacconelli, E., Carrara, E., Savoldi, A., Harbarth, S., Mendelson, M., Monnet, D.L., *et al*. (2018) Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. Lancet Infect Dis 18: 318–327.

Tsekoa, T.L., Singh, A.A., and Buthelezi, S.G. (2020) Molecular farming for therapies and vaccines in Africa. Curr Opin Biotechnol 61: 89–95.

Urayama, S.-I., Yoshida-Takashima, Y., Yoshida, M., Tomaru, Y., Moriyama, H., Takai, K., and Nunoura, T. (2015) A New Fractionation and Recovery Method of Viral Genomes Based on Nucleic Acid Composition and Structure Using Tandem Column Chromatography. Microbes Environ.

Ventola, C.L. (2015) The antibiotic resistance crisis: part 1: causes and threats. P T 40: 277–83.

Voelkerding, K. V, Dames, S.A., and Durtschi, J.D. (2009) Next-Generation Sequencing: From Basic Research to Diagnostics. Clin Chem 55: 641–658.

Vollmer, W., Joris, B., Charlier, P., and Foster, S. (2008) Bacterial peptidoglycan (murein) hydrolases. FEMS Microbiol Rev 32: 259–286.

Vollmer, W. and Seligman, S.J. (2010) Architecture of peptidoglycan: more data and more models. Trends Microbiol 18: 59–66.

Wang, L., Ding, M.Y., Wang, J., Gao, J.G., Liu, R.M., and Li, H.T. (2022) Effects of Site-Directed Mutagenesis of Cysteine on the Structure of Sip Proteins. Front Microbiol 13:.

Wasik, B.R. and Turner, P.E. (2010) On the Biological Success of Viruses. Annu Rev Microbiol 67: 519–541.

2 CHAPTER TWO: LITERATURE REVIEW

2.1 Introduction

2.1.1 Bacterial Cell Wall

The bacterial cell envelope is a complex multi-layered structure that serves as the first and major line of defence against threats from unpredictable and often hostile environments. The cell wall provides structure to the cells, maintains homeostasis and also provides protection from extreme conditions from outside environments. Bacterial cell envelope components minimally consist of a cytoplasmic membrane (CM) and a cell wall (CW); however, Gram-negative bacteria have an addition of an outer membrane, which is outer the PG layer. On the outer membrane, Gram-negative bacteria have lipopolysaccharide (LPS) molecules that protect bacteria from harmful substances in their environment. Between the two membranes of lipopolysaccharide (LPS) molecules of phospholipid in the cell envelope of Gram-negative bacteria lays a thin layer of peptidoglycan (PG) Figure 2.1.The major component targeted in the cell wall is the peptidoglycan. PG is the major structural component of the bacterial cell wall (Schmelcher *et al.*, 2012). It serves as a platform for attaching proteins, polysaccharides and ribitol-phosphate polymers (Vollmer and Seligman, 2010). The thickness of crossed-linked PG ranges between 20 – 80 nm for Gram-positive bacteria, while the thickness in Gram-negative is lean (<10 nm) (Mai-Prochnow *et al.*, 2016b). For antimicrobial/antibiotic treatments, often the cell wall is targeted. Strategies to increase the pipeline of new antibiotic drugs include repurposing existing drugs and the development of analogues or entirely new chemistries. Novel alternatives to these predominantly small molecule-based strategies include the development of biological drugs like antibodies, antimicrobial peptides, bacteriophages and lytic proteins (including phage lytic proteins like endolysins or so-called enzybiotics) as new-generation antibiotics that could disrupt the field (Parisien *et al.,* 2008; Wittebole *et al.,* 2014).

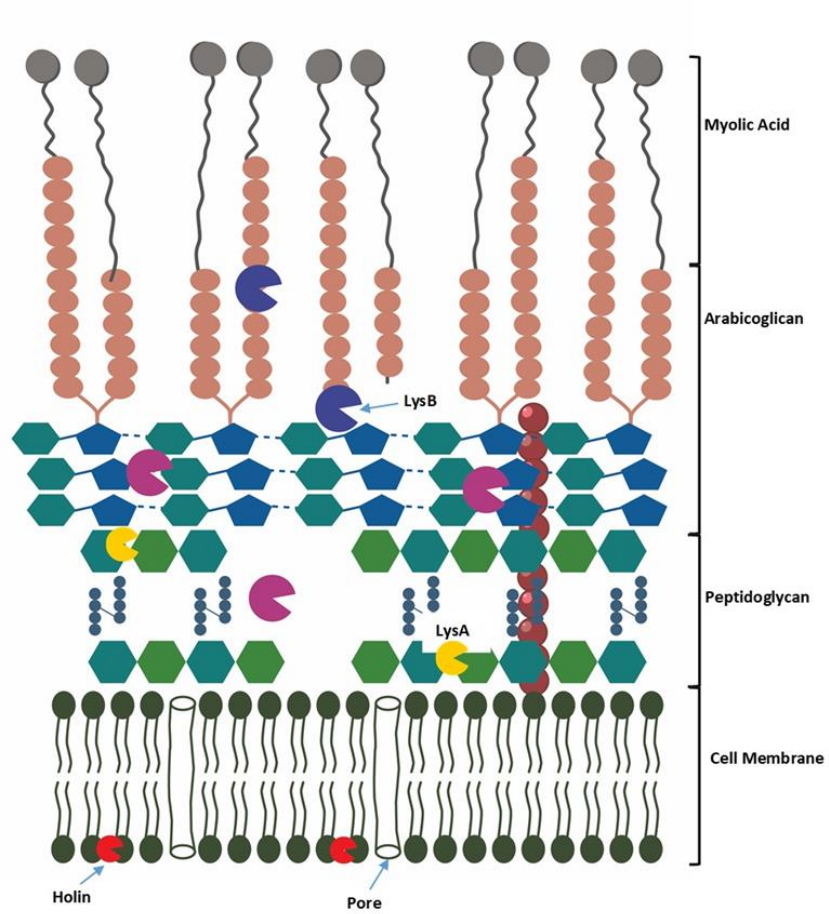


Figure 2.1 Structure of the mycobacterium cell envelope with a graphic elucidation of the biological / mode of action. LysA, which acts on the PG cell wall and hydrolyses the glycoside bonds. LysB disrupts the bonds between mycolic acids and the arabinogalactan layer. Holin attaches to the cell wall and the resultant pores are observed.

2.2 Current antibiotic resistance crisis

The current antibiotic resistance crisis poses a significant global threat to public health. The emergence of antibiotic resistance is due to but not limited to i) overuse, ii) inappropriate prescribing, iii) extensive agricultural use, and iv) few availabilities of new antibiotics/drugs. These factors have rendered many conventional antibiotics ineffective against infections. The spread of MDR is caused by a contribution of multiple biochemical factors and is presenting a serious challenge in fighting infectious diseases. Drug resistance is achieved in a number of ways, but the five major ones are: i) mutation or enzymatic alteration, ii) inactivated by enzymatic degradation and modifications, iii) bypass of the inhibited mechanism, iv) overexpression of drug target & v) drug penetration can be slowed down or inhibited by certain group of proteins and cell walls. Moreover, microbes may employ more than one mechanism, making it even more challenging to combat multidrug resistance.

The growing resistance against antibiotics is alarming as it limits treatment options, prolongs illnesses, increases healthcare costs, and escalates mortality rates. The urgency to find new antimicrobials stems from the decreasing efficacy of existing treatments. Infections that were once easily treatable are now becoming life-threatening due to antibiotic resistance. If left unchecked, this trend could lead to a future where common infections and routine medical procedures become perilous. Developing novel antimicrobials, such as endolysins, becomes crucial to combat this crisis. Endolysins offer a promising solution by targeting specific bacterial pathogens while potentially evading existing resistance mechanisms. Their ability to disrupt bacterial cell walls presents a unique approach that might help overcome the challenges posed by antibiotic resistance. Addressing this crisis demands immediate action in researching and developing alternative antimicrobial strategies, like endolysins, to ensure continued effectiveness in combating infections and safeguarding public health.

2.3 Bacteriophages and their derivatives

Bacteria-infecting viruses are known as bacteriophages (phages) and are found widely in nature. In their lytic life cycle, they kill their hosts. Over the years, this function has sparked massive interest in using them as antimicrobial agents to combat the increase in antibiotic-resistant bacteria (Rodríguez-Rubio *et al.,* 2013; Yang *et al.,* 2014). Their application can be used for numerous functions (Figure 2.2), i.e., but not limited to (i) detecting pathogens in the environment or food (O'Flaherty *et al.,* 2009), (ii) disinfection of medical apparatuses and devices, (iii) prevention of bacterial biofilm formation on industrial surfaces (O'Flaherty *et al.,* 2009; Fenton *et al.,* 2013; Zhang *et al.,* 2018); or (iv) vehicles for drug delivery (Rodríguez-Rubio *et al.,* 2013; Domingo-Calap *et al.,* 2016).

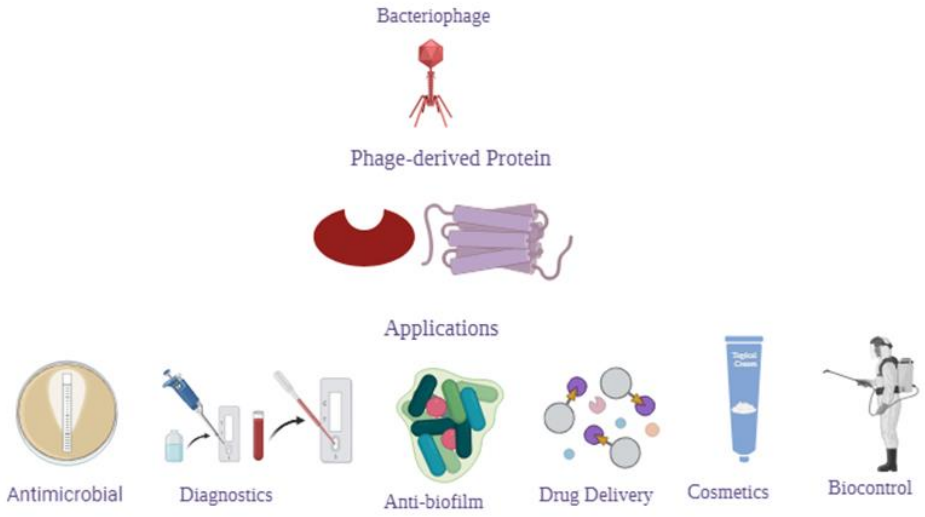


Figure 2.2:Schematic summary of potential applications of phage-derived protein.

Bacteriophage peptidoglycan (PG) degrading enzymes (endolysins) or phage-derived proteins are increasingly being investigated as antibacterial agents, with several products for topical use against Gram-positive infections already on the market (Tiberi *et al.,* 2018; Pontali *et al.,* 2019), They are classified into three groups: endolysins, holins and virion-associated peptidoglycan hydrolases (VAPGHs). This classification is based on mode of action (Gordillo Altamirano and Barr, 2019) which differs in some steps of infection cycles (Rodríguez-Rubio *et al.,* 2013). The term endolysins came after comprehending the enzymes capabilities and biochemical characteristics. These particular enzymes are synthesised and sequestered within the cell's cytoplasm, assembling phage particles (Fischetti, 2018). They pose a viable alternative to conventional antibiotics by being specific and highly effective, and the potential for the development of resistance to lysins is low (Pastagia *et al.,* 2011).

Endolysins are proposed that when administered as recombinant proteins for antibacterial function should be called enzybiotics (Rodríguez *et al.,* 2011; Gerstmans *et al.,* 2018; Heselpoth *et al.,* 2018). Given the promise of cell wall degrading enzymes in treating Gram-positive and Gram-negative infections, they have gained interest as next generation antimicrobial (van Schie *et al.,* 2021). Novel and practical options for treating drug resistant bacteria are required, as drug resistance continues to grow and is crippling the economy globally. The field of phage lytic enzymes has drastically expanded, first for Gram-positive pathogens, and later also for Gram-negative pathogens through protein engineering (Briers, 2019). While growing focus has been on lytic enzymes that target Gram-positive bacteria, there is a pressing requirement to address Gram-negative infections, where the outer membrane presents a recurring obstacle. Currently, several strategies have been developed to break the barriers posed by the outer membranes of Gram-negative bacteria to endolysins. Physical and chemical permeabilizers to permeabilize the outer membrane resulting in increased efficacy of endolysins(Briers *et al.*, 2009; Yang *et al.*, 2015)

2.3.1 Endolysins – Peptidoglycan Degrading Enzymes

Endolysins, also known as peptidoglycan hydrolases, are a class of bacteriophage-encoded enzymes that degrade the cell wall of bacterial (Lai *et al.,* 2015; Gondil *et al.,* 2020). These enzymes can also cleave the bacterial peptidoglycan (PG) from within the bacterial cell, leading to the lysis of the bacteria (Grabowski *et al.,* 2021). Bacteriophage endolysins, also known as lysins, represent a novel group of antimicrobial agents that are increasingly gaining recognition as alternative options for preventing and treating bacterial infections. Lysins consist of a catalytic component responsible for breaking specific bonds in bacterial peptidoglycan, and they share a classification with hydrolases (Pastagia *et al.,* 2011). The binding domain allows each lysin to target a specific substrate in the bacterial cell wall (usually

carbohydrate), offering some species specificity (Pastagia *et al.,* 2011). When they are applied exogenously to Gram-positive bacteria, purified native or recombinant lysins are able to degrade the cell wall of susceptible bacteria and cause log fold cell lysis within seconds to minutes (Pastagia *et al.,* 2011). Lysins typically lack signal sequences, which means they may rely on holin to access the PG (Fenton *et al.,* 2010).

2.3.2 Holins – Cytoplasmic Membrane Degrading Enzymes

Holins are small phage-encoded small transmembrane proteins that perforate the bacterial cytoplasmic membrane (Wang *et al.,* 2000; Cahill and Young, 2019). This process allows phage-encoded peptidoglycan hydrolases to act on the cell wall, resulting in host cell lysis and phage release (Kim *et al.,* 2018). Holin's synergistic role thus involves making holes in the membrane making the membrane permeable and susceptible to other lytic proteins (Wang *et al.*, 2000). Their function results in pore within the outer membrane, in turn creating pore through which endolysins use and reach the peptidoglycan (PG). Holins exhibit significant diversity, often displaying unique amino acid sequences. Generally, they are compact proteins, usually containing fewer than 150 amino acids, and are characterized by at least one anticipated transmembrane domain (TMD). Additionally, holins typically possess a hydrophilic and heavily charged C-terminal region (Catalão *et al.,* 2013). Holins are categorized into three classes I, II and III, and this is classification is based on their mechanism of action (number of TMDs the gene has) (Wang *et al.*, 2000). Class I holins are small membrane proteins that create a single hole in the bacterial cell membrane, resulting in the sudden and simultaneous release of phage progeny. In contrast, Class II holins, which are larger, form multiple holes in the membrane, leading to a more gradual and slower release of phage progeny. Class III holins are distinct from Classes I and II, as they employ a unique dual-start mechanism that allows them to release phage progeny without causing complete cell lysis. Each class of holin plays a crucial role in the successful in the bacteria cell lysis and in also aiding in the function of endolysins to reaching the PG.

2.3.3 VAPGH –Virion-Associated Peptidoglycan Hydrolases

VAPGHs act in the first phase of the lytic cycle, while endolysins act in the late phase of gene expression in the lytic cycle (Rodríguez-Rubio *et al.,* 2013). These phage lytic proteins can target Gram-positive and Gram-negative bacteria (Gordillo Altamirano and Barr, 2019), making them good candidates for combating drug-resistant bacterial-causing agents. The primary purpose of VAPGHs is to assist in the release of newly formed phage progeny from the host bacterial cell by degrading the bacterial cell wall, which is composed of peptidoglycan. Their action is part of a coordinated process that involves other phage-encoded proteins such as holins and endolysins and unlike endolysins (Rodríguez-Rubio *et al.*, 2016; Gordillo

Altamirano and Barr, 2019). VAPGH do not possess any cell-wall binding domains (CBD), but they have highly conserved domains (Aslam et al., 2021).

2.4 Structure and mode of action of endolysins

Endolysin typically exhibit distinct structural differences; they can either be globular or modular and this is based on whether they target Gram-negative or Gram-positive bacteria. These differences reflect the difference between the cell wall compositions of these bacterial groups. The role of endolysins is to perform both substrate recognition and enzymatic hydrolysis (Loessner, 2005). Endolysins are usually composed of two functional domains, the enzymatic domain and the cell wall binding domain (CBD) (Loessner, 2005), however not all endolysins harbour CBD (Schmelcher *et al.*, 2012). Gram-positive infecting endolysins have a modular structure with cell wall binding domain on the C-terminal with the catalytic domain on the N-terminal (Rodríguez-Rubio et al., 2016; Gerstmans et al., 2018). These endolysin can have one or two catalytic domains separated by a linker sequence. Endolysins that infect Gram-negative have an inverted molecular orientation compared to the Gram-positive infecting endolysins. They have a globular organisation with one catalytic domain (Gerstmans et al., 2018); however, there are reports that show modular Gram-negative endolysins. The modular organisation has one or two cell wall binding domain on the N-terminal and one catalytic domain on the C-terminal (Gerstmans et al., 2018). The catalytic domain is responsible for the hydrolytic cleavage of highly conserved bonds in the peptidoglycan (Rodríguez-Rubio *et al.*, 2016; Briers, 2019).

The cell wall binding domain function is to bind to the peptidoglycan, which is needed for efficient cleavage of the bacterial cell wall (Shrivastava *et al.*, 2019) and it is usually found on the 3' of the gene. It is critical for the mode of action of endolysins and also plays a significant role in their specificity. Unlike the EAD, the CBD amino acid sequence is short and several of those amino acids are responsible for interacting with certain components of the peptidoglycan (Schmelcher *et al.*, 2012). Different endolysins have different CBDs that bind to specific peptidoglycan structures. While endolysin specificity usually targets a single species (Fenton et al., 2010), two lysins isolated from a *Pseudomonas* phages (KZ144 and EL188) are exceptions with their CBD on the N-terminus and an EAD at the C-terminus (Gondil *et al.*, 2020). And they exhibit strong affinity for *Pseudomonas aeruginosa* cell walls, however, one of them (KZ144) can also identify peptidoglycan (PG) in other Gram-negative bacteria.

The EAD of endolysins is typically modular, meaning that it is made up of several different domains that work together to catalyze the hydrolysis reaction (Nelson *et al.*, 2012b; Schmelcher *et al.*, 2012). The enzymatic active domain of endolysins is responsible for breaking down the peptidoglycan layer of bacterial cell walls, leading to bacterial cell lysis.

Peptidoglycan is a polymer made up of sugar and amino acid chains that forms the rigid outer layer of bacterial cells. The domain comprises of enzymes such as peptidases or glycosidases which target specific components within the PG structure weakening the cell wall and ultimately causing the bacterial cell to burst. Based on the target of the component of the PG, EAD can be categorized into five groups targeting certain peptide bonds within the peptidoglycan structure (i) *N*-acetylmuramidases (ii) lytic transglycosylases (iii) *N*-acetyl-β-D-glucosaminidases, (iv) *N*-acetyl-β-d-glucosaminidases and endopeptidases (Fenton et al., 2010; Gerstmans et al., 2018; Heselpoth et al., 2018). *N*-acetylmuramidases, and lytic transglycosylases both cleave the *N*-acetylmuramoyl-β-1,4-*N*-acetylglucosamine bond, which is one of the two alternating glycosidic bonds of the sugar strand, *N*-acetyl-β-D-glucosaminidases cut the other glycosidic bond (*N*-acetylglucosaminyl-β-1,4-*N*-acetylmuramine) in the sugar strand, *N*-acetylmuramoyl-L-ala- nine amidases- hydrolyze the amide bond between MurNAc and the first amino acid of the stem peptide (L- alanine) and the Endopeptidases- These enzymes cleave within the peptides making up the interconnecting stem portion of the peptidoglycan units

2.5 Applications of Phage Lytic Proteins

Endolysin have several potential applications in various field is this can be attributed to their mechanisms of action and gene arrangement. The specificity of applications of endolysins varies on the enzyme's characteristics, target bacteria, and the intended use. Their ability to kill target bacteria makes them potential alternatives to antibiotics and also alternatives to the disruption of biofilms (Table 2.1). They can also be applied for control and contamination in the food industry as well as in the agricultural space. The binding domain of the endolysins can be used in diagnostic tests to detect the presence of specific bacterial pathogens. They have use in cosmetic for topical application to combat problems arising from bacterial infections.

Table 2.1: List of phage-derived proteins with the target pathogens application.

Name	Applications	Target Pathogen	Result summary	Reference
LysH5	Biofilm	*Staphylococcus*	Bacterial load reduction of S. aureus and S. epidermidis strains 1–3 log units	(Gutiérrez *et al.*, 2018)
P. aeruginosa M4 phage	Biofilm	*P. aeruginosa*	Reduction of bacteria in 24 hs from 6.87 to 4.03 log	(Łusiak-Szelachowska *et al.*, 2020)
D29 phage	Therapeutics	*M. tuberculosis*	Demonstrates prophylactic efficacy against M. tuberculosis infection through inhalation administration.	(Düzgüneş *et al.*, 2021)
PM-477	Biofilm	*Gardnerella strains*	Eliminated Gardnerella bacteria, breaking down biofilms in the vaginal microbiome without harm.	(Landlinger *et al.*, 2021)
T7, P4, PDRPv, Bo4, Bxz2	*Therapeutics*	*M. smegmatisi, M. tuberculosis, M. bovis*	Reduced infections in multiple organs and lesions in some	(Bajpai *et al.*, 2018; Azimi *et al.*, 2019)
lys08/PHB08	Biofilms	*Enterococcus faecalis*	Effectively fought E. faecalis biofilms, especially with added Mn2+ ions.	(Yang *et al.*, 2020)
Ply500	Food Protection	*Listeria Monocytogenes*	Reduced Listeria monocytogenes by 4 logs in Iceberg Lettuce	(Stone *et al.*, 2022; Nazir *et al.*, 2023)
CBD-P35, and CBD500	Biocontrol/Pathogen detection	*Listeria Monocytogenes*	Detection of Listeria strains using cell wall binding domains in a flourescence assay	(Schmelcher *et al.*, 2010)
phi11	Veterinary and Food	*Staphylococcus spp*	The purified protein efficiently lysed untreated staphylococcal mastitis pathogens	(Donovan *et al.*, 2006)
PE204	Biocontrol	*R. solanacearum*	Application of phage PE204 stopped bacterial wilt.	(Bae *et al.*, 2012)
PlyG	Biocontrol	*Bacillus anthracis*	Able to break down vegetative cells and attack the endospores.	(Kikkawa *et al.*, 2008)
PlyF307	planktonic and biofilm	*Acinetobacter baumannii*	Displayed strong antimicrobial action (>5-log reduction)	(Lood *et al.*, 2015)
(ENDL) from P100.1 C	Skin Care	Propionibacterium acnes	Exhibits antimicrobial activity	(Collins, 2023)
Staphefekt SA.100	Cosmetics	*Staphylococcus aureus*	Successfully treated chronic and recurrent S. aureus-related skin conditions	(Totté *et al.*, 2017)

2.5.1 Therapeutic Potential

The susceptibility of a pathogen infection and the effectiveness of any antimicrobial therapeutic agent including phages are influenced by the individual's immune status. Some

mathematical models of phage and bacteria interactions propose cooperation of the immune system for successful treatment using phage therapy (Chatain-ly, 2014; Viertel *et al.,* 2014). Phage therapy for medical products has been withheld in some countries due to the demand for a high level of clinical evidence regardless of the intrinsic efforts of academia, regulators and biotech companies. One of the responses to high doses in phage treatment may be inflammation, triggered by the release of endotoxins. Due to the nature of phages, they stimulate an adaptive immune response. Their effects are typically caused by the toxicity of contaminants (i.e. LPS) from the preparation of the phages (Krut and Bekeredjian-Ding, 2018). For immunocompromised patients, phage therapy may offer little benefit in targeting nosocomial infections. Notably, the amounts of LPS do not exceed that of antibiotic treatment (Krut and Bekeredjian-Ding, 2018; Luong *et al.,* 2020). CBDs of endolysins have also been explored as potential therapeutic agents. CBDs can be used to target and bind to specific bacterial cells, without harming human cells. This makes them a promising new approach for the treatment of bacterial infections. They can also be used in combination therapy to enhance the efficiency of the standard drug administered or reduce the dosage required for effective treatment. Holins can also be used to aid in drug delivery as they function to disrupt membrane by creating holes in the membrane.

As illustrated in Figure 2.1, the cell wall comprises of a thick layer of PG that is intrinsically linked to arabinogalactan that has been esterified with mycolic acids. Developing lytic phages that meet regulatory requirements and match clinical needs is crucial to combat bacterial and antibiotic-resistant pathogens. In the age of multi-drug resistance, phage lytic protein treatment can be seen as a complement or alternative to antibiotics. Most phage species employ two major protein classes during the lysis of the bacterial host. One is the transmembrane protein, holin, and the other is a peptidoglycan cell wall hydrolase called endolysin (lysin). These two proteins work together to trigger the lysis of the bacterial cell (Roach and Donovan, 2015; Puiu and Julius, 2019). Phage lytic enzymes can be explored not only for their benefits within healthcare but in any scenario that involves harmful bacteria.

2.5.2 Diagnostics

Rapid and accurate pathogen identification is crucial for disease prevention and treatment, particularly in the context of food safety. Phage derived proteins have been used for the detection of foodborne bacteria using phage proteins. Some of the pathogens responsible for substantial economic losses, foodborne illnesses, and even fatalities *include L. monocytogenes, S. aureus, C. perfringens,* and *Salmonella*. The food industry requires rapid and accurate pathogen detection tool to detect microbial contamination swiftly and reliably. The current methods used for pathogen detection in the food industry are known for being

time-consuming and requiring substantial manual labour which is one of the most notable limitations is with the PCR-based approach. PCR produces positive results because it cannot differentiate between live and nonviable cells. This distinction is particularly critical in food diagnostics as the inability to poses a significant challenge in food safety assessments.

Endolysins can successfully be used for rapid detection of bacterial pathogens. Because of their specificity (CBD) and degrade the peptidoglycan layer found in bacterial cell walls, they can be combined with other detection methods, such as fluorescence labelling or biosensors, to help identify the presence of specific bacteria quickly and accurately. Endolysins could be incorporated into portable diagnostic devices for swift, on-site pathogen detection, a feature especially beneficial in healthcare setups where timely pathogen identification holds paramount importance. Their capacity to precisely pinpoint bacteria and distinguish between viable and non-viable cells renders them valuable assets in the realms of pathogen detection and ensuring food safety. CBD fused with a fluorescence tag and can be expressed together and be immobilised into a kit. One other detection kit with results in 15 min was was developed using a PlyG lysin (Fenton *et al.,* 2010), this shows the turnaround time at which results can also be available in diagnosis when using lytic proteins in detection kits.

2.5.3 Agricultural Industry

Within the agricultural space, studies on lysins have shown promise in controlling and eliminating phytopathogenic bacteria. A study by (Fenton et al., (2010), they have demonstrated protection against a Gram-negative bacterium that cause potatoes to rot (soft). Another demonstration of lysins capabilities was done by Kim et al., (2004), where they demonstrated the inhibition of infection in fruit. They found that the preparations used contained catalytic domains essential for lytic function but determined that the binding domain was the result of the protection. Endolysins offer potential avenue for managing bacterial infection even in plants as well as offering alternatives for bactericides (Fenton *et al.,* 2010). Endolysin can be incorporated into bactericides for less harmful to nature alternatives and yet effective and specific targeting.

2.6 Phage Lytic protein safety and specificity

From a safety perspective, the high specificity of endolysins is among their most beneficial properties in food safety and medical applications. These enzymes specifically destroy a target pathogen without affecting the desired commensal microflora, conferring a significant advantage over many commonly used antibiotics or chemical preservatives (Schmelcher et al., 2012 and Puiu and Julius, 2019). A significant concern with systemic administration of lysins in humans or animals is the release of proinflammatory cellular debris associated with bacterial lysis, such as teichoic acids, lipoteichoic acids, and PG, which can lead to serious adverse effects such as bacterial infection and failure of several organs (Nau and Eiffert, 2002). Due to their proteinaceous nature, endolysins are noncorrosive and biodegradable – this is considered an additional advantage (Nelson et al., 2012; Schmelcher et al., 2012; Heselpoth et al., 2018).

Due to their proteinaceous nature, it is not surprising that antibodies can be raised against these enzymes (Schmelcher *et al.,* 2012; Gutiérrez *et al.,* 2018; Vázquez *et al.,* 2018; Gondil *et al.,* 2020). Mice exposed to the pneumococcal phage lysin Cpl developed antibodies (IgG) against it, but these antibodies didn't seem to affect how Cpl inhibited the bacteria *in vivo*, suggesting they weren't effective at stopping it. (Loeffler *et al.,* 2003). In addition, the use of lysins with two catalytic domains, each with different peptide glycan specificity, may further reduce the likelihood of resistant strains emerging (Fenton et al., 2010).

2.7 Regulatory Pathways for Approving Endolysins as Therapeutic

Recombinantly produced endolysins are usually classified as biologics and approval of biologics is more stringent than medical devise such as diagnostics kits (Morrow and Felcone, 2004). A comprehensive preclinical and clinical testing to ensure safety and efficacy is needed. Endolysins, with proven efficacy in clinical trials and approval could pave the way, particularly in healthcare that is grappling with resistant bacteria that traditional antibiotics struggle to treat. The market for endolysins as therapeutics is expected to continue growing and this is driven by the continuing prevalence of antibiotic resistance bacteria (Haddad Kashani *et al.*, 2017).

The FDA in the US or the EMA in Europe oversee and assess trails and data generated from the studies before approving any drug (Joppi *et al.*, 2020). The rigorous process applies for phage therapy as well as phage derivatives such as endolysins. Although the approval of endolysins as antimicrobials into the market depends on the specific characteristics of the endolysins, the intended route of administration and ensuring safety in humans. South Africa's (SA) own regulation body known as South African Health Products Regulatory Authority (SAHPRA) approves any drugs in SA just as the FDA and recombinant proteins fall under

biological medicines in our regulatory body (SAHPRA). For any biological medicine, the process is as rigorous as the FDA. Compliance with SAHPRA is essential for approval. Guidelines from SAHPRA are also informed by the Human and Veterinary Medicines, European Union (EU) guidelines and the WHO Guideline for approvals (South African Health Products Regulatory Authority, 2022).

2.8 References

Aslam, B., Arshad, M.I., Aslam, M.A., Muzammil, S., Siddique, A.B., Yasmeen, N., *et al*. (2021) Bacteriophage Proteome: Insights and Potentials of an Alternate to Antibiotics. Infect Dis Ther.

Aslam, B., Arshad, M.I., Aslam, M.A., Muzammil, S., Siddique, A.B., Yasmeen, N., Khurshid, M., Rasool, M., Ahmad, M., Rasool, M.H., 2021. Bacteriophage proteome: insights and potentials of an alternate to antibiotics. Infectious Diseases and Therapy 10, 1171-1193.

Azimi, T., Mosadegh, M., Nasiri, M.J., Sabour, S., Karimaei, S., and Nasser, A. (2019) Phage therapy as a renewed therapeutic approach to mycobacterial infections: A comprehensive review. Infect Drug Resist 12: 2943–2959.

Bae, Y.J., Wu, J., Lee, H.J., Jo, E.J., Murugaiyan, S., Chung, E., and Lee, S.W. (2012) Biocontrol potential of a lytic bacteriophage PE204 against bacterial wilt of tomato. J Microbiol Biotechnol 22: 1613–1620.

Bajpai, U., Mehta, A.K., Eniyan, K., Sinha, A., Ray, A., Virdi, S., *et al*. (2018) Isolation and characterization of bacteriophages from india, with lytic activity against Mycobacterium tuberculosis. Can J Microbiol 64: 483–491.

Briers, Y. (2019) Phage Lytic Enzymes. Viruses 11: 113.

Briers, Y., Schmelcher, M., Loessner, M.J., Hendrix, J., Engelborghs, Y., Volckaert, G., and Lavigne, R. (2009) The high-affinity peptidoglycan binding domain of Pseudomonas phage endolysin KZ144. Biochem Biophys Res Commun 383: 187–191.

Cahill, J., Young, R., 2019. Phage lysis: multiple genes for multiple barriers. Advances in virus research 103, 33-70.

Catalao, M.J., Gil, F., Moniz-Pereira, J., Sao-Jose, C., Pimentel, M., 2013. Diversity in bacterial lysis systems: bacteriophages show the way. FEMS microbiology reviews 37, 554-571.

Catalão, M.J., Pimentel, M., 2018. Mycobacteriophage lysis enzymes: targeting the mycobacterial cell envelope. Viruses 10, 428.

Chatain-Ly, M.H., 2014. The factors affecting effectiveness of treatment in phages therapy. Frontiers in microbiology 5, 51.

Collins, K.M. (2023) Production and characterization of a phage endolysin with putative antimicrobial activity against Cutibacterium acnes.

Domingo-Calap, P., Georgel, P., Bahram, S., 2016. Back to the future: bacteriophages as promising therapeutic tools. Hla 87, 133-140.

Donovan, D.M., Lardeo, M., and Foster-Frey, J. (2006) Lysis of staphylococcal mastitis pathogens by bacteriophage phi11 endolysin. FEMS Microbiol Lett 265: 133–139.

Düzgüneş, N., Sessevmez, M., and Yildirim, M. (2021) Bacteriophage therapy of bacterial infections: The rediscovered frontier. Pharmaceuticals 14: 1–16.

Fenton, M., Mcauliffe, O., O'mahony, J., and Coffey, A. (2010) Recombinant bacteriophage lysins as antibacterials.

Fischetti, V.A., 2018. Development of phage lysins as novel therapeutics: a historical perspective. Viruses 10, 310.

Gerstmans, H., Criel, B., and Briers, Y. (2018) Synthetic biology of modular endolysins. Biotechnol Adv 36: 624–640.

Gondil, V.S., Harjai, K., and Chhibber, S. (2020) Endolysins as emerging alternative therapeutic agents to counter drug-resistant infections. Int J Antimicrob Agents 55:.

Gordillo Altamirano, F.L. and Barr, J.J. (2019) Phage therapy in the postantibiotic era. Clin Microbiol Rev 32:.

Grabowski, Ł., Łepek, K., Stasiłojć, M., Kosznik-Kwaśnicka, K., Zdrojewska, K., Maciąg-Dorszyńska, M., Węgrzyn, G., Węgrzyn, A., 2021. Bacteriophage-encoded enzymes destroying bacterial cell membranes and walls, and their potential use as antimicrobial agents. Microbiological research 248, 126746.

Guo, S., Ao, Z., 2012. Phage in the diagnosis and treatment of tuberculosis. Front Biosci 17, 2691-2697.

Gutiérrez, D., Fernández, L., Rodríguez, A., and García, P. (2018) Are phage lytic proteins the secret weapon to kill staphylococcus aureus? mBio 9:.

Haddad Kashani, H., Schmelcher, M., Sabzalipoor, H., Seyed Hosseini, E., and Moniri, R. (2017) Recombinant Endolysins as Potential Therapeutics against Antibiotic-Resistant Staphylococcus aureus: Current Status of Research and Novel Delivery Strategies. Clin Microbiol Rev 31:.

Heselpoth, R.D., Swift, S.M., Linden, S.B., Mitchell, M.S., and Nelson, D.C. (2018) Enzybiotics: Endolysins and Bacteriocins.

Joppi, R., Bertele, V., Vannini, T., Garattini, S., and Banzi, R. (2020) Food and Drug Administration vs European Medicines Agency: Review times and clinical evidence on novel drugs at the time of approval. Br J Clin Pharmacol 86: 170.

Kikkawa, H.S., Ueda, T., Suzuki, S.I., and Yasuda, J. (2008) Characterization of the catalytic activity of the gamma-phage lysin, PlyG, specific for Bacillus anthracis. FEMS Microbiol Lett 286: 236–240.

Kim, M., Jo, Y., Hwang, Y.J., Hong, H.W., Hong, S.S., Park, K., Myung, H., 2018. Phage-antibiotic synergy via delayed lysis. Applied and environmental microbiology 84, e02085-02018.

Kim, W.S., Salm, H., and Geider, K. (2004) Expression of bacteriophage phiEa1h lysozyme in Escherichia coli and its activity in growth inhibition of Erwinia amylovora. Microbiology (Reading) 150: 2707–2714.

Krut, O., Bekeredjian-Ding, I., 2018. Contribution of the immune response to phage therapy. The Journal of Immunology 200, 3037-3044.

Lai, M.-J., Liu, C.-C., Jiang, S.-J., Soo, P.-C., Tu, M.-H., Lee, J.-J., Chen, Y.-H., Chang, K.-C., 2015. Antimycobacterial activities of endolysins derived from a mycobacteriophage, BTCU-1. Molecules 20, 19277-19290.

Landlinger, C., Tisakova, L., Oberbauer, V., Schwebs, T., Muhammad, A., Latka, A., *et al*. (2021) Engineered Phage Endolysin Eliminates Gardnerella Biofilm without Damaging Beneficial Bacteria in Bacterial Vaginosis Ex Vivo. Pathogens 10: 1–19.

Lin, K., 2018. Exploiting Bacterial Lysis for Tuberculosis Drug Target Validation and Vaccine Development. Weill Medical College of Cornell University.

Loeffler, J.M., Djurkovic, S., Fischetti, V.A., 2003. Phage lytic enzyme Cpl-1 as a novel antimicrobial for pneumococcal bacteremia. Infection and immunity 71, 6199-6204.

Loessner, M.J. (2005) Bacteriophage endolysins - Current state of research and applications. Curr Opin Microbiol 8: 480–487.

Lood, R., Winer, B.Y., Pelzek, A.J., Diez-Martinez, R., Thandar, M., Euler, C.W., *et al*. (2015) Novel phage Lysin capable of killing the multidrug-resistant gram-negative bacterium Acinetobacter Baumannii in a mouse bacteremia model. Antimicrob Agents Chemother 59: 1983–1991.

Luong, T., Salabarria, A.-C., Roach, D.R., 2020. Phage therapy in the resistance era: where do we stand and where are we going? Clinical Therapeutics.

Łusiak-Szelachowska, M., Weber-Dąbrowska, B., and Górski, A. (2020) Bacteriophages and Lysins in Biofilm Control. Virol Sin 35: 125.

Mai-Prochnow, A., Clauson, M., Hong, J., and Murphy, A.B. (2016) Gram positive and Gram negative bacteria differ in their sensitivity to cold plasma. Sci Rep 6: 38610.

Morrow, T., & Felcone, L. H. (2004). Defining the difference: What Makes Biologics Unique. Biotechnology Healthcare, 1(4).

Nau, R., Eiffert, H., 2002. Modulation of release of proinflammatory bacterial compounds by antibacterials: potential impact on course of inflammation and outcome in sepsis and meningitis. Clinical microbiology reviews 15, 95-110.

Nazir, A., Xu, X., Liu, Y., and Chen, Y. (2023) Phage Endolysins: Advances in the World of Food Safety. Cells 12: 2169.

Nelson, D.C., Schmelcher, M., Rodriguez-Rubio, L., Klumpp, J., Pritchard, D.G., Dong, S., and Donovan, D.M. (2012) Endolysins as Antimicrobials. Adv Virus Res 83: 299–365.

O'Flaherty, S., Ross, R.P., Coffey, A., 2009. Bacteriophage and their lysins for elimination of infectious bacteria. FEMS microbiology reviews 33, 801-819.

Pastagia, M., Euler, C., Chahales, P., Fuentes-Duculan, J., Krueger, J.G., Fischetti, V.A., 2011. A novel chimeric lysin shows superiority to mupirocin for skin decolonization of methicillin-resistant and-sensitive Staphylococcus aureus strains. Antimicrobial agents and chemotherapy 55, 738-744.

Pastagia, M., Schuch, R., Fischetti, V.A., Huang, D.B., 2013. Lysins: the arrival of pathogen-directed anti-infectives. Journal of medical microbiology 62, 1506-1516.

Pontali, E., Raviglione, M.C., Migliori, G.B., 2019. Regimens to treat multidrug-resistant tuberculosis: past, present and future perspectives. European Respiratory Review 28.

Puiu, M., Julius, C., 2019. Bacteriophage gene products as potential antimicrobials against tuberculosis. Biochemical Society Transactions 47, 847-860.

Roach, D.R., Donovan, D.M., 2015. Antimicrobial bacteriophage-derived proteins and therapeutic applications. Bacteriophage 5, e1062590.

Roach, D.R., Leung, C.Y., Henry, M., Morello, E., Singh, D., Di Santo, J.P., Weitz, J.S., Debarbieux, L., 2017. Synergy between the host immune system and bacteriophage is

essential for successful phage therapy against an acute respiratory pathogen. Cell host & microbe 22, 38-47. e34.

Rodríguez, L., Martínez, B., Zhou, Y., Rodríguez, A., Donovan, D.M., García, P., 2011. Lytic activity of the virion-associated peptidoglycan hydrolase HydH5 of Staphylococcus aureus bacteriophage vB_SauS-phiIPLA88. BMC microbiology 11, 1-11.

Rodríguez-Rubio, L., Gutiérrez, D., Donovan, D.M., Martínez, B., Rodríguez, A., and García, P. (2016) Phage lytic proteins: Biotechnological applications beyond clinical antimicrobials. Crit Rev Biotechnol 36: 542–552.

Rodríguez-Rubio, L., Martínez, B., Donovan, D.M., Rodríguez, A., García, P., 2013. Bacteriophage virion-associated peptidoglycan hydrolases: potential new enzybiotics. Critical reviews in microbiology 39, 427-434.

Schmelcher, M., Donovan, D.M., and Loessner, M.J. (2012) Bacteriophage endolysins as novel antimicrobials. Future Microbiol 7: 1147.

Schmelcher, M., Shabarova, T., Eugster, M.R., Eichenseher, F., Tchang, V.S., Banz, M., and Loessner, M.J. (2010) Rapid multiplex detection and differentiation of Listeria cells by use of fluorescent phage endolysin cell wall binding domains. Appl Environ Microbiol 76: 5745–5756.

Shrivastava, A., Shrivastava, N., and Singh, P.K. (2019) Enzymes in Pharmaceutical Industry. Enzymes in Food Biotechnology 591–602.

South African Health Products Regulatory Authority (2022) Biological Medicine Amendment Guideline.

Stone, E., Pennone, V., Reilly, K., Grant, I.R., Campbell, K., Altermann, E., and McAuliffe, O. (2022) Inhibition of Listeria monocytogenes by Phage Lytic Enzymes Displayed on Tailored Bionanoparticles. Foods 11:.

Tiberi, S., du Plessis, N., Walzl, G., Vjecha, M.J., Rao, M., Ntoumi, F., Mfinanga, S., Kapata, N., Mwaba, P., McHugh, T.D., 2018. Tuberculosis: progress and advances in development of new drugs, treatment regimens, and host-directed therapies. The Lancet Infectious Diseases 18, e183-e198.

Totté, J.E.E., van Doorn, M.B., and Pasmans, S.G.M.A. (2017) Successful Treatment of Chronic Staphylococcus aureus- Related Dermatoses with the Topical Endolysin Staphefekt SA.100: A Report of 3 Cases. Case Rep Dermatol 9: 19–25.

van Schie, L., Borgers, K., Michielsen, G., Plets, E., Vuylsteke, M., Tiels, P., Festjens, N., Callewaert, N., 2021. Exploration of synergistic action of cell wall-degrading enzymes against Mycobacterium tuberculosis. bioRxiv.

Vázquez, R., García, E., García, P., 2018. Phage lysins for fighting bacterial respiratory infections: a new generation of antimicrobials. Frontiers in immunology 9, 2252.

Viertel, T.M., Ritter, K., Horz, H.-P., 2014. Viruses versus bacteria—novel approaches to phage therapy as a tool against multidrug-resistant pathogens. Journal of Antimicrobial Chemotherapy 69, 2326-2336.

Vollmer, W. and Seligman, S.J. (2010) Architecture of peptidoglycan: more data and more models. Trends Microbiol 18: 59–66.

Wang, I.N., Smith, D.L., and Young, R. (2000) Holins: The protein clocks of bacteriophage infections. Annu Rev Microbiol 54: 799–825.

Yang, D., Chen, Y., Sun, E., Hua, L., Peng, Z., and Wu, B. (2020) Characterization of a Lytic Bacteriophage vB_EfaS_PHB08 Harboring Endolysin Lys08 against Enterococcus faecalis Biofilms. Microorganisms 8: 1–12.

Yang, H., Wang, M., Yu, J., and Wei, H. (2015) Antibacterial Activity of a Novel Peptide-Modified Lysin Against Acinetobacter baumannii and Pseudomonas aeruginosa. Front Microbiol 6: 1471.

Yang, H., Yu, J., Wei, H., 2014. Engineered bacteriophage lysins as novel anti-infectives. Frontiers in microbiology 5, 542.

Zhang, Y., Cheng, M., Zhang, H., Dai, J., Guo, Z., Li, X., Ji, Y., Cai, R., Xi, H., Wang, X., 2018. Antibacterial effects of phage lysin LysGH15 on planktonic cells and biofilms of diverse staphylococci. Applied and Environmental Microbiology

3 CHAPTER THREE: SEQUENCE-BASED SCREENING FOR THE IDENTIFICATION OF PHAGE-DERIVED PROTEINS

Abstract

The need for alternatives to combat antibiotic-resistant microorganisms is well-established due to the increased prevalence of antibiotic resistance. Therefore, there is a clear necessity to introduce novel approaches to address this pressing issue. Phage-derived lytic proteins (endolysins) have gained attention as viable alternatives due to their ability to degrade the peptidoglycan (PG) layer. Using bioinformatics, an online phage database with special focus on UKZN sequences was screened for phage-derived lytic proteins. Four open reading frames (ORFs) were selected based on their sequence identity and homology to phage-derived enzymes. These sequences displayed domains typical of phage-derived lytic proteins such as lysozyme-like, amidase, and PG-binding domains. The four genes were synthesised and cloned into a pUC57 vector for plasmid propagation and subsequently subcloned into the pEAQ-HT for expression. They were subcloned for expression with His_6 nucleotides sequences on the 5' end of the gene and their sizes were 396 bp, 1341 bp, 1374 bp and 1383 bp, each without signal peptide for cytosolic expression. The results of the phylogenetic analysis revealed all the genes were from the *Mycobacterium* genus and three of the four shared domains critical for phage lytic proteins. This observation strengthened our hypothesis that they function as endolysins, capable of specific targeting and disruption of the cell walls unique to certain bacterial species, thereby highlighting their potential in various applications, including being formulated as a disinfectant.

3.1 Introduction

The escalating antibiotic resistance crisis demands innovative antimicrobial approaches. Phage-derived proteins which are abundant and ubiquitous in prokaryotic environments (Clokie *et al.*, 2011), specifically endolysins, exhibit precise targeting of the bacterial cell walls, showcasing their potential as potent antimicrobials. Exploring these proteins, especially through gene mining of a phage database offers a promising opportunity. Phages have long been used to treat bacterial infections (Reina and Reina, 2018; Górski et al., 2019; Morrisette et al., 2020). Their derivatives have also gained interest (Aslam et al., 2021) and the pursuit of identifying phage lytic proteins to combat antibiotic-resistant bacteria holds promise, potentially transforming antimicrobial therapies.

Recombinant DNA technology has revolutionised biology, medicine, and biotechnology by allowing the manipulation of genes and protein material with unprecedented precision (Adrio and Demain, 2014; Nwankwo and Abalaka, 2014; Mtimka *et al.*, 2020). After identifying target genes from phage genomes, the next step is DNA cloning, allowing the creation of exact copies of specific DNA sequences. Cloning of phage lytic genes enables the production of these valuable proteins. Cloning, which requires multiple steps to create a replica of the target DNA, is used to produce the desired proteins or DNA in large quantities which can then be used to study the function of genes of interest, manipulate genes, and characterise specific genes (Brown, 2002). These cloned genes can be expressed in host organisms such as bacteria, plants, yeast, or mammalian cells, in order to produce them sufficiently for further research and application.

A notable illustration of the impact of cloning in biopharmaceuticals is insulin production (Loenen *et al.*, 2014). Gene mining of phage lytic enzymes has gained substantial attention due to their potential to combat antibiotic resistance, a pressing concern within the healthcare sector (Lai *et al.*, 2013). The cloning process typically involves DNA amplification through PCR, DNA cleavage using restriction enzymes, ligation of the cleaved gene into a vector, and transformation to obtain plasmids harbouring genes of interest. Subsequently, these genes can be expressed using preferred expression systems, marking a promising intersection of gene mining and cloning technologies to address critical challenges.

The choice of approach to pursue novel genes depends on various factors, including the nature of the gene library, the database, specific genes of interest, time constraints, and resource availability. Sequence-based methods to screen for phage-derived lytic proteins offer considerable advantages, such as discovering potential candidates without time-consuming functional screening assays and identifying novel genes that might not be easily identified using conventional screening approaches. This chapter's central objective is to explore the

sequence-based approach to mining of phage-derived lytic proteins with the potential to serve as next-generation antimicrobials from an online phage database (https://phagesdb.org/) focusing on University of KwaZulu-Natal (UKZN) sequences.

3.2 Methods and Materials

3.2.1 Phage Database Screening

Phage-derived lytic proteins were screened from a phage database using a sequence-based approach focusing on the University of KwaZulu-Natal (UKZN) phage database (https://phagesdb.org/institutions/KRIT/) (Russell and Hatfull, 2017). The database consists of an Actinobacteriophage Database comprising of 332 phages obtained from functional screening of environmental samples taken around Durban, Kwa-Zulu Natal. The coordinates for the soil sampling were 29.81802 S, 30.94418 E, 29.816519 S, 30.945196 E and 29.8587 S, 31.0218 E. The soil is not known to be enriched, and the soil samples were taken from a superficial level. From the 332 phages in the database, 65 (19.6%) have been sequenced and available on GenBank. Sequences were analysed using BioEdit (http://www.ncbi.nlm.nih.gov/) (Hall, 1999) and CLC Bio WorkBench (Scholz *et al.*, 2014) software programmes. Sequence translation of the sequences was conducted using BioEdit using Nucleic Acid Sort Six-Frame translation with a minimum ORF number of 100 and start codon of ATG. Molecular weight and theoretical pI (isoelectric point) were computed using the Compute pI/Mw tool from Expasy (Gasteiger et al., 2005). Homology searches were conducted using the basic local alignment search tool (BLAST) (http://www.ncbi.nlm.nih.gov/BLAST/) was used. Determining conserved domains within the sequences was achieved using the NCBI conserved domain function (Altschul *et al.*, 1990). The prediction of signal peptides was performed using the SignalP 4.1 server (http://www.cbs.dtu.dk/services/SignalP/).

3.2.2 Phylogenetic Analysis

Nonreductant sequences from NCBI blast database were used as a reference to determine the phylogenetic position of the identified phage lytic proteins. The selected sequences were accessed using sequence identity, closely related sequences based on homology. The phylogenetic tree was constructed using Molecular Evolutionary Genetic Analysis (MEGA v11). The alignment of the sequences was done through MEGA using MUSCLE (Hall, 2013). Following sequence alignment, phylogenetic tree was constructed using the maximum likelihood statistical method with 1000 bootstrap replicates (Kumar *et al.*, 1994; Hall, 2013) under the Jones-Taylor-Thornton model.

3.2.3 Gene Synthesis

After completing the above bioinformatic analyses on the sequences, a subset of four genes was chosen for further investigations. The selected genes were synthesized and cloned into the pUC57 vector (Figure A.1) using the services of GenScript (Piscataway, USA). The following restriction enzymes NruI, NdeI, and AgeI were inserted on the 5' and XhoI on the 3' of the genes. The restriction enzymes were specifically chosen because of their compatibility with the gene sequences, ensuring that they remained in-frame.

3.2.4 Development of Expression Constructs

3.2.4.1 Chemical Transformation

Before transformation, chemically competent cells (*E. coli* DH5) were given 5 to 10 min to thaw on ice. Two µl of the plasmid DNA (pUC_LysA, pUC_LysB, pUC_Lysin and pUC_Holin) were added to 50 µl of cells in 1.5 ml Eppendorf tubes on ice and left for 10 mins. Subsequently, a 45-second incubation period at 42 ° C was followed by a 2 min cold shock period on ice. After cold shock, 500 µl of SOC media was added to each tube to recover the cells which was then followed by a 1 h incubation period at 37 °C, 200 rpm shaking. After incubation, 100 µl of the grown cells were plated onto LB agar plates with 50 mg/ml of kanamycin (kan50) and incubated overnight at 37 °C. The following day, random colonies were inoculated in 5 ml of LB supplemented with kan50 and incubated at 37 °C, 180 rpm for the extraction of the plasmid.

3.2.4.2 Isolation and Quantification of Plasmid DNA from the different vectors (pUC57, pEAQ-HT and pET-30b(+))

Plasmid DNA was extracted using the GeneJet MiniPrep from Thermo Fisher Scientific (Waltham, Massachusetts) with minor modifications. The modifications involved two washes with 600 µl of washing solution. Using 30 µl of nuclease-free water, bound DNA was eluted. The Hidex Sense microplate reader (Hidex, Deutschland) was used to measure the absorbance at 280 nm, 260 nm, and 230 nm to determine the purity and concentration of isolated plasmid DNA. One absorbance unit at 260 nm equals 50 µg of dsDNA/ml per. Isolated DNA has a purity range of 1.5 to 2.0 at the A260 nm to A280 nm ratio and 1.8 to 2.3 at the A260 nm to A230 nm ratio, according to Marmur (1961).

3.2.4.3 DNA Separation by Agarose Electrophoresis

Based on the molecular weight of the DNA fragments, and plasmids were separated on 1% and 1.5% (w/v) agarose gels. The agarose gel was prepared using 1 x TAE buffer containing 5 µl of GelRed (Sambrook and Russell, 2001). The samples were combined with 6x DNA loading dye prior to electrophoresis. NEB's 1 kb Gene Ruler molecular weight marker (Cat. NEB #B7025) was utilized unless stated otherwise. The same buffer utilised to prepare the

agarose gel was used in the electrophoresis buffer chamber of the apparatus (PEQLAB Biotechnologie GmbH, Germany). The visualisation was carried out under 300 nm ultraviolet light and photographed using a ChemiDoc MP system (Bio-Rad, Universal Hood III, California, USA).

3.2.4.4 Polymerase Chain Reaction (Gradient PCR)

Primer sets were designed to amplify the gene from pUC57 and introduce the *Age*I restriction site on the 5' of the Lysin and Holin gene while introducing *Nru*I for LysA and the LysB, and ATG start codon. On the 3' end of all genes, *Xho*I was introduced together with the TGA stop codon after the restriction site to exclude the C-terminal His_6 tag. Amplification was achieved using a BioRad Thermo Cycler (C1000 Touch^{TM} thermal cycler) and gradient PCR was used. A Q5 PCR kit from NEB Biolabs (New England) was used for the reactions. Amplicons with a molecular mass corresponding to our target genes were purified using a gel extraction kit (Thermo Fischer). The NanoDrop 2000 Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) was used to measure at an absorbance of 280 nm, 260 nm, and 230 nm to determine the concentration and purity of DNA. Table 3.1 shows the primer sets used to amplify the genes.

Table 3.1: Primer sets designed to amplify genes out of pUC57 introduce restriction sites for cloning into the pEAQ-HT vector.

Name	Sequence	Primer
LysA	5’ CCACCACCACTACCCGGGTGGCA	Forward1 and 2
	5’ tcgcgaATGCACCACCACCACCACCAC	
	3’CTCGAGAATCGGACGCAGCGCCGCCG	Reverse
LysB	5’ CACCACCACGCGACC	Forward1 and 2
	5’ tcgcgaATGCACCACCACCACCACCAC	
	3' CTCGAGAATACGACGGAAACCCGC	Reverse
Lysin	5’ CCACCACCACACCGCGATC	Forward1 and 2
	5’ accggtATGCACCACCACCACCACCAC	
	3’ CTCGAGAATGTTGCCTTTCAG	Reverse
Holin	5’ ATGCACCACCACCACCACCACGATAAAATAAGGC	Forward1 and 2
	5’ CATATGACCGGTATGCACCAC	
	3’ TGACTCGAGGACACGTGGACGAGCGGCCG	Reverse

3.2.4.5 Digestion with Restriction Enzymes

Amplicons after gel purification were prepared for ligation by cleaving the ends with the respective restriction enzymes. Based on Sambrook and Russell’s (2001) methods, a total

reaction volume of 20 µl was prepared to contain x µl of dH2O, 1-5 µl (~50-100 ng/µl) of DNA, 2 µl of 10x Fast Digest Buffer and 1µl (1U/µl) of restriction enzyme (per enzyme in a double digest reaction). The reaction was stopped by heat inactivation at 75 ℃ for 5 min. To that, a sample loading dye was added. The samples were then loaded into a 1% agarose gel for gel electrophoresis. The GeneJet Gel Purification kit was used to purify the DNA from the gel.

To generate pEAQ-HT_LysA, pEAQ-HT_LysB, pEAQ-HT_Lysin, and pEAQ-HT_holin, the PCR products corresponding to the molecular mass of the selected genes were purified and ligated into a prepared pEAQ-HT vector. The ligation reactions were carried out using T4 DNA ligase based on Sambrook and Russell's (2001) description. In a 20 µl reaction volume, the insert to vector was 3:1. Ligations were performed at room temperature unless otherwise indicated. To deactivate the reactions, the reactions were subjected to heat shock at 70 ° C for 5 min.

These ligations were electroporated in *E. coli* DH10B (1.44 kV, 200 Ω and 25 µF). Electroporated cells were recovered with LB media and incubated at 37 °C, 200 rpm for an hour of recovery before plating on LB agar plates supplemented with streptomycin 50 mg/l, kanamycin 50 mg/l, and rifamycin 20 mg/l. Single colonies were inoculated in LB media supplemented with respective antibiotics and incubated overnight (37 °C at 200 rpm). The inoculum was then extracted for plasmid and stored at –20 °C until use.

3.3 Results

An online phage database (focussing on UKZN sequences) was screened using bioinformatic tools for lytic proteins leveraging on the established efficacy of phages as antimicrobials. The phage database with the sequences from UKZN had, in total had 19,6% of unique sequences and an abundance of uncharacterised sequences. The lytic sequences contained overhang sequences with 10 bases (CGGTCGGTTA and CGCTTGTCA and for protein expression, a start codon is required at the start of the protein sequence therefore, in the six-frame translation, a command was made to translate the sequences and have each ORF commence with a starting codon (ATG) to facilitate starting protein for protein expression. NCBI Blast tool was used for homology search of all identified sequences, from which several sequences of potential relevance were detected. Finally, four were selected for further studies. Table 3.2 summarises the amino acid information and sizes of the four phage lytic genes chosen then translated to protein for further characterisation. The table also shows the theoretical pI of the translated proteins together with the minimum and maximum hydropathy scores from ProtScale. Signal peptide analysis of the four ORFs revealed that their N-terminal has no signal peptide (Table 3.3), with high reliability ranging from 0.96-1.0. Checking whether the presence/ absence of genes contains a signal peptide is crucial for determining where the

expression will take place and the functionality of the protein expressed. If the gene contains the tag, it must be removed for the protein to be expressed intracellularly and not compromise its functionality post-expression.

Figure 3.1 shows a group of uncharacterized genes encoding for phage lytic proteins with properties suggesting promise as novel antimicrobials, there were identified with the aid of NCBI BLASTp. These proteins encompass domains such as lysozymes, amidases, and proteins containing PG-binding domains (CBDs). The diversity of the enzymes potentially allows them to target various components of a microbial cell wall, while the CBDs specifically bind to peptidoglycan, a key structural element. It can also be observed that the CBD positions on LysA and LysB are on different terminal ends of the gene.

Table 3.2: Summary of amino acid sequences of the selected putative phage-derive proteins.

Gene	Nucleotides (bp)	pI	Scale: Hydropathicity.	Molecular Weight (kDa)	Translated Amino Acid Sequence
Lysin A (LysA)Designated SM03)	1341	6.80	MIN: -2.311 MAX: 2.344	48.62	MYPGGTTIGTPARPAAERKTRFWWDIVRSIDNGFAVILNWVVPPARKPIKAVKGSTNPSYGS GTTYHYVTAVGWSDEGNGGRPAVLIADSGFSPNVYWVDLDTAFALIHTDLWKGYAFADLPLI APPPPGVEVPPGIPVKIGDPPIAVTPPAPVPPTQPPTVVKITDPFTGAIWSPNRYSPRGLGTP GWIAVHTQEGGRTARDLALFLANPANEVSYHSVNDDVEVLKVVAEGDASWSASNANKYAF HHCFAGSYAGWSRDKWLSPDASDGKNEDVQLTKGAHVVAWWCDKYGIPAEWIGGRAQPP WGARGILGHVDLGQWGGGHFDPGGNFPVNEFIRRVVKFLTGEDQPPLVPLPPVVVPGTNP DAYSDWMLVRGDPRNDVDRVMRVQRRLKNAYKGYAGHLAVDGDFGPATQAAVREFQRRS NLVADGIVGPMTAAALRP
Lysin B (LysB)Designated SM29	1374	7.77	MIN: -2.767 MAX: 2.344	50.28	MATGSHPRHGVQSRLGWPVVGAPLRPLPPHRQETMMPELKLGYQGPLYAPWFDWFTRKY SQTAPLLGRKDGYYGSDEKRAVETLQRNLGIVVDGVFGDRTAAAAGYTWPGASAPPVVTP RRPIWFYSCPGSGADWWLGPSYDVGQMVAGTGWNEPGRRSLNINHQPVGFPKGGYLGL MGGDPTFSYIEVITAQKIEFARLLRTNPDVQLAMEARQRDRNARVDVELWPSGYSQSADGM CDAVAELFGDGGEFELIRDRINGLILFGNPATPVTGIARKTYPAWLTALMRNINMSDDFYAVA KDRIRPAFYAEIIKAEMELPFFVHVLRIAVPIVLSWASTLLPFLTPLLGGLGPGVQLGLGMISGL QGLGSNPALTQLLGMAGSGKDVQTTERVEDILSPTGILSNIPDLIAMLGALPGLQSHGLYHAT APARPEFGNRVGTQFAYDIIAGFRR
Endolysin (Lysin)Designated SM07	1383	5.82	MIN: -3.111 MAX: 2.000	51.62	MTAIITRQRAQEVHDRARARKGLPYAYGGAFTNDPRRSTDCSGLVLQTAAWYMGRTDWVG NRYGSTESFRLDHKIVYDLGFKRLPPGGIAALGFTPVMLVGLQHGGGGMYSHTACTLMTMD IPGGPVKVSKRGVDWESHGNRNGVGVDLYDGARAWNDPLFHDFWYLDAKLEDAPAPEDD AVEILSQATGLSTERALEILPAVRDGLIASECVNANRIAMWLAQVGHESASFKYTEEIAKNGR

					YAPYIGRTWIQITWDYNYRAFSQWCFDRGLVSTPDYFVRNYTDLALLKWAGLGAAWYWTV ARTDINALSDRRDLETVTLRINGGHNGIADRRERYDRASALGDRLLALISDAQPIDPFEEEMM REVESFSIYATPGEPKIPLFVMLQSLDAHGPHEPYVEEQARHGDRDAISRVIRTAAGKGKYG TASGPVNQASRVLAELEDSGVLSNYLKGN
Holin Designated SM31	396	6.06	MIN: -1.989 MAX: 2.522	13.23	MDKIRQYYYLAAAAIAGLIPILNILHVLSDDQILSVNQIIASIGSLIGAAGVGTAGVILGKQRKEGV LEKLDPADKVVTGIQEAVTAAAHANDQLDRVKEVVNTLNPGTVIGNLVEQAIAAAARPRV

Table 3.3: Signal peptide prediction report.

Protein ID	**Predicted Signal**	**Reliability**	**Cleavage Site**
Holin (SM31)	Transmembrane*	0.96	-
LysA (SM03)	None#	0.99	-
LysB (SM29)	None#	1.0	-
Lysin (SM07)	None#	1.0	-

*a membrane-spanning segment is predicted at the N-terminus of the protein

Neither a signal sequence nor a membrane-spanning part is expected at the N-terminus of the protein

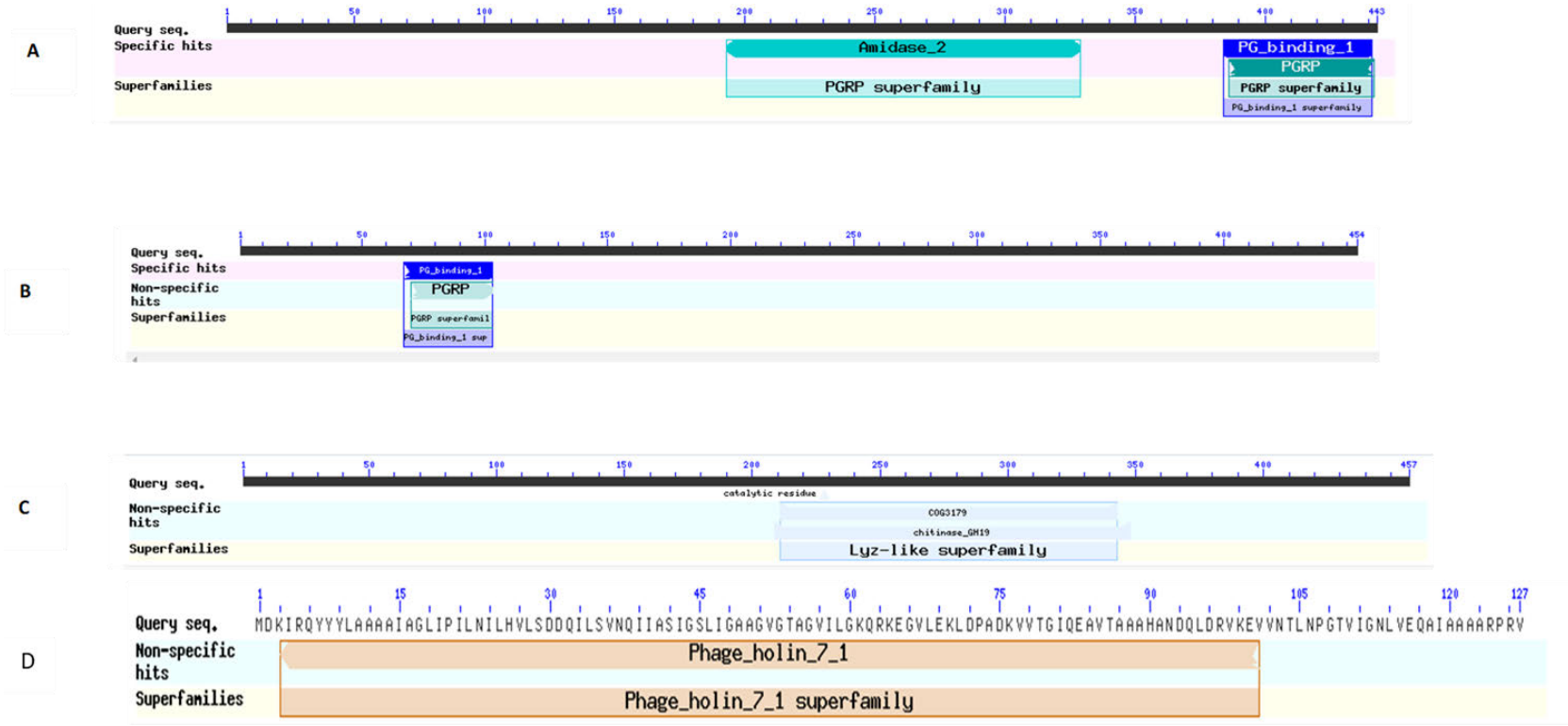


Figure 3.1: Graphic summary of putative conserved domains detected in (A) LysA (SM03), (B) LysB (SM29), (C) Lysin (SM07) and (D) Holin (SM31). The PG domain found in (A) and (B) is crucial for the degradation of bacterial cells. The central domain scan for (C) shows a relation to the lyz-like superfamily (lysozyme), which plays a role in PG cleavage.

A phylogenetic tree was constructed based on the amino acid sequences of endolysins (PG hydrolase) and cytoplasmic membrane degrading enzymes (holins) genes to analyse the conservative core sequences of the genes. Using MUSCLE alignment method, phylogenetic analysis was done using maximum likelihood, yielding the phylogenetic trees shown from Figure 3.2 to Figure 3.5. From analysis of the trees, it can be deduced that most of the sequences are novel and mainly related to those isolated from the *Mycobacterium* genus with most of the nodes across above 75, affirm the relation between the nodes. From the phylogenetic tree analysis, a correlation can be made between the domain searches.

Figure 3.2: Phylogenetic tree of LysA (SM03) using maximum likelihood analysis.
LysA marked with a red circle.

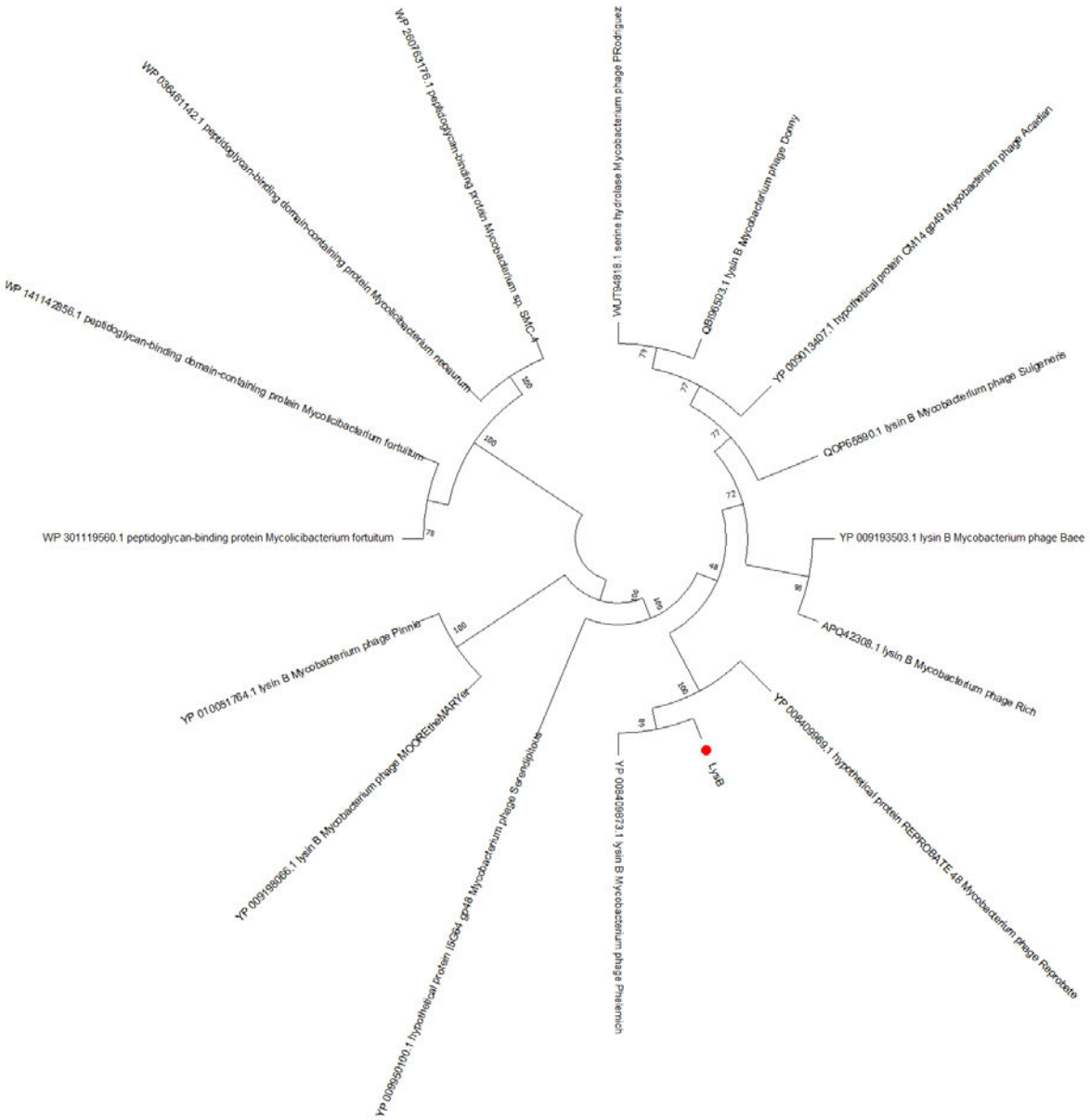


Figure 3.3: Phylogenetic tree of LysB (SM29) using maximum likelihood analysis.
LysB marked with a red circle.

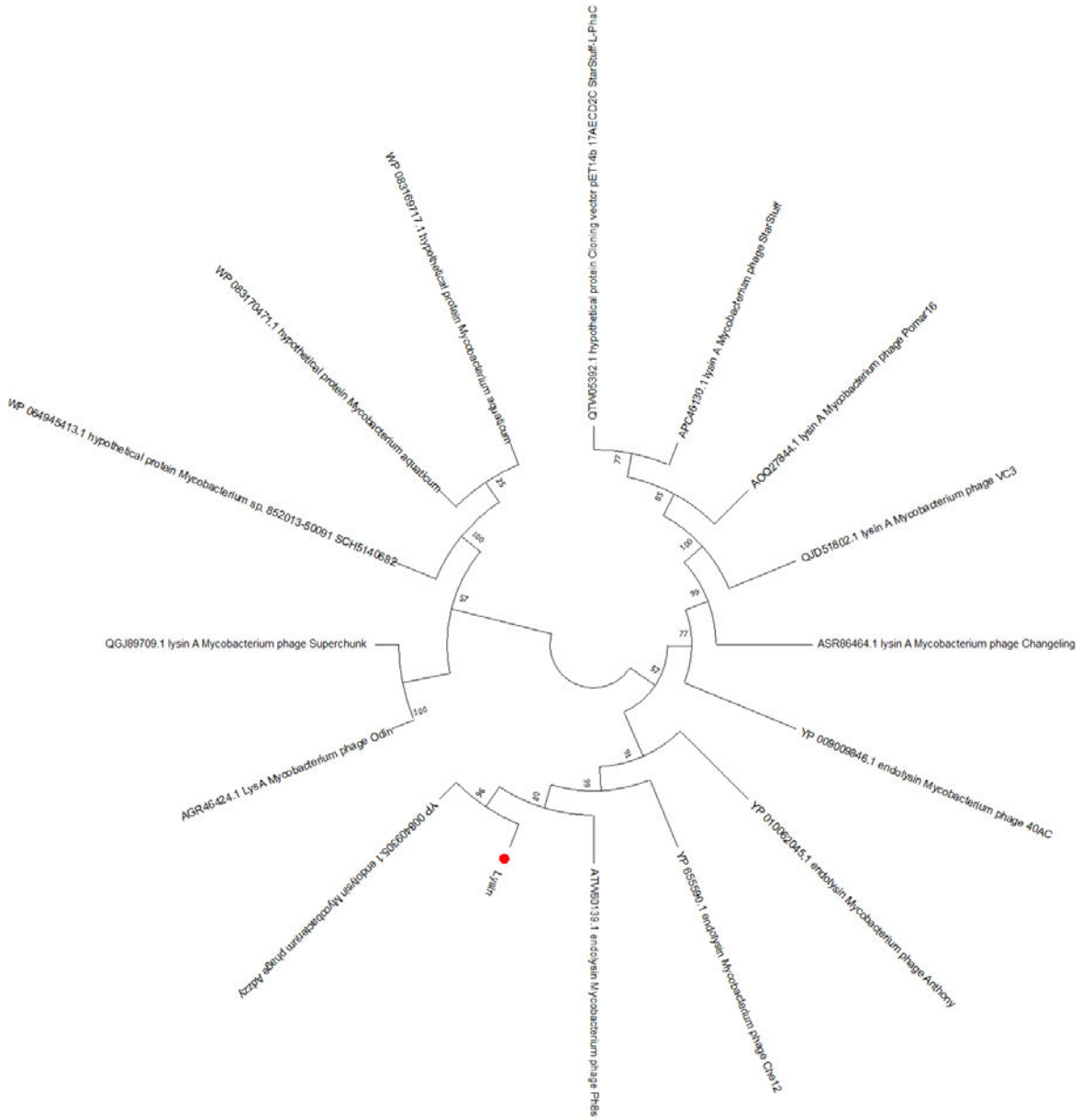


Figure 3.4: Phylogenetic tree of Lysin (SM07) using maximum likelihood analysis. Lysin marked with a red circle.

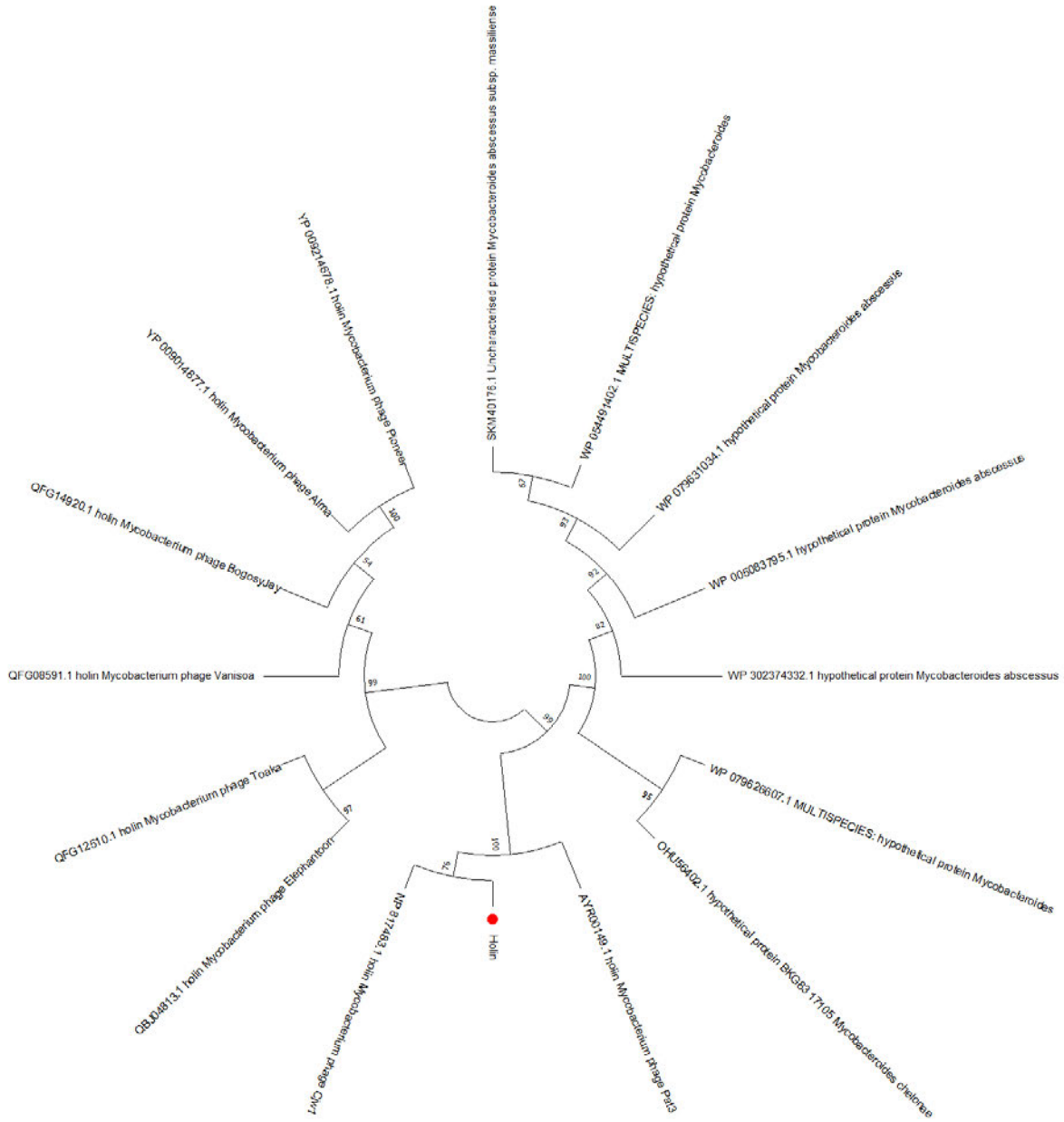


Figure 3.5: Phylogenetic tree of Holin (SM31) using maximum likelihood analysis.
Holin marked with a red circle.

For transient expression in *N. benthamiana*, GenScript synthesised the genes and cloned them into the pUC57 vector for plasmid DNA propagation. Primer sets were designed to amplify the genes in pUC57 to subclone into an expression vector. Figure 3.6 shows the plasmid from the PCR amplicon that were prepared for ligation into the expression vector.

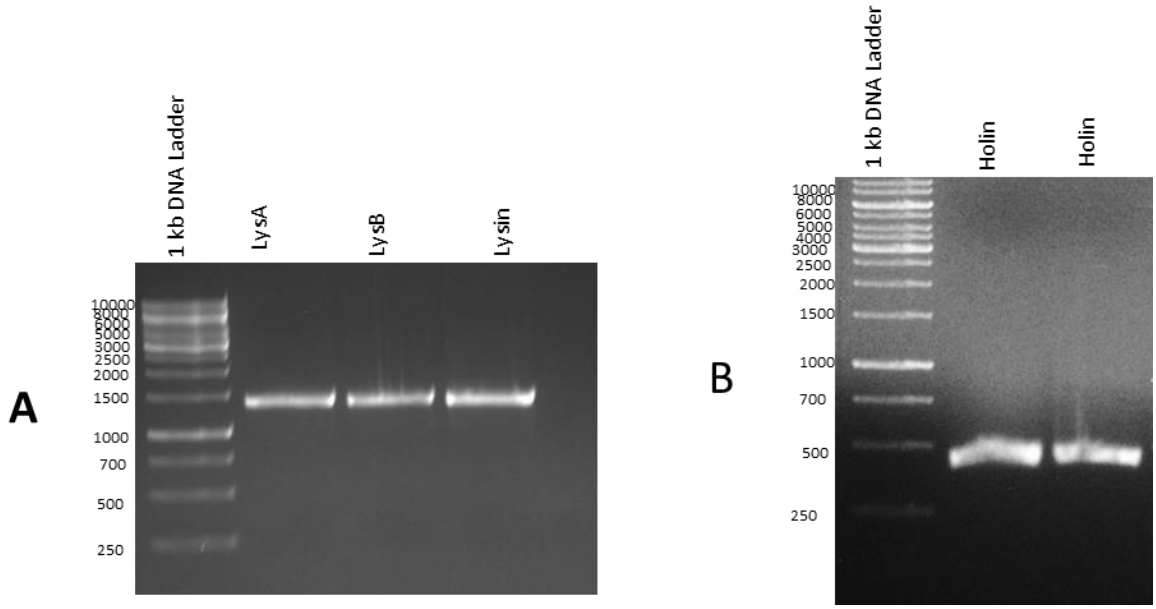


Figure 3.6: PCR amplicon from pUC57 prepared for ligation into pEAQ-HT vector.
(A) LysA, LysB and Lysin genes prepared for ligation into an expression vector and (B) Holin gene prepared for ligation into pEAQ-HT post PCR amplification from the gene's synthesised by Genscript.

3.4 Discussion

Lytic proteins offer a promising solution as next generation antimicrobials to address the growing problem of antibiotic resistance and provide a viable alternative to treating bacterial infections (Gondil *et al.*, 2020; Aslam *et al.*, 2021a). A phage database was screened for genes encoding for lytic proteins, leveraging on the established efficacy of phages as antimicrobials. Studying phage databases offers several advantages such as avoidance of the functional screening approach, fast access to available and uncharacterised sequences, and by exploring the online databases the characterisation of uncharacterised sequences can reveal new information by using bioinformatics tools This approach is a powerful tool for efficiently screening and identifying target molecules. From the database, uncharacterised sequences with putative lytic functions were examined. These enzymes are known to be instrumental in the degradation of peptidoglycan (PG), a critical component of bacterial cell walls (Schmitz *et al.*, 2010; Luong *et al.*, 2020).

The homology search prioritized uncharacterized proteins containing critical functional domains, aiming to identify novel lytic proteins with potential for future applications. Endolysins often feature an enzymatic domain with no cell wall binding domain (CBD) or one or two CBDs (Nelson *et al.*, 2012b; Heselpoth *et al.*, 2018a) located at either end of the protein. The selected ORFs displayed variations in domains, all of which are necessary for lytic function. These domains include a catalytic domain and a CBD. The identified catalytic domains play a crucial role in the cleavage of bonds such as β-(1-4)-glycosidic bonds and the amide link between *N*-

acetylmuramic acid and L-alanine residues found in PG (Carlos Jorge Sousa de São-José, 2016; Shrivastava *et al.*, 2019).

Alignment analysis of the LysA sequence revealed a catalytic domain located at the N-terminal and a cell wall binding (CWB) domain at the C-terminus. This domain arrangement demonstrates typical endolysin. In the case of LysB, the alignment results strongly support its classification as a hydrolase, a classification commonly associated with endolysins (Fenton *et al.*, 2010b). Also observed via the alignment and analysis of LysB was the identification of a fragment on the N-terminal region involved in cell wall binding. This analysis emphasises its potential role as a next-generation antimicrobial. When aligning Lysin sequence with other related sequences, it was evident that Lysin exhibits a CBD with a specificity for Gram-negative bacteria.

Furthermore, the analysis showed homology to amidase and chitinase in the catalytic domains. Both are recognized enzymatic activities of endolysins. Lastly, alignment of holin sequences with other sequences indicated homology to phage holins associated with Gram-negative and Gram-positive bacteria. These alignments and the combined phylogenetic tree data suggest that the genes encoding for phage-derived proteins share a common function, reinforcing the view that the selected genes pose the functionality of phage lytic proteins. They have the potential to target and disrupt bacterial cell walls. Based on the phylogenetic tree analysis, the genes are all from the *Mycobacterium* genus and there is a distinction with the Holin as observation can be made that they have homology to *Mycobacterium abscessus.*there while the other three show homology to *Mycobacterium phage*. LysB indicates a strong PG-binding domain and Lysin shows homology to chitinase and indicating some specificity toward Gram-negative bacteria. Therefore, if there appears to be shared commonality amongst the three enzymes as under Lysin sequences other lysins A can be observed and under LysA, endolysins can be seen. There is a possibility that all three might also demonstrate specificity towards Gram-negative bacteria as well. However, confirming this will require further characterisation work to verify their full functionality.

The significance of crucial domains extends beyond functionality but also to their ability to be recombinantly expressed. Assessing the presence of signal peptides within ORFs is important due to the role they play in mediating protein targeting and translocation within both prokaryotic and eukaryotic cells (K Terpe, 2003a; Vollmer *et al.*, 2008b). A signal peptide analysis of the genes revealed the absence of a leader peptide, indicative of protein retention within the cytoplasm post expression (Nakai and Imai, 2018). However, some proteins have been documented to exhibit extracellular functionality despite the lack of signal peptides; the

absence of a signal peptide does not necessarily prevent the possibility of protein translocation outside of the cellular cytoplasm, nor does it mean that it will remain biologically active after expression. No cleavage site was identified within our selected genes and this is doesn't have any implications on our target protein. However, had a signal peptide identified and no cleavage site identified, challenges with expressing the proteins might be encountered. Such problems would include misplacement of the protein, functionality of the protein, and in some instances, protein degradation.

The pI values indicate the pH at which each protein will have a net charge with LysB with the highest pI and Lysin with the lowest at 5.82. LysA and Holin on the other hand have pI values closer to physiological pH. This information on pI is valuable for protein purification, especially for techniques such as ion exchange chromatography. The hydrophobicity of the proteins was analysed using Kyte-Doolittle hydropathy scale. A variation of values can be seen with the Holin having a maximum value of 2.522 indicating a potentially hydrophobic region and Lysin with a minimum value of -3.111 suggesting a region with a high concentration of hydrophilic amino acids.

To elucidate the functionality of the selected genes, the genes were sub-cloned for transient expression in plants using pEAQ-HT vector. PCR amplification gel in Figure 3.6 revealed no contamination suggesting and had bands corresponding with the target genes suggesting a good PCR. The next time was to the sub-clone then into an expression vector and the products will be ready for downstream processes. Because these proteins actively function against specific bacteria, transient expression in plants enables the production of these enzymes without the need for special cell cultivation and/ or expression conditions to avoid complications of bacterial expression.

3.5 Conclusion

Gene mining of the phage database yielded significant results and four putative endolysin phage lytic proteins were discovered. The discovery of these genes encoding for putative phage lytic proteins not only illustrates the advantages of mining phage databases but also underscores their potential for finding next generation antimicrobials for therapeutic applications. Observation in the gene arrangement of the four encoding for putative phage lytic proteins provided valuable insights on them and their potential functional. The gene arrangement displayed typical arrangements of endolysin characters. They possessed CBDs, catalytic domains and had homology to other classes of enzymes with typical phage lytic protein functions. Moreover, phylogenetic analysis indicates a shared ancestry among these genes with three genes showing close homology to each other, reinforcing their potential

functions as endolysins capable of targeting and disrupting the cell walls of specific bacteria. The observed theoretical pI provides insight to the net charge properties of the four proteins and can be used further to predict protein behaviour and guide other functions and interaction. The hydropathy scale will help in comprehension of the potential function of these proteins and other factors that can be affected might be the folding as more hydrophobic regions might fold inwards and those that are hydrophilic might be on the surface of the protein. Mining the phage database proved to be a valuable tool for the discovery of target molecules. The discovery of these genes encoding for putative phage-derived proteins provides some structural and potentially functional characteristics of the putative phage lytic proteins and their functionality as antimicrobials; however, furthermore work needs to be conducted to fully elucidate their functionality to pave the way for novel antimicrobial agents.

3.6 References

Adrio, J.L. and Demain, A.L. (2014) Microbial enzymes: tools for biotechnological processes. Biomolecules 4: 117–39.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. J Mol Biol 215: 403–10.

Aslam, B., Arshad, M.I., Aslam, M.A., Muzammil, S., Siddique, A.B., Yasmeen, N., *et al*. (2021) Bacteriophage Proteome: Insights and Potentials of an Alternate to Antibiotics. Infect Dis Ther.

Brown, T.A. (2002) Studying DNA. In Genomes. Oxford: Wiley-Liss.

Carlos Jorge Sousa de São-José, D. (2016) Endolysins as Antibacterial Agents: from Engineering Approaches to the Uncovering of Holin as a Key Factor Influencing Lytic Activity.

Clokie, M.R., Millard, A.D., Letarov, A. V, and Heaphy, S. (2011) Phages in nature. Bacteriophage 1: 31–45.

Fenton, M., Mcauliffe, O., O'mahony, J., and Coffey, A. (2010) Recombinant bacteriophage lysins as antibacterials.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). The Proteomics Protocols Handbook Protein Identification and Analysis Tools on the ExPASy Server. http://www.expasy.org/tools/.

Gondil, V.S., Harjai, K., and Chhibber, S. (2020) Endolysins as emerging alternative therapeutic agents to counter drug-resistant infections. Int J Antimicrob Agents 55:.

Górski, A., Miedzybrodzki, R., and Borysowski, J. (2019) Phage therapy: A practical approach, Górski, A., Międzybrodzki, R., and Borysowski, J. (eds) Springer.

Hall, B.G. (2013) Building phylogenetic trees from molecular data with MEGA. Mol Biol Evol 30: 1229–1235.

Hall, T.A. (1999) BioEdit: A user-friendly biological sequence alignment program for Windows 95/98/NT. Nucleic Acids Symp Ser 41: 95–98.

Heselpoth, R.D., Swift, S.M., Linden, S.B., Mitchell, M.S., and Nelson, D.C. (2018) Enzybiotics: Endolysins and Bacteriocins.

Kumar, S., Tamura, K., and Nei, M. (1994) MEGA: Molecular Evolutionary Genetics Analysis software for microcomputers.

Lai, M.J., Soo, P.C., Lin, N.T., Hu, A., Chen, Y.J., Chen, L.K., and Chang, K.C. (2013) Identification and characterisation of the putative phage-related endolysins through full genome sequence analysis in Acinetobacter baumannii ATCC 17978. Int J Antimicrob Agents 42: 141–148.

Loenen, W.A.M., Dryden, D.T.F., Raleigh, E.A., Wilson, G.G., and Murrayy, N.E. (2014) Highlights of the DNA cutters: A short history of the restriction enzymes. Nucleic Acids Res 42: 3–19.

Luong, T., Salabarria, A.C., and Roach, D.R. (2020) Phage Therapy in the Resistance Era: Where Do We Stand and Where Are We Going? Clin Ther 42: 1659–1680.

Morrisette, T., Kebriaei, R., Lev, K.L., Morales, S., and Rybak, M.J. (2020) Bacteriophage Therapeutics: A Primer for Clinicians on Phage-Antibiotic Combinations. Pharmacotherapy 40: 153–168.

Mtimka, S., Pillay, P., Rashamuse, K., Gildenhuys, S., and Tsekoa, T.L. (2020) Functional screening of a soil metagenome for DNA endonucleases by acquired resistance to bacteriophage infection. Mol Biol Rep 47: 353–361.

Nakai, K. and Imai, K. (2018) Prediction of protein localization. Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics 1–3: 53–59.

Nelson, D.C., Schmelcher, M., Rodriguez-Rubio, L., Klumpp, J., Pritchard, D.G., Dong, S., and Donovan, D.M. (2012) Endolysins as Antimicrobials. Adv Virus Res 83: 299–365.

Nwankwo, D.C. and Abalaka, M.E. (2014) Restriction Enzymes and Their Uses in Specific Sequencing To Produce Predictable Fragment of Dna Making Genetic Engineering Simply. Journal of Pharmaceutical Research & Opinion 1: 148–152.

Reina, J. and Reina, N. (2018) Phage therapy, an alternative to antibiotic therapy? Revista Espanola de Quimioterapia 31: 101–104.

Russell, D.A. and Hatfull, G.F. (2017) PhagesDB: The actinobacteriophage database. Bioinformatics 33: 784–786.

Sambrook, Joseph. and Russell, D.W. (David W. (2001) Molecular cloning : a laboratory manual, Cold Spring Harbor Laboratory Press.

Schmitz, J.E., Schuch, R., and Fischetti, V.A. (2010) Identifying Active Phage Lysins through Functional Viral Metagenomics †. Appl Environ Microbiol 76: 7181–7187.

Scholz, M., Lo, C.-C., Chain, P.S.G., Scholz, M.B., Miller, J.R., Earl, D., *et al*. (2014) Improved Assemblies Using a Source-Agnostic Pipeline for MetaGenomic Assembly by Merging (MeGAMerge) of Contigs.

Shrivastava, A., Shrivastava, N., and Singh, P.K. (2019) Enzymes in Pharmaceutical Industry. Enzymes in Food Biotechnology 591–602.

Terpe, K. (2003) Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. Appl Microbiol Biotechnol 60: 523–533.

Vollmer, W., Joris, B., Charlier, P., and Foster, S. (2008) Bacterial peptidoglycan (murein) hydrolases. FEMS Microbiol Rev 32: 259–286.

4 CHAPTER FOUR: RECOMBINANT PROTEIN PRODUCTION OF PHAGE DERIVED PROTEIN FOR ANTIMICROBIAL ACTIVITY

Abstract

Phage-derived proteins have gained significant attention because of their applications in various fields, including their use as reagents and therapeutic molecules. Their recombinant production is a crucial step towards their application in combatting the looming effects of antibiotic resistance. Although the transient expression in the *Nicotiana benthamiana* system might be favoured over the bacterial expression system for phage-derived proteins due to their nature, the bacterial expression system offers simplicity and easy gene manipulation and has been shown to perform well for phage lytic proteins. In this chapter, the aim was to describe the recombinantly produced phage-derived proteins. The phage-derived protein genes were transiently expressed in *Nicotiana benthamiana* plants using the non-replicating Cowpea mosaic virus (CPMV)-based vector pEAQ-HT, and in *Escherichia coli* using pET-30b(+) expression vectors. Proteins were expressed as 6 × histidine fusions to facilitate purification and analysis. The His_6 tag was at the N-terminus for plant translated genes, and for bacterial expression, the His_6 tag was at the N- or C-terminus of the translated gene The purification was conducted using immobilised metal affinity chromatography (IMAC) under normal and denaturing conditions, and a second purification step was conducted using ion exchange chromatography. Transient expression in *N. Benth* with LBA4401 host cells produced truncated proteins for LysA (SM03), LysB (SM29), and Lysin (SM07) according to western blot analysis, while no expression was observed for Holin (SM31) protein. Heterologous expression was achieved in BL-21(AI) for SM07, SM03 and SM29 with a final concentration of IPTG at 0.1 mM, 0.2% arabinose at 37 °C over a period of 4 h post-induction. The N-terminal His_6 tagged-SM07 was successfully purified under normal conditions having been induced with both IPTG (0.1 mM) and arabinose (0.2%) and confirmed by peptide mass fingerprinting with 3 peptides (DLETVTLR, ALEILPAVR and YAPYIGR) to be an endolysin. Subsequently, SM07 was successfully subjected to a second purification step using anion exchange (CaptoQ) and quantified to 0.0006 g/l. The bacterial expression in BL-21 system outperformed transient expression in *N. benth* with respect to the production of recombinant phage lytic proteins. Furthermore, we successfully streamlined SM07 by shifting the histidine tag from the C-terminus to the N-terminus and implementing anion exchange as a secondary step.

4.1 Introduction

Finding effective therapeutic agents has become challenging due to several factors such as slow discovery of new drugs, improvement of efficiency on currently existing drugs, rapid mutations, misuse of currently available treatment regimens and defaulting. Antimicrobial proteins, such as phage-derived enzymes, have gained interest as next-generation antimicrobials to tackle antibiotic resistance. Bacteria gain resistance either via horizontal gene transfer that involves antibiotic degradation enzymes, or pathways that allow them to bypass the effects of the antibiotics, or by spontaneous mutation that changes the antibiotic target or up-regulates the efflux pumps (Munita and Arias, 2016; Aslam *et al.*, 2018b; C Reygaert, 2018; Mtimka *et al.*, 2024).

Recombinant protein production has been used to produce essential proteins in molecular biology research, drug discovery, and industrial applications (K Terpe, 2003b; Terpe, 2006; Adrio and Demain, 2014). With this approach, complex proteins, peptides, and antibodies with high specificity and efficacy can be produced, improving large-scale production efficiency through cost-effective and scalable processes with minimal variation. Several systems are used for the recombinant production of proteins, including but not limited to the transient plant expression system, bacterial expression system, mammalian protein expression system, and yeast expression system (Starkevič *et al.*, 2015; Mardanova *et al.*, 2017). Bacterial systems are frequently used to produce proteins that do not require intricate modifications (Rosano and Ceccarelli, 2014). In contrast, plant systems can offer desirable glycosylation patterns for therapeutic proteins (Singh *et al.*, 2021; Coates *et al.*, 2022). In cases where complex post-translational modifications or specific protein folding mechanisms are required, animal and algae expression systems may be favoured (Mauro, 2018a), although insect cell and yeast can also be considered. Plant and bacterial systems have rapid growth rates and efficient protein expression capabilities. They can produce high protein yields rapidly, have lower upstream costs and greater ease of scalability, making them suitable for time sensitive applications (Rosano and Ceccarelli, 2014; Mardanova *et al.*, 2017; Xu *et al.*, 2023a).

The use of these systems has facilitated the production of specific proteins which can assist in addressing intricate biological hurdles. One notable example is insulin production through recombinant methods, making it accessible to people with diabetes (Tripathi, 2016a). The global rise in antibiotic resistance can be combatted with biotechnological processes that produce next-generation antimicrobials. However, approaches to uncovering the next generation go beyond the production of recombinant proteins. There are numerous innovative strategies that can be explored such as nanotechnology, peptide-based therapy, CRISPR-based approaches, and combination therapy. Each approach offers promising avenues for

addressing the challenges of antimicrobial resistance and developing more effective treatments against various pathogens.

Over the past two decades, transient expression methods, employing systems such as *Agrobacterium* for the delivery of genetic material, have been pivotal in producing pharmaceutical proteins such as recombinant monoclonal antibodies and vaccines (Tsekoa *et al.*, 2020). Plant protein expression systems predominantly use the pEAQ vector via the CaMV 35S promoter and the Magnicon system (Komarova *et al.*, 2010b; Merlin *et al.*, 2014) *Agrobacterium* can be used to transform plants by infecting the plants through wounds and openings(De Saeger *et al.*, 2021). There are several *Agrobacterium* strains (Ach5 and C58) used within plant biotechnology but the most frequently used are derivatives of C58. AGL1 is a C58 derivative and LBA4404 has chromosome background of Ach5 but clusters with C58 as well (De Saeger *et al.*, 2021). Both have selection markers of rifampicin and AGL has an additional selection marker of carbenicillin (Krenek *et al.*, 2015; De Saeger *et al.*, 2021). These strains both contain the tumour-inducing (Ti) plasmids and influences transformation efficiency of the strains, however, strains from C58 harbouring octopine-type plasmid and pCH32 helper plasmid are degraded the best for transformation efficiency (Krenek *et al.*, 2015; Hwang *et al.*, 2017). Although phage-derived proteins hold promise as potent antimicrobial agents, their production remains challenging and optimizing expression strategies is crucial to achieve high yields of pure proteins. Recent advances have demonstrated successful replication of phage-derived proteins using plant expression systems. However, a study by Kovalskaya et al., (2019) revealed that the bacterial expression of phage-derived proteins performed better than those expressed in plants, specifically against *Staphylococcus aureus*. Regardless of the chosen expression system, challenges persist, including low expression levels, failure to express, and potential toxicity toward the host organism, all impacting protein functionality and purification for various applications. Overcoming these expression hurdles remains critical to obtain functional and purified proteins for pharmaceutical use.

Once proteins have been successfully expressed, purification is the next critical step in the production process. Recombinant protein production includes purifying the protein, which is crucial for obtaining a clean protein sample (K. Terpe, 2003; Komarova *et al.*, 2010a). The most employed and adaptable approach to purifying recombinant proteins in research is His tagging, which involves immobilized metal affinity chromatography using columns containing nickel or cobalt, providing enhanced purification capabilities (K. Terpe, 2003; Rosano and Ceccarelli, 2014). Unlike stable protein expression, which involves integrating the recombinant gene into the host genome, transient expression entails temporary expression of the gene in host cells, commonly within plants, as highlighted by Canto (2016) and Xu et al., (2023).

The chapter investigates the comparative expression approach which focuses on transient plant expression and heterologous protein expression, aiming to explore different protein production methods. Subsequently, attention is directed towards the steps required to obtain a pure protein sample which can be utilised further to characterise.

4.2 Methods and Materials

4.2.1 Recombinant Production of the Protein

4.2.1.1 Agro-Infiltration of the mutant *N. benthamiana (ΔXTFT)*

The *Agrobacterium-mediated* transient expression system in *N. benth* was used. The constructs of the phage lytic proteins (pEAQ-HT_LysA, pEAQ-HT_LysB, pEAQ-HT_Lysin, and pEAQ-HT_Holin; the schematic representation of the genes in pEAQ-HT can be seen in (Figure A.2) were transformed into AGL1 and LBA4404 electro-competent cells. Cultures were grown in media with respective antibiotics and incubated at 28 °C with shaking. Cells were harvested by centrifugation, resuspended in an infiltration buffer (Table A.1), and prepared for plant infiltration using a process described by Marillonnet *et al.,* (2004) with the presence or absence of an inducer. Different infiltrations OD_{600nm} were used for different cell lines. For AGL1, cells were resuspended and diluted to an OD_{600nm} of 1 or 2, while LBA4404 was 0.4 and 0.6 for optimisation. The infiltrated *N. benth (ΔXTFT)* plants were kept under controlled growth conditions, including a 16/8 h light/dark cycle. The leaf samples were harvested at various times and stored at -80 °C for subsequent analysis.

4.2.1.2 Extraction of expressed protein(s) from the *N. benthamiana* leaves post-infiltration

The harvested leaves of *N. benth* (ΔXTFT) leaves were crushed using a mortar, pestle, and liquid nitrogen. The leaf material was crushed until the material was fine and powdery. A 1:2 dilution was performed using extraction buffer one (15 mM KH_2PO4, 81 mM $NaHPO_4$, 27 mM KCl and 140 mM NaCl, pH 7.4) or extraction buffer two (20 mM sodium phosphate pH 7.2, 20 mM EDTA, 1 mM PMSF, 20 mM Ascorbic acid, 1 mM DTT and 1% Triton-X100) to resuspend the leaf extract. The extract was clarified using the Eppendorf 5415R centrifuge (Eppendorf, HH, Germany) at 10 000 × *g* for 30 min at 4 °C. The resulting supernatant was kept for downstream processing, and the pellet was discarded.

4.2.2 Purification and Analysis of Proteins

4.2.2.1 Immobilised metal ion affinity chromatography (IMAC)

The recovered supernatant was purified using a prepacked IMAC Protino Ni-TED resin column (Macherey-Nagel, Germany)(Porath, 1992). This was done in accordance with the

manufacturer's instruction (http://www.mn-net.com/Portals/8/attachments/Redakteure_Bio/Protocols/Protino/UM_ProtinoNi_TED.pdf). A 1 × elution buffer was used to elute the bound protein. All the fractions of the purification steps were collected and analysed on sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) and western blot to visualise the purified protein.

4.2.2.2 SDS-PAGE and Western Blot Analysis

The sodium dodecyl sulphate page (SDS PAGE) was used to analyse different protein fractions (Laemmli, 1970). Protein samples were mixed with loading dye and boiled at 95 °C for 5 min before loading onto gel. Electrophoresis was performed in 1 x SDS running buffer at 180 V for 40 min. The gels were stained with coomassie staining solution (10% (v/v) acetic acid, 0.125% (w/v) Coomassie brilliant blue G, 25% (v/v) Isopropanol) and followed by destaining using a destaining solution (10% (v/v) acetic acid, 5% (v/v) isopropanol).

Proteins separated on a polyacrylamide gel were transferred to Immobilon- PVDF membranes using a semi-dry transfer apparatus. The membranes were blocked with PBST and then incubated with a monoclonal anti-poly histidine antibody. After washing, the target proteins were detected using Bio-Rad Clarity Western Blot ECL Substrate unless otherwise stated. For visualisation, the ChemiDoc MP System digital imager was used.

4.2.3 Heterologous expression of the lytic protein in *E. coli*

4.2.3.1 Cloning in the pET-30b expression vector

Genscript synthesised plasmids were cleaved and prepared for ligation into a pET-30b(+) vector pET-30b (+) prepared for *Nde*I and *Xho*I. This cloning in pET-30b (+) enabled the His_6 to be incorporated at the C-terminal end of the protein due to the absence of stop after the restriction site on the 3'. Following successful cloning, the plasmid was extracted and stored at -20 °C, and glycerol stocks were prepared from the same inoculum used for plasmid extractions.

4.2.3.2 Protein Expression Studies

To acquire *E. coli* cells harbouring our constructs, the plasmids extracted from section 4.2.3.1 were transformed into *E. coli* BL-21(DE3) cells. Single colonies from of each were inoculated in 5 ml of LB broth supplemented with kanamycin (50 µg/ml) and were incubated at 37 ºC, shaking at 180 rpm for ± 16 h. Overnight (±16 h) grown cultures were subsequently used to inoculate two separate flasks for each construct containing 50 ml of LB broth (unless otherwise indicated) supplemented with respective antibiotics and incubated at 37 ºC, shaking at 200 rpm. When the cultures reached an OD_{600nm} of 0.4 to 0.6, one culture was induced with 0.1

mM IPTG (isopropyl β-D-1-thiogalactopyranoside), and the other flask was not induced. 1 ml fractions were taken every hour for five hours, with an overnight fraction taken at ± 16 h for protein determination of expression levels.

4.2.3.3 Immobilised Metal Ion Affinity Chromatography (IMAC)

Cells from IPTG-induced cultures were pelleted at 10 000 × *g* for 20 min and lysed with B-PER (Thermo Fisher Scientific, Waltham, MA USA). The recovered soluble fraction was purified using an immobilised metal ion affinity chromatography (IMAC) purification system (Porath, 1992), in accordance with the manufacturers recommendations, and SDS-PAGE and western blot analysis were conducted to visualise the purified protein. To optimise purification, modifications were made to the purification process. These modifications varied by purification optimisation (Table A.2).

4.2.3.4 Ammonium Sulphate Precipitation

Solid ammonium sulphate was added to the crude soluble protein fractions to 15% (w/v) ammonium sulphate saturation. The suspension was incubated at 4 °C for 16 h with continuous mixing using a roller drum machine (New Brunswick Scientific, USA). The precipitated proteins were removed by centrifugation (27 000 x *g*, 30 min, 4 °C), and the recovered supernatant fraction was loaded onto the hydrophobic interaction chromatography (HIC) column.

4.2.3.4.1 Hydrophobic interaction chromatography (HIC)

The recovered supernatant fraction from the section 4.2.3.4 was loaded onto pre-packed phenyl Sepharose 26/10 HIC column (GE Healthcare, USA). The column was connected to an AKTA avant 150, using the UNICORN protein purification software. The column containing the unbound proteins was washed with three CVs of equilibration buffer (300 mM $(NH_4)_2SO_4$, 20 mM Tris, pH7.5), at flow rate of 1.5 ml/min followed by gradient elution with HIC elution buffer (1 M Tris. pH 7.5) at the same rate over 20 CVs. The fractions collected were analysed on SDS-PAGE and western blot.

4.2.3.5 Cloning of Lysins to Introduce His_6 Tag on the N-terminal

4.2.3.5.1 Polymerase Chain Reaction (Gradient PCR)

Primer sets were designed to introduce the His_6 tag to the 5' of the gene already in the pET-30b(+) vector, and a stop codon was introduced before the restriction site on the 3' to exclude the C-terminal His_6 tag. Table 4.1 shows the primers designed for the shifting of the His_6 tag. Amplification was achieved using a BioRad Thermo Cycler (C1000 Touch™ thermal cycler). A Q5 PCR kit from NEB Biolabs (New England) was used for the reactions. The thermal

cycling conditions were as follows: 98 °C for 1 min followed by 98 °C for 30 sec; (×) °C for 30 sec; 72 °C for 30 °C sec, and 94°C for 30 sec with go to step 2 for 24 cycles. Then followed by 72 °C for 5 min and last step at 4°C forever. The excepted PCR products were purified using a gel extraction kit (ThermoFischer). The extraction was carried out according to the owner's manual. After extraction, a NanoDrop 2000 Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) was used to determine the concentration and purity of the DNA by measuring the absorbance at 280 nm, 260 nm, and 230 nm.

Table 4.1: Primer sets to introduce the His_6 tag on the N-terminal of the three genes.

Name	Sequence	Primer Set
SM03	5' catatgCACCACCACCACCACCACTACCCGGGTGGCACCA	Forward 1 and 2
	5' cCACCACCACtacccgggtggca	
	3'CTCGAGAATCGGACGCAGCGCCGCCG	Reverse
SM29	5' catatgCACCACCACCACCACCACGCGACCGGTAGCCACC	Forward 1 and 2
	5' caccaccacgcgacc	
	3' CTCGAGAATACGACGGAAACCCGC	Reverse
SM07	5' catatgCACCACCACCACCACCACaccgcgatcattaccc	Forward 1 and 2
	5' cCACCACCACaccgcgatc	
	3' CTCGAGAATGTTGCCTTTCAG	Reverse

4.2.3.5.2 DNA Ligation, Expression, and Purification

The ligation reaction was carried out using a 1:3 insert ratio. Before ligation, DNA was cleaved with the respective enzymes. The reaction comprised of DNA x µl (50 ng), 10X DNA Buffer 1 µl, T4 DNA Ligase 1 µl and vector 1 µl, dH_2O x µl to a final volume of 10 µl. The ligation reaction was used to transform *E. coli* DH5α cells. Colonies were randomly selected and filtered for corrected size amplicon using cleavage using the engineered restriction sites. The confirmed transformants resulted in the expression plasmids designated pET30_LysA, pET30_LysB, and pET30_Lysin for labelling purposes.

The successfully ligated clones were transformed into *E. coli* BL-21(AI) cells. The expression of the cells harbouring the constructs was carried out under similar conditions as in section 4.2.3.2. However, only one temperature profile was used (37 °C) with an additional inducer (0.2% arabinose). Protein fractions were analysed using SDS-PAGE and western blot. Purification was achieved using IMAC under normal conditions.

For larger culture volumes, purification chromatography was automatically monitored using AKTA Avant 150 and the UNICORN protein purification software. Approximately 25 ml of Protino® Ni-TED resin was packed into an XK/16 column. The column was equilibrated with 5 CV of 1 × LEW (50 mM $Na_2H_2PO_4$, 300 mM NaCl, pH 8) buffer at 2 ml/min, and the soluble crude fraction containing a targeted His-tagged protein was loaded onto the column at 1 ml/min. The unbound column proteins were washed with 4 CVs of 1 × LEW buffer at 2 ml/min, followed by elution of the targeted protein with the 1 × elution buffer (50 mM $Na_2H_2PO_4$, 300 mM NaCl, 250 mM imidazole pH 8) at 1ml/min over 7 CVs. Protein fractions were analysed using SDS-PAGE and western blot.

4.2.3.6 Second Purification Step Using Ion Exchange

The successfully purified protein(s) was dialysed to remove imidazole and introduce a buffer (20 mM Tris pH 6.5) compatible with ion exchange for the second purification step. The AKTA Avant 150 Unicorn protein purification software was used. The connected 1 ml column containing the ion exchange resin was washed with 5 CVs at 1 ml/min of start buffer (20 mM Tris pH 6.5), followed by 5 CVs at 1ml/min of the elution buffer, and equilibrated with 5 CVs at 1 ml/min of start buffer before loading the protein at 1 ml/min. Once the protein was loaded onto the column, the column was washed with 5 CVs at 1.5 ml/min of the start buffer to remove unbound proteins. The bound proteins were eluted over 20 CVs at 1 ml/min of elution buffer (20 mM Tris, 1 M NaCl pH 6.5) using gradient elution starting at 0% to 100%. The eluted fractions were loaded on an SDS-PAGE to observe our target protein, and western blot was also conducted. The second step purified protein was quantified using BSA on an SDS-PAGE.

4.2.3.6.1 Bradford Quantification

The soluble fraction protein concentration was determined using gel quantification. The assay involved loading two different volumes of eluted protein (diluted/undiluted) on an SDS-PAGE with standards of BSA (20 mg/ml) (10 µl). A standard curve was constructed as follows using six standards 0.0781 µg, 0.156 µg, 0.312 µg, 0.625 µg, 1.25 µg and 2.5 µg. Absolute quantification was conducted using BioRad Image Lab 4.1 and ChemiDoc MP System digital imaging system (Bio-Rad, Universal Hood III, California, USA).

4.2.4 Peptide Mass Fingerprinting Analysis

The purified protein(s) was analysed using peptide mass fingerprinting. The protein band of interest was cut from a gel and processed for analysis. The gel piece was treated with ammonium bicarbonate and methanol to remove stains, followed by reduction and alkylation steps. Trypsin was used to digest the protein overnight. The peptides were extracted, dried,

and then resuspended in a solution. The samples were analysed using liquid chromatography-electrospray ionisation tandem mass spectrometry (LC-ESI-MS/MS).

The analysis used a TripleTOF MS instrument and a nano-high-performance liquid chromatography system. The peptides were separated using reverse phase chromatography. The sample was first desalted using a C18 trap column and then eluted from the C18 RSLC column using a buffer B. Electrospray ionisation was used to ionize the peptides for mass spectrometry analysis. The data obtained covered a specific range of mass-to-charge ratios, and the collision energies were adjusted accordingly. Protein Pilot software was used for data analysis(Seymour and Hunter, 2015).

4.3 Results

4.3.1 Transient plant expression in *N. benthamiana*

Transient plant expression of the four genes successfully cloned into pEAQ-HT was tested for expression using two strains of *Agrobacterium* at two OD_{600nm} values. For the AGL1 strain, OD_{600nm} of 1 and 2 were tested and harvested at seven days post infiltration (d.p.i) (Figure 4.1 A and B), while for LBA4404, the OD_{600nm} values tested were 0.4 and 0.6 (Figure 4.1C and D). Infiltration was achieved using infiltration buffer 1, with extraction buffer 1 being used for extraction of total protein from leaf material in a 1:2 ratio. The expected molecular sizes of the respective proteins in order of samples loading on the gels are 48.62 kDa, 50.28 kDa, 51.62 kDa, and 13.23 kDa, with no expression observed.

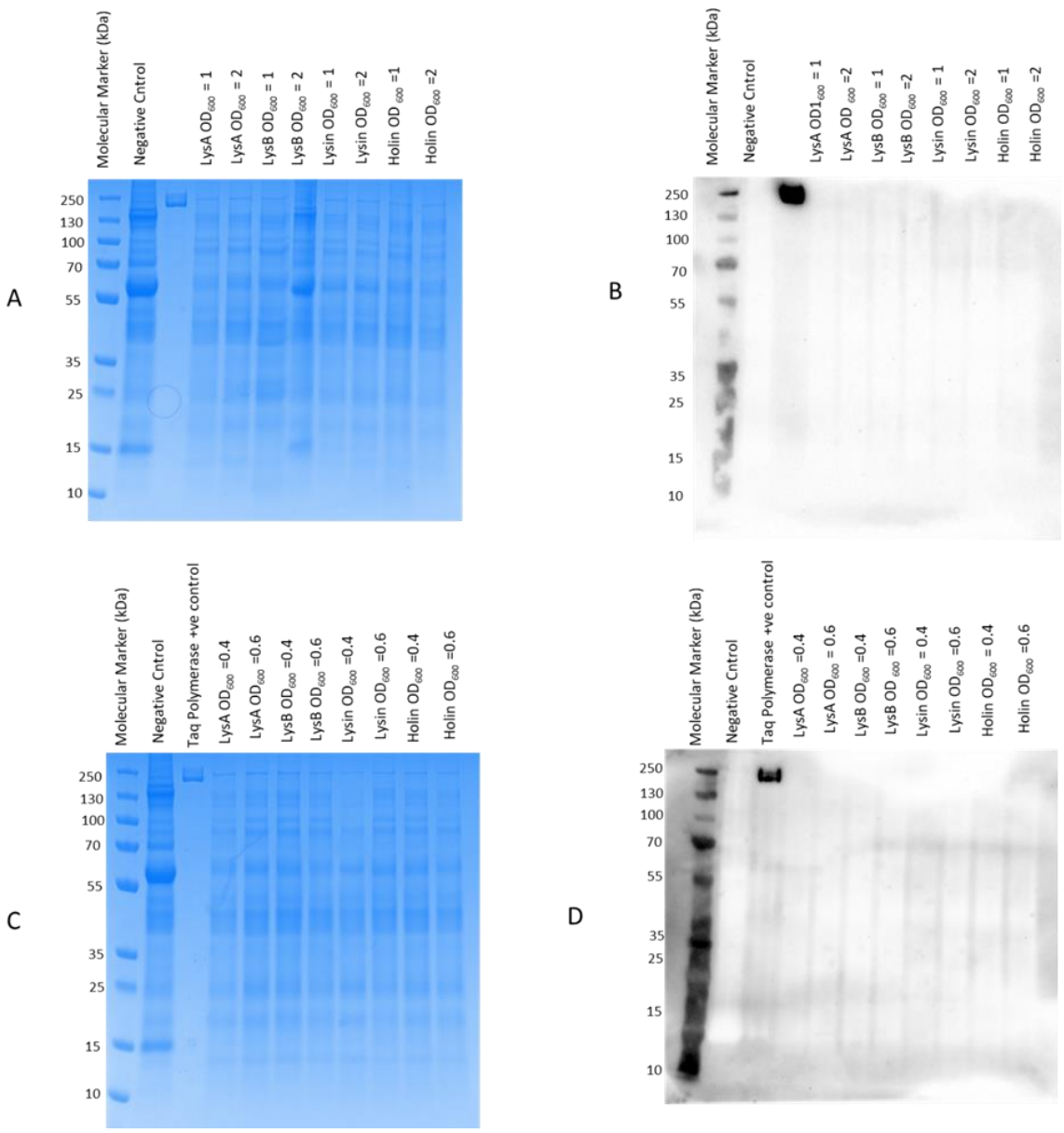


Figure 4.1: SDS-PAGE (left) and western blot (right) analysis of transient expression in *N. benth* of the lytic proteins.

Plants harvested seven d.p.i. (A) and (B) depict two different OD_{600nm} used for AGL1. While (C) and (D) depict two different OD_{600nm} used for LBA4404.

The LBA4404 strain was selected for further expression studies. These studies included the optimisation of the infiltration buffer and the protein extraction buffer (Table A.1). Time profile studies of the expression were carried out up to 15 d.p.i, with harvesting every second day from day 3 post infiltration. Expression was commenced when cells reached an OD_{600nm} of 0.4. Figure 4.2 shows the comparison results using infiltration buffers 1 and 2 with different extraction buffers. Figure 4.2 A and B show extraction with extraction buffer 1, while C and D

used extraction buffer 2. Although the estimated molecular mass for our proteins LysA (SM03), LysB (SM29), and Lysin (SM07) is estimated to range from 48 kDA to 52 kDa, a band can be seen between 15 kDa and 25 kDa on the western blot (Figure 4.2 D) except for LysA (SM03) using infiltration buffer 1. No band is observed for Holin (SM31) under any of the conditions tested.

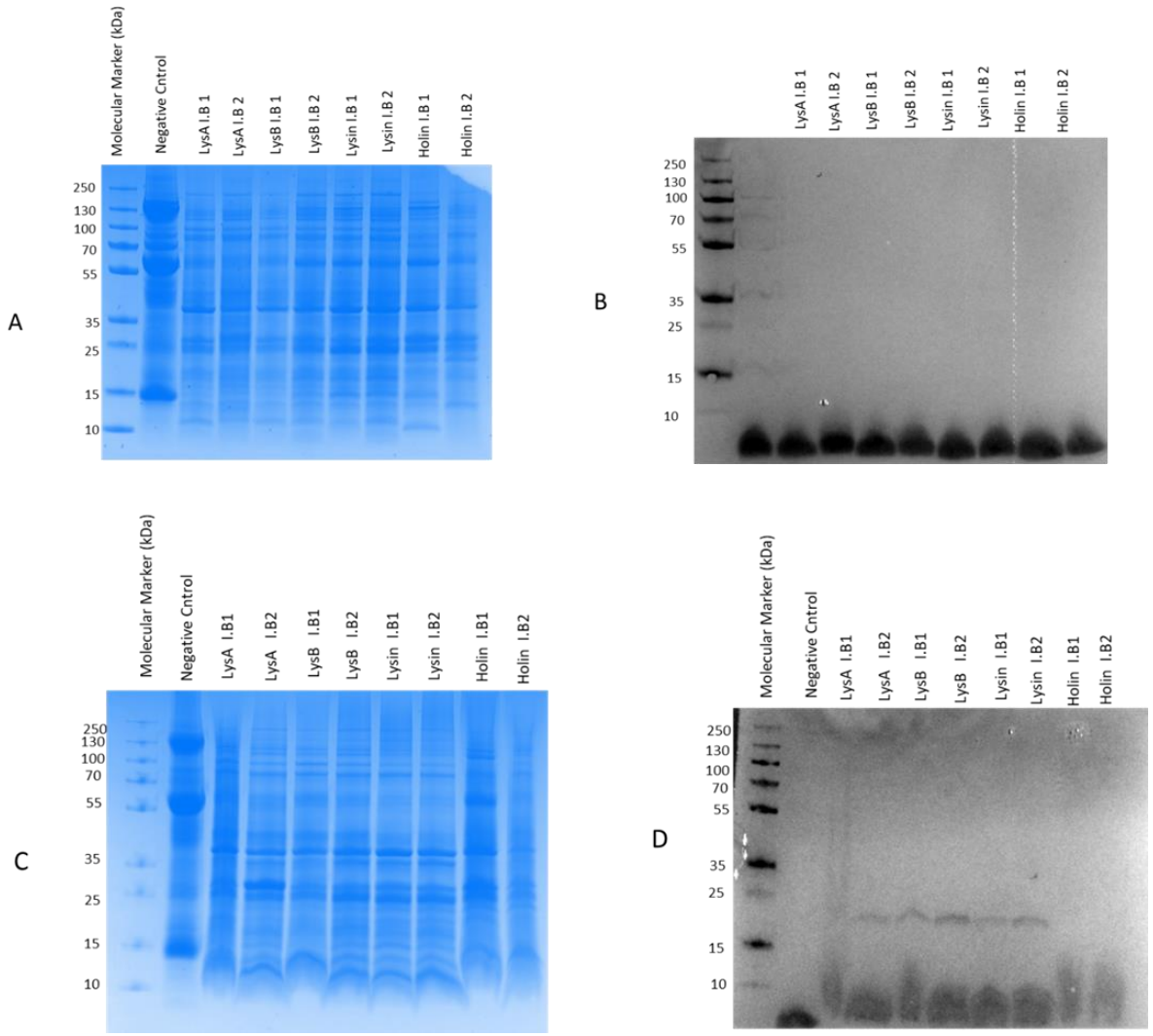


Figure 4.2: SDS-PAGE (left) and western blot (right) analysis of two extraction buffers used. Expression for LysA (SM03), LysB (SM29), Lysin (SM07) and Holin (SM31) were conducted in LBA OD_{600nm} 0.4.

A time profile study was conducted using expression conditions that resulted in the appearance of a His_6-tagged band on a western blot gel. In some cases, the band detected was not the anticipated size. Leaf discs were harvested from day three post infiltration, and every two days thereafter until day 15. The expression was also carried out in the presence of an inducer, acetosyringone (aceto), at a final concentration of 100 µM. Figure A.3 and

Figure A.*4* show the expression time profile of the four proteins in the presence and absence of aceto. The western blot analysis shows a band (truncated version (± 20 kDa)) of the proteins except for the Holin (SM31). A variation of the expression levels of the truncated version can be observed either in the absence of aceto or, for some, in both the absence and presence of aceto when analysing the protein western blots.

Purification was attempted using clarified lysates of the proteins containing a truncated protein visible by western blot analysis. The proteins were tagged with a His_6 to facilitate easy purification using IMAC Ni-TED resin. As seen in Figure 4.3, the purification did not yield any positive results, as only a band at ± 20 kDa was observed in the western blot (Figure 4.3B). Due to the poor results obtained from the transient plant expression work, it became necessary to explore alternative approaches. Consequently, we shifted our focus towards using a bacterial expression system to address the objective of our study.

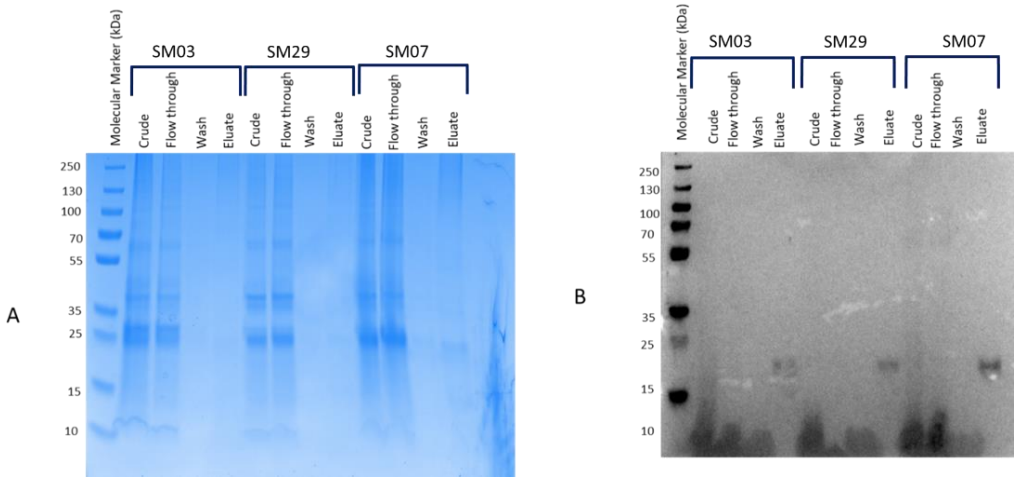


Figure 4.3: Purification of LysA (SM03), LysB (SM29), and Lysin (SM07) using Ni-TED from LBA at 15d.p.i, infiltrated with infiltration buffer 2 and extracted with buffer 2.

4.3.2 Subcloning and Bacterial Expression

Following the challenges encountered with plant expression, a different approach was taken to clone the target gene(s) into a bacterial expression system. The vector selected for bacterial

expression was pET-30b(+) under the T7 promoter. The vector was prepared for ligation by restriction endonuclease digestion using *Nde*I and *Xho*I. Figure 4.4 shows the confirmation of the ligated constructs into pET-30b(+). Expression results for these genes showed inconclusive transient expression in plants. Figure A.5 is a graphic representation of the genes ligated into pET-30b(+). Double digesting of the plasmid was done to confirm successful ligation before expression studies.

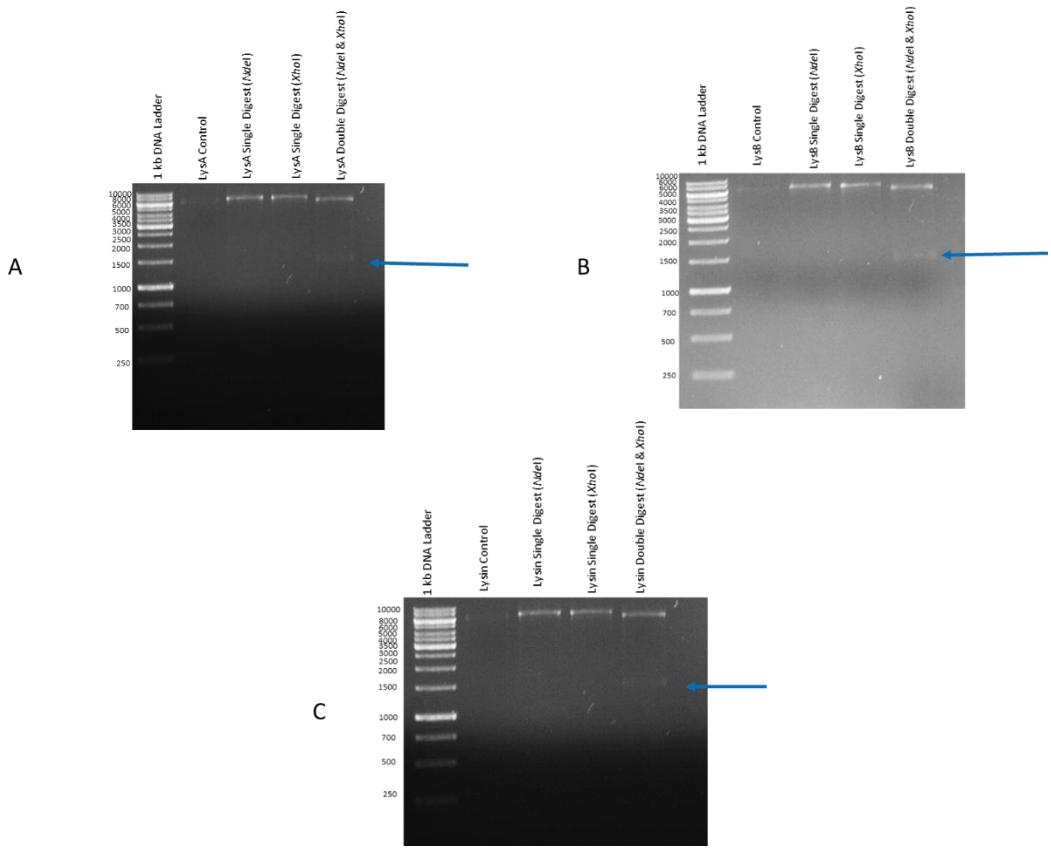


Figure 4.4: 1% Agarose gel images showing the insert confirmations ligated into pET-30b(+).
Key: LysA (SM03), LysB (SM29) and Lysin (SM07).

Having successfully cloned the target genes into the pET-30b(+) expression vector, expression studies were performed. The expression studies were carried out under different incubation temperatures (17 °C, 25 °C, 30 °C, and 37 °C), in the presence or absence of IPTG (final concentration of 0.1 mM IPTG used).

Time profile expression studies were carried out by growing a single colony of the transformants in BL-21(DE3) *E. coli* in 5 ml of LB overnight (±16 h) at 37 °C, 200 rpm supplemented with the respective antibiotics. From the overnight inoculum, 1 ml was taken into 50 ml of LB and incubated at 37 °C, 200 rpm until the OD_{600nm} was between 0.4 to 0.6.

Once the desired OD_{600nm} was obtained, the flasks were induced with 0.1 mM final concentration IPTG and incubated at the mentioned temperatures. Samples were taken hourly for five hours post induction, with an overnight sample taken at ± 16 h.

The gels depicting the studies of the three proteins under these different conditions can be seen in Figure 4.5 to Figure 4.7 with the insoluble fraction analysis in Figure A.6 to Figure A.8. SDS-PAGE analysis shows that the proteins, more specifically SM07, are visible in the insoluble fraction. The gels demonstrating expression profiles both in the presence and absence of IPTG do not show any data variation with the soluble expression levels being deficient.

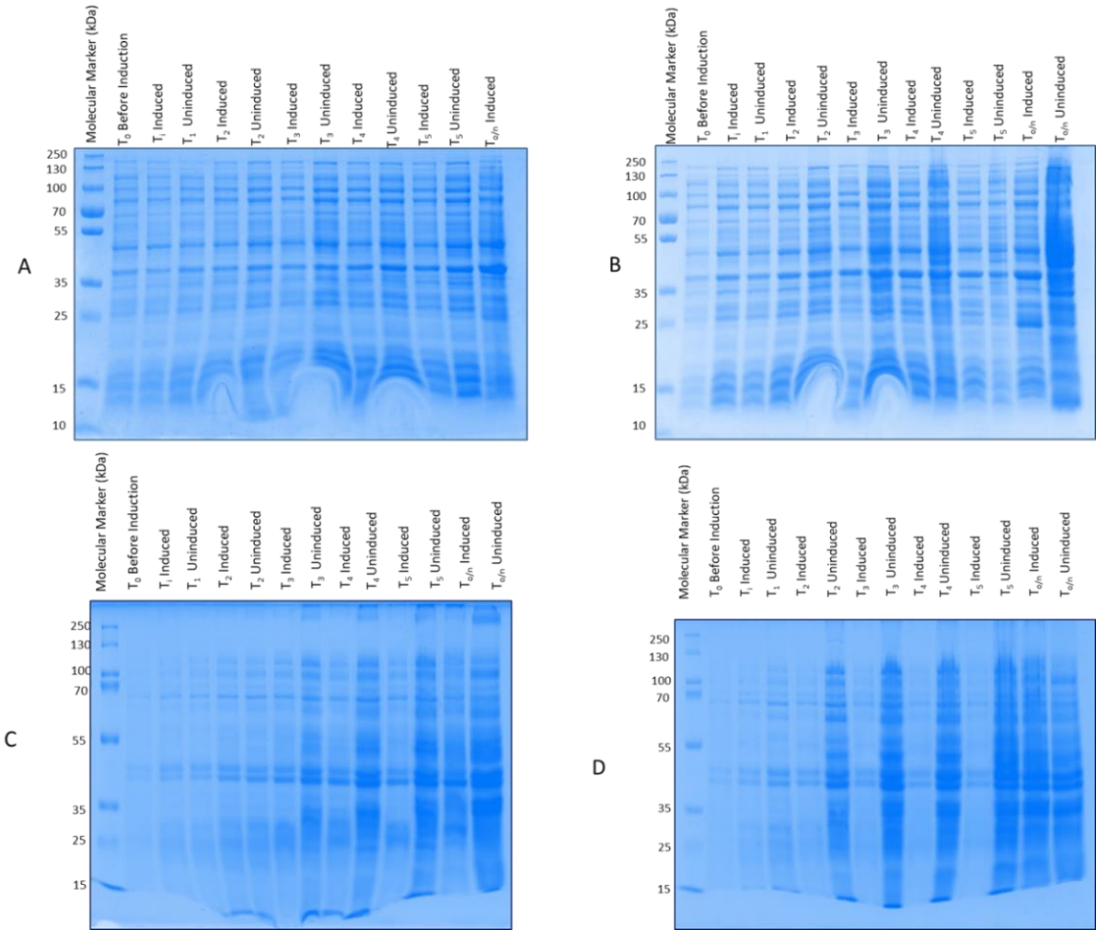


Figure 4.5: SDS-PAGE analysis of the time profile of SM07 using LB in BL-21 (DE3) cells at four different temperature points.

(A) 17 °C, (B) 25 °C, (C) 30 °C, and (D) 37 °C, respectively. The figures depict the induced and uninduced fractions.

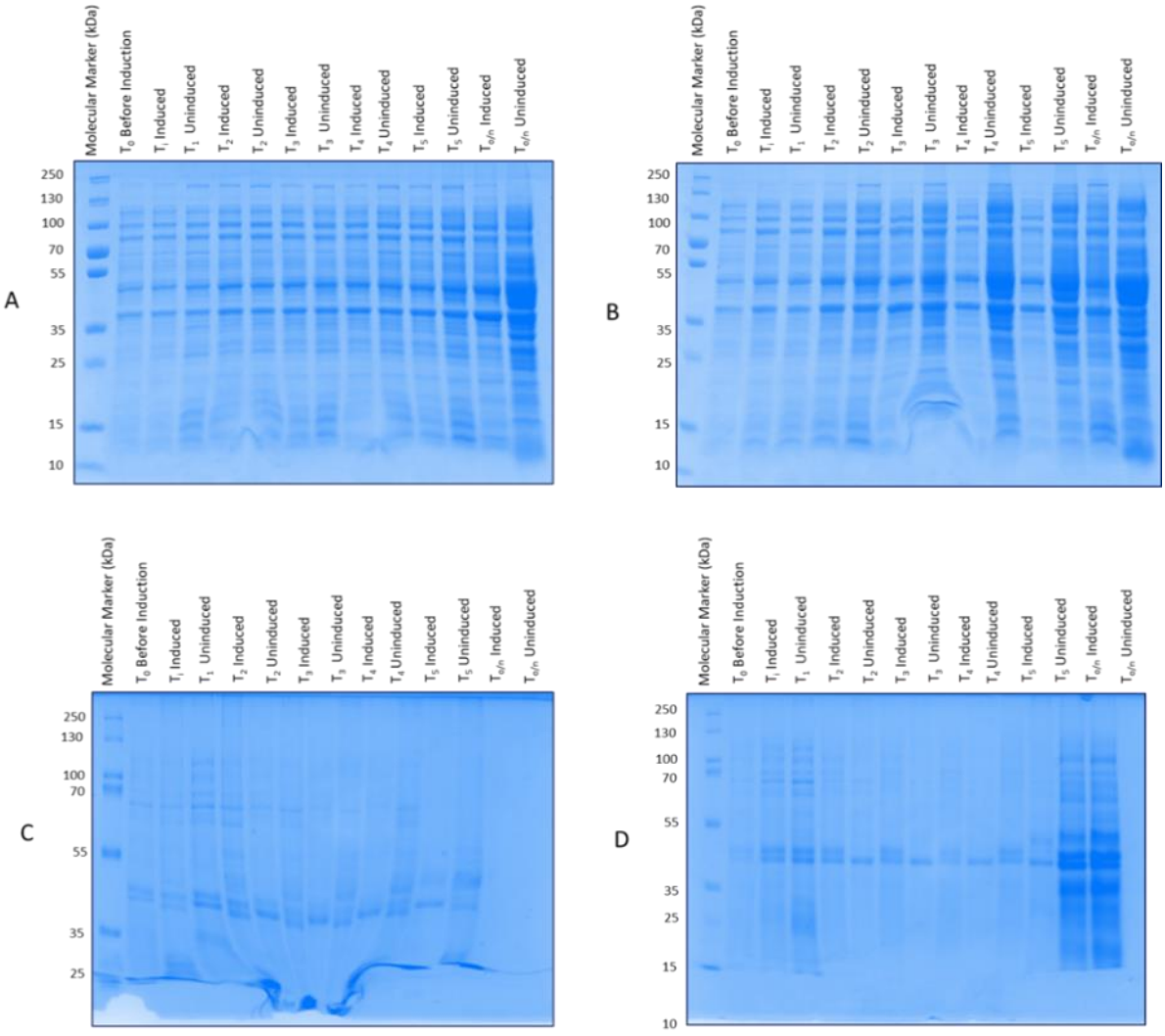


Figure 4.6:SDS-PAGE analysis of the time profile of SM03 using LB in Bl21 (DE3) cells at four different temperature points.

(A) 17 °C, (B) 25 °C, (C) 30 °C, and (D) 37 °C, respectively. The figures depict the induced and uninduced fractions.

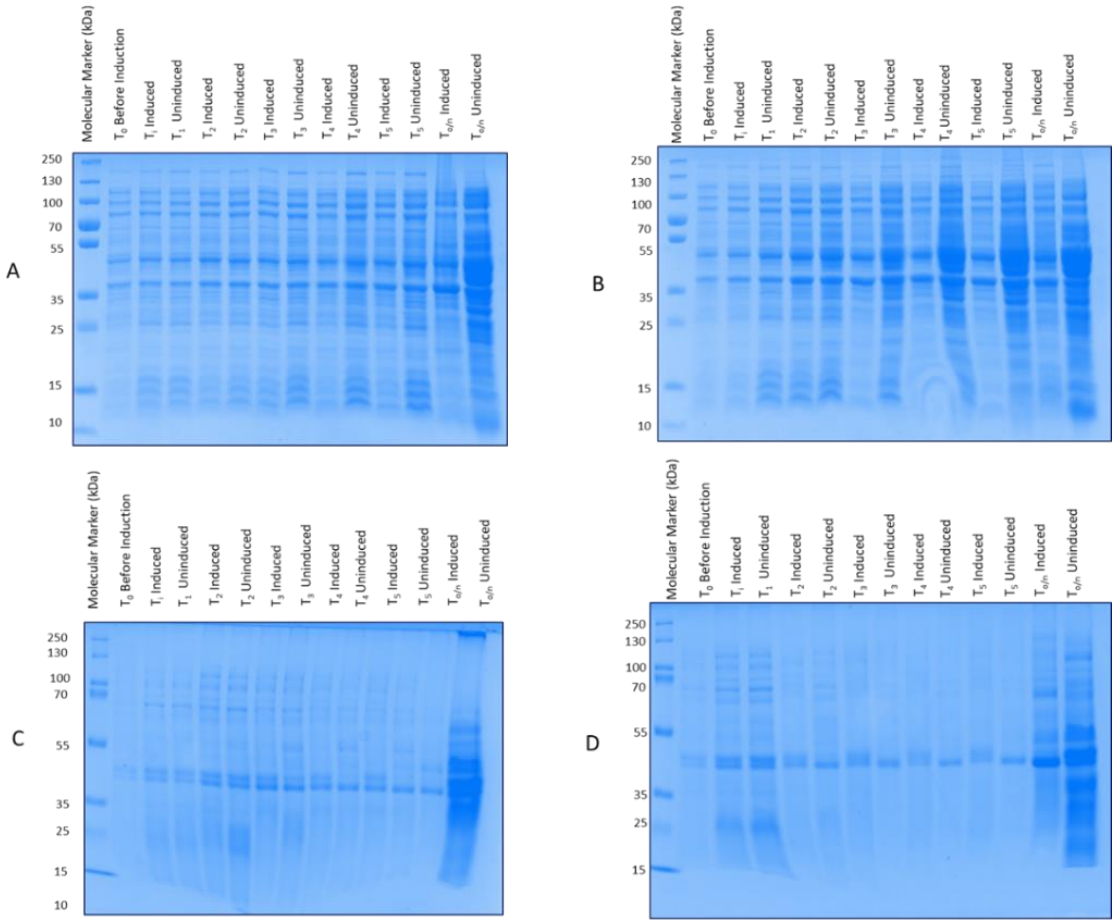


Figure 4.7:SDS-PAGE analysis of the time profile of SM29 using LB in BL21 (DE3) cells at four different temperature points.

(A) 17 °C (B) 25 °C, (C) 30 °C, and (D) 37 °C, respectively. The figures depict the induced and uninduced fractions.

Although only a small amount of protein was visible from the expression gels shown above, protein purification using immobilised metal affinity chromatography was attempted (this would involve protein capture via a C-terminal linked histidine tag). Clarified lysates were purified from expression studies carried out at 37 °C, with a 0.1 mM IPTG concentration being harvested 4 h post-induction (Figure 4.8).

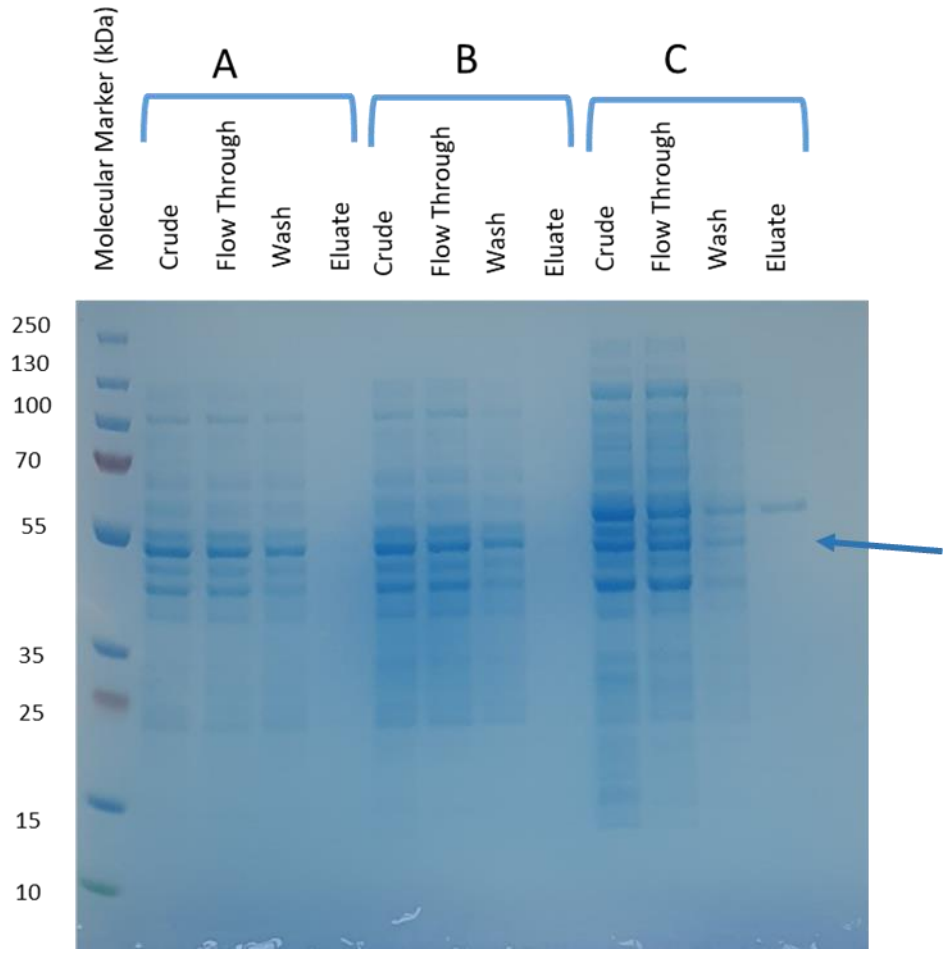


Figure 4.8: SDS-PAGE purification analysis of (A) SM03, (B) SM29 and (C) SM07 using Ni-TED prepacked columns under normal conditions.

Despite our initial attempts at purifying the required proteins, Ni-TED yielded unsatisfactory results. Further optimisation of the Ni-TED purification protocol was needed to enhance the efficiency of the purification step. These measures included altering the protein extraction conditions (Table A.2). DTT addition or removal from the lysing buffer showed little effect (Figure 4.9). A lot of the proteins can be observed in the flow-through and wash fractions, indicating poor protein binding onto the resin.

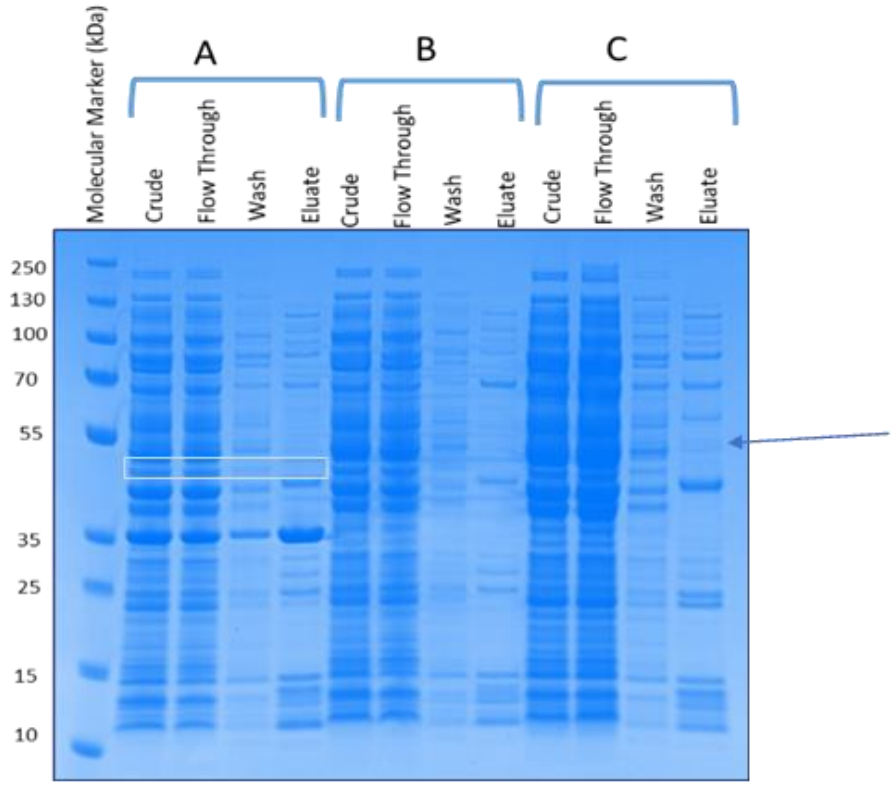


Figure 4.9: SDS Page analysis purification using Ni-TED prepacked columns with elution buffer containing an additional 0.5 mM DTT. (A) SM03, (B) SM29, and (C) SM07. The arrow pointing to the band corresponding with SM07 and the box (white) corresponding to SM03 purification fractions.

To address the potential issue of poor protein binding to the resin, the hypothesis was that the observed difficulty could be the result of protein folding during expression, potentially hiding the histidine tag. To address this concern, denaturing salts and detergents were introduced into the sample to relax the protein structure, thus potentially improving the proteins' binding capacity. These however proved unsuccessful (Figure 4.10) with no corresponding band being observed via SDS-PAGE analysis of the eluate. Bands of interest were observed in the flow through as well as in the crude fraction.

Having shown little to no improvement with Ni-TED, even after the introduction of denaturing salts and the addition of detergents, hydrophobic interaction chromatography was attempted. Ammonium sulphate was added to a clarified lysate at 15% (w/v) and the protein precipitated overnight at 4 °C, with sample mixing by rolling. The recovered lysate (post centrifugation) was purified using a phenyl sepharose resin. Protein purification using HIC was successful for SM07 but did not work for SM03 and SM29, respectively (Figure 4.11). Confirmation of SM07 purification was confirmed through western blot analysis since not much could be observed from the SDS-PAGE run.

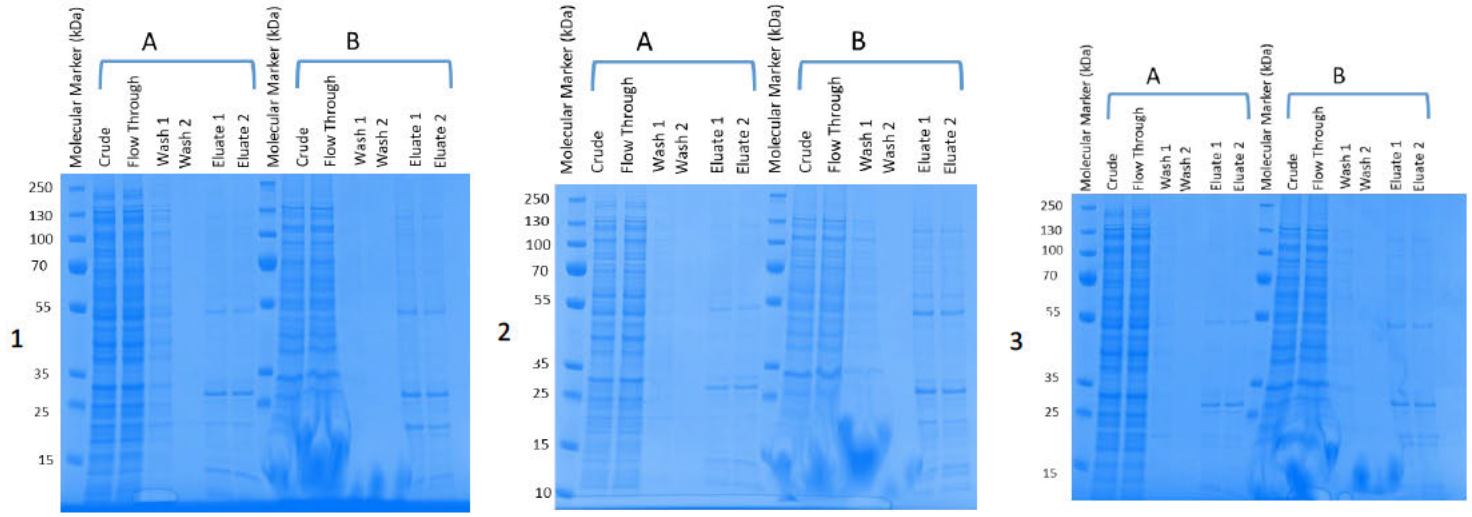


Figure 4.10: 12% SDS-PAGE analysis purification (1) SM03, (2) SM29, and (3) SM07 using Ni-TED prepacked columns. (A) Triton and (B) Urea in the LEW buffer and standard elution buffer.

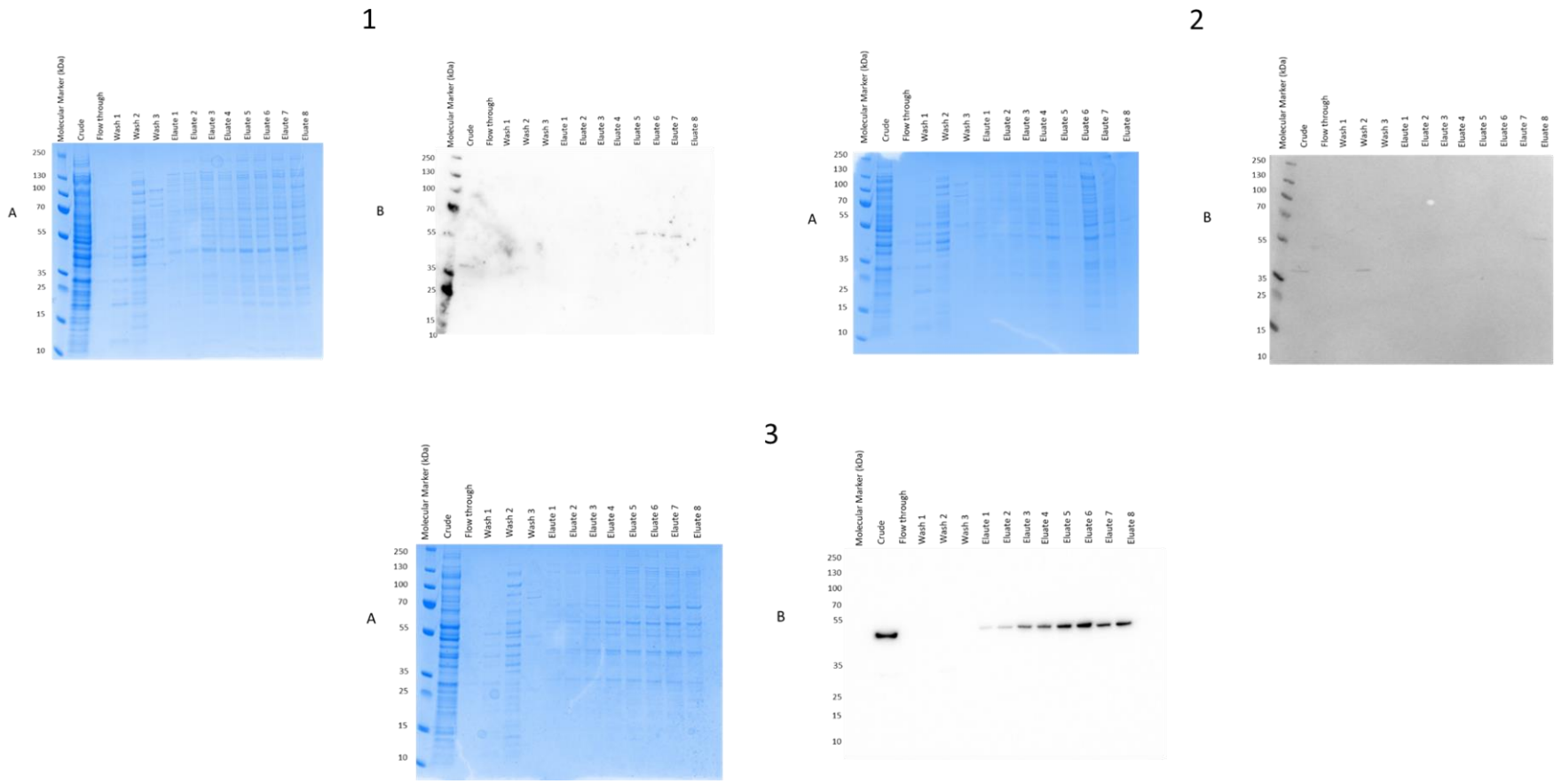


Figure 4.11: SDS-PAGE and western blot analysis of HIC purification after 15% precipitation with ammonium sulphate. Precipitation was conducted overnight at 4 °C, rolling. Key: 1-SM03, 2-SM29, and 3-SM07 .

Positioning of the His_6 tag at the C-terminal end could potentially limit protein accessibility to the nickel resin, thereby impacting the protein's ability to efficiently bind to the resin (thus impacting purification yields). To address these challenges, optimisation of the position of the His tag was deemed necessary to ensure efficient purification without compromising protein integrity or functionality. Using PCR techniques, the His_6 tag was introduced on the 5' of the gene and introduced a stop codon after the restriction site on the 3' end to exclude a His tag that would be on the C-terminus of the protein. This approach was anticipated to facilitate protein purification using Ni-TED column purification, eliminating the need for multiple purification approaches that would be time-consuming and costly.

The BL-21(AI) cell line was used for expression while maintaining the experimental conditions (37 °C, 0.1 mM IPTG with the addition of arabinose) established in the previous section to improve expression

Figure 4.12 to

Figure *4.14*) and purification.

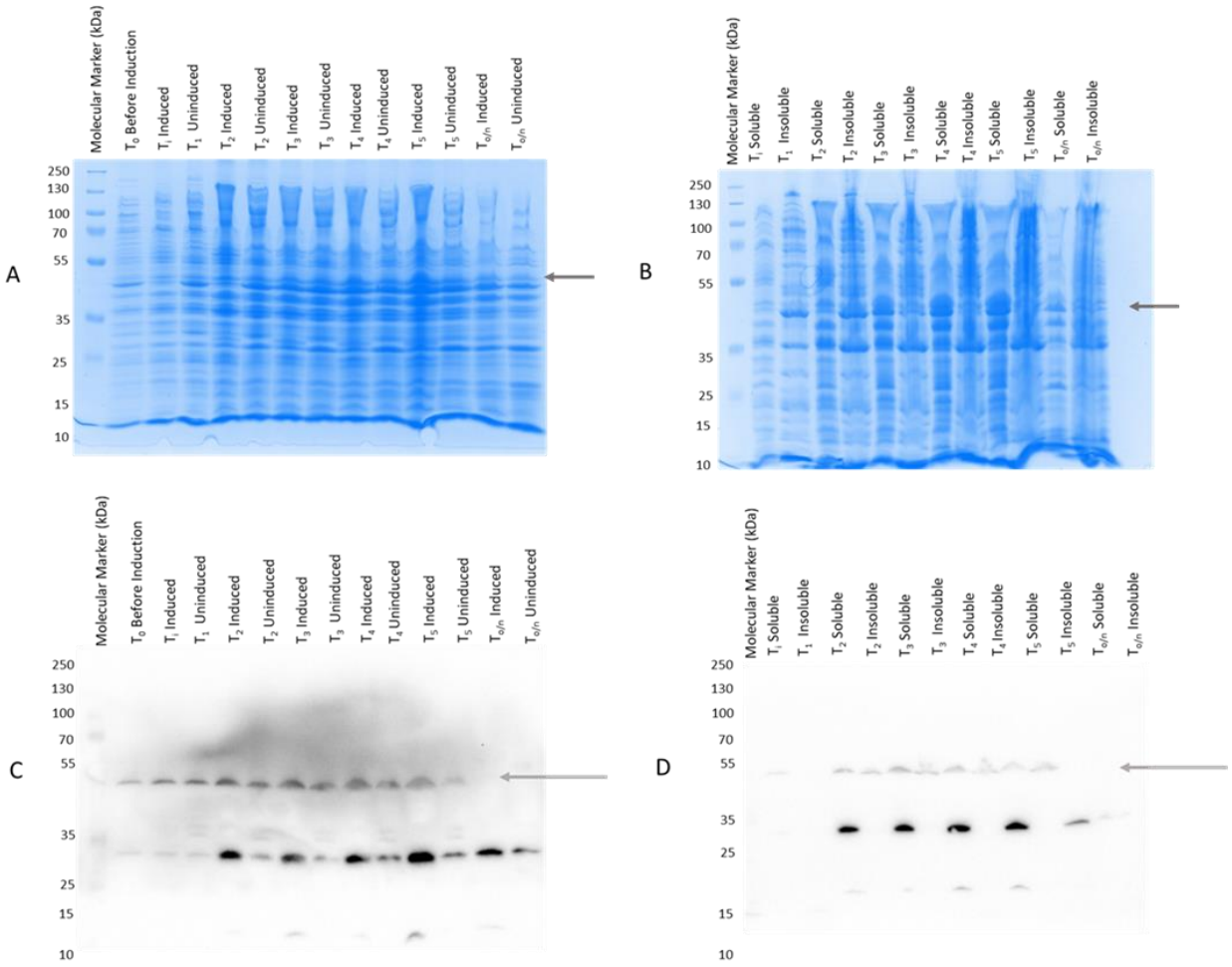


Figure 4.12: SDS-PAGE (top) and Western Blot (bottom) analysis of LysA (SM03) time profile in BL-21(AI) at 37 °C.
(Left)Time profile analysis of the induced and uninduced expression fraction and (Right) Time profile analysis of the soluble and insoluble expression fraction.

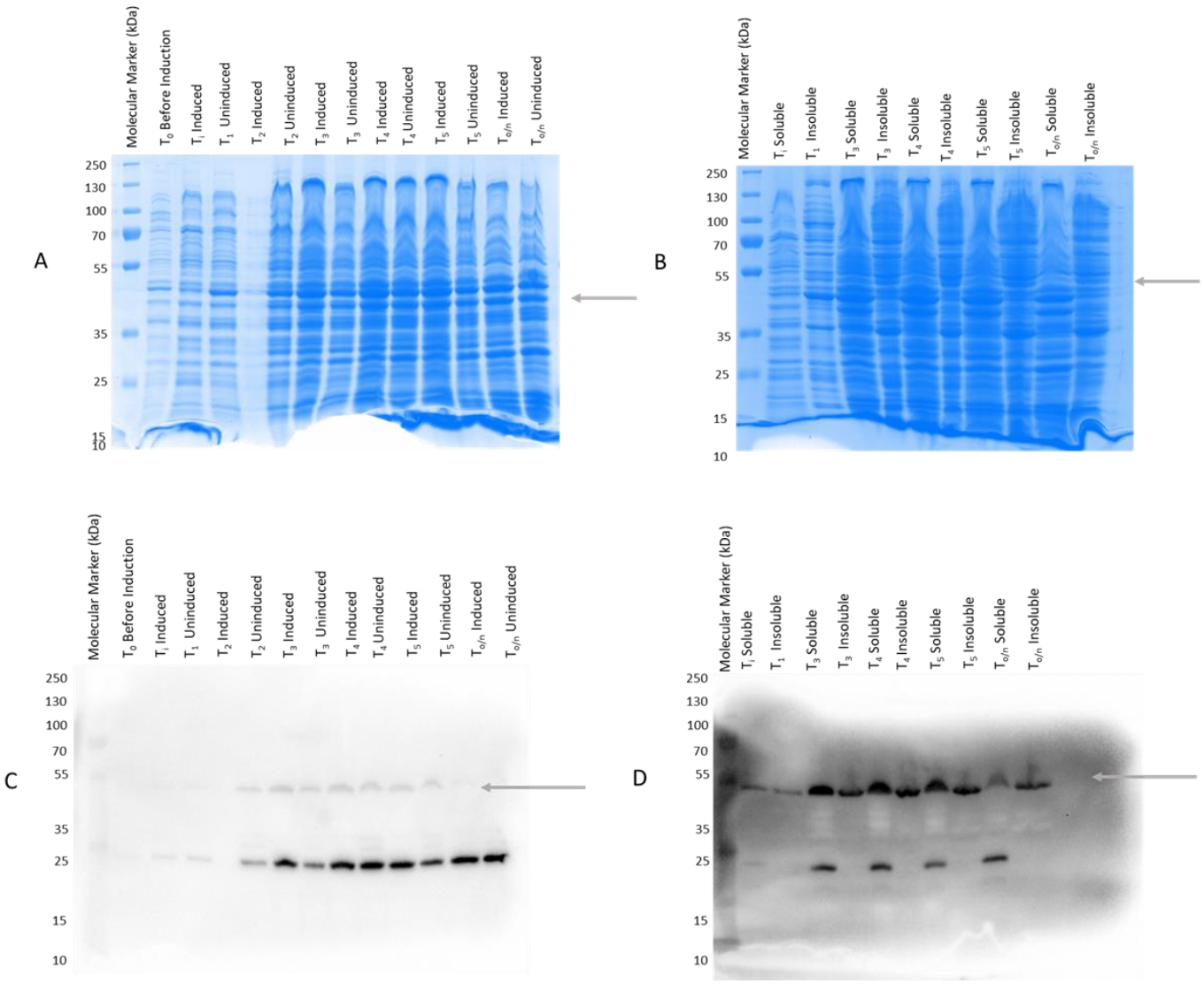


Figure 4.13: SDS-PAGE (top) and western blot (bottom) analysis of LysB (SM29) time profile in BL-21(AI) at 37 °C.

(Left)Time profile analysis of the induced and uninduced expression fraction and (Right) Time profile analysis of the soluble and insoluble expression fraction.

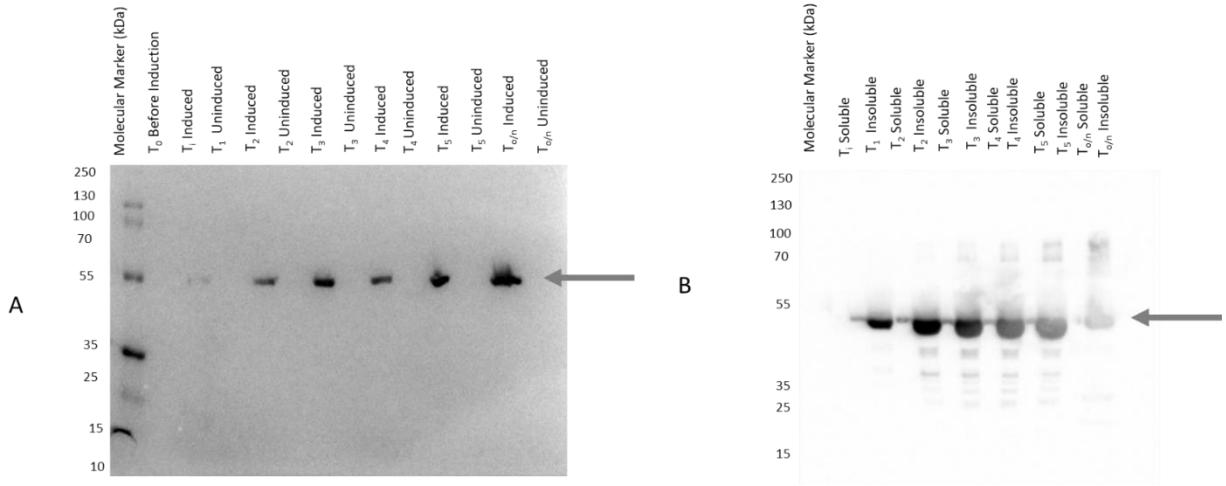


Figure 4.14: Western blot analysis of Lysin (SM07) time profile in BL-21(AI) at 37 °C.

(A) Time profile analysis of the induced and uninduced expression fraction and (B) Time profile analysis of the soluble and insoluble expression fraction.

After expression studies, purification was conducted using IMAC Protino Ni-TED columns with cells harvested 4 h after induction. SM03 was not eluted from the column, indicating inability to purify the protein, and a light band was observed with SM29. However, with SM07, a band corresponding to the estimated size of SM07 was observed as seen in Figure 4.15.

.

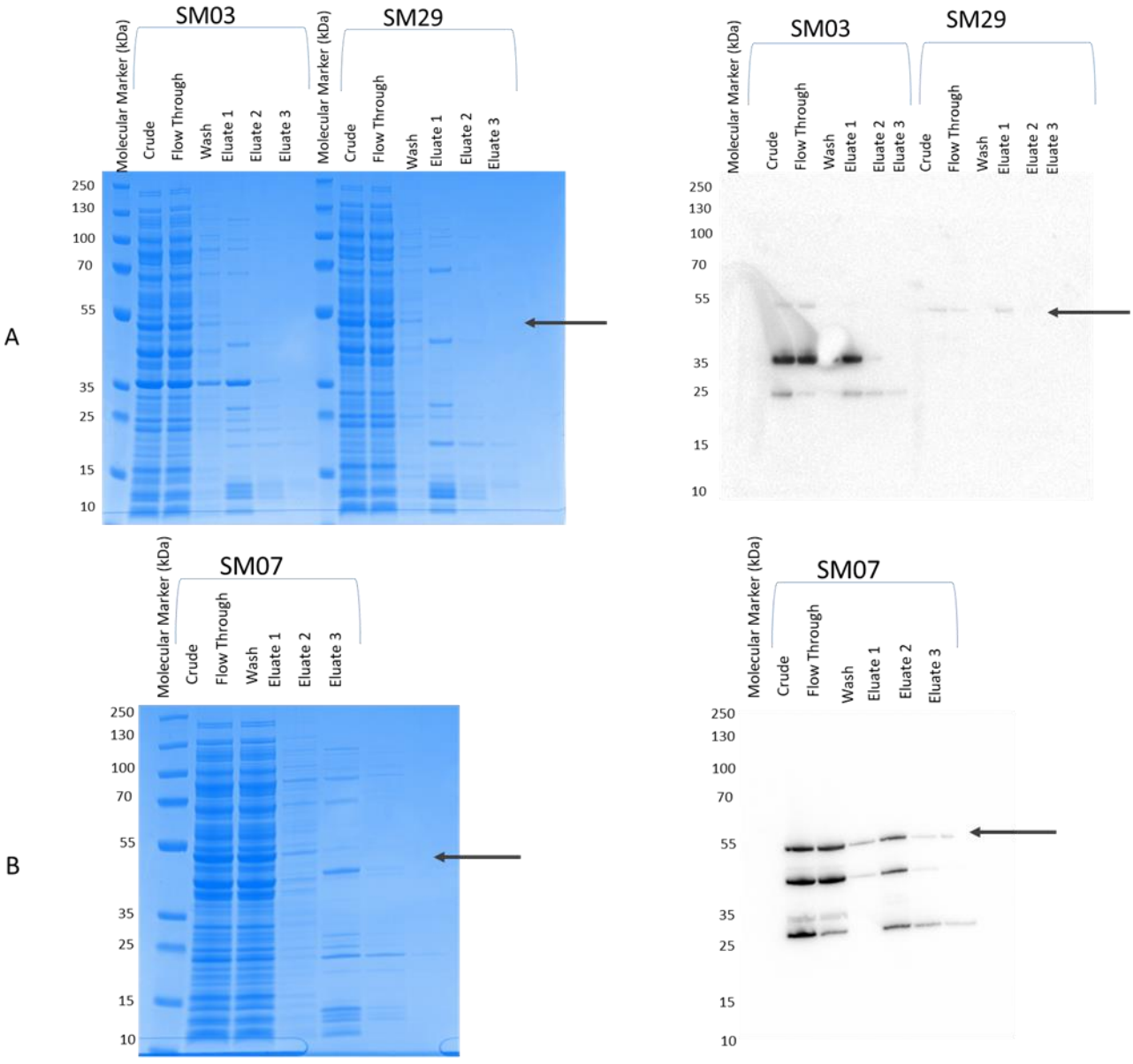


Figure 4.15: Ni-Ted purification of the lytic proteins expressed in BL-21 (AI) at 37 °C and harvested 4 h post induction.
(A) SDS-PAGE on the left of SM03 and SM29 and western blot analysis on the right. (B) SDS-PAGE on the left of SM07 and western blot analysis on the right.

The successfully purified SM07 was upscaled to increase biomass. Purification was achieved using an AKTA Avant 150 and was subsequently subjected to a second purification step to remove any background proteins. This was achieved using ion exchange chromatography. The two-anion resins (CaptoDEAE and CaptoQ) were used. Figure 4.16 reveals the resins'

performance; the best-performing resin can be seen as the CaptoQ. Figure 4.17 illustrates the purification process, from IMAC to anion exchange and the fractions taken from each process. A bovine serum albumin (BSA) standard was used to accurately quantify the polished protein (Figure A.9).

Despite the challenges posed by low expression levels, the observation of a His-tagged band on the western blot indicates the presence of SM07 protein. The sample was submitted for peptide mass fingerprinting analysis to again identify the protein. The results of the peptide mass fingerprinting yielded endolysins with a 5.3% coverage demonstrating 95% confidence. Six peptides were identified from that coverage and three with a confidence of 95% and above. Figure 4.18 shows the three peptides identified with a confidence of more than 95%.

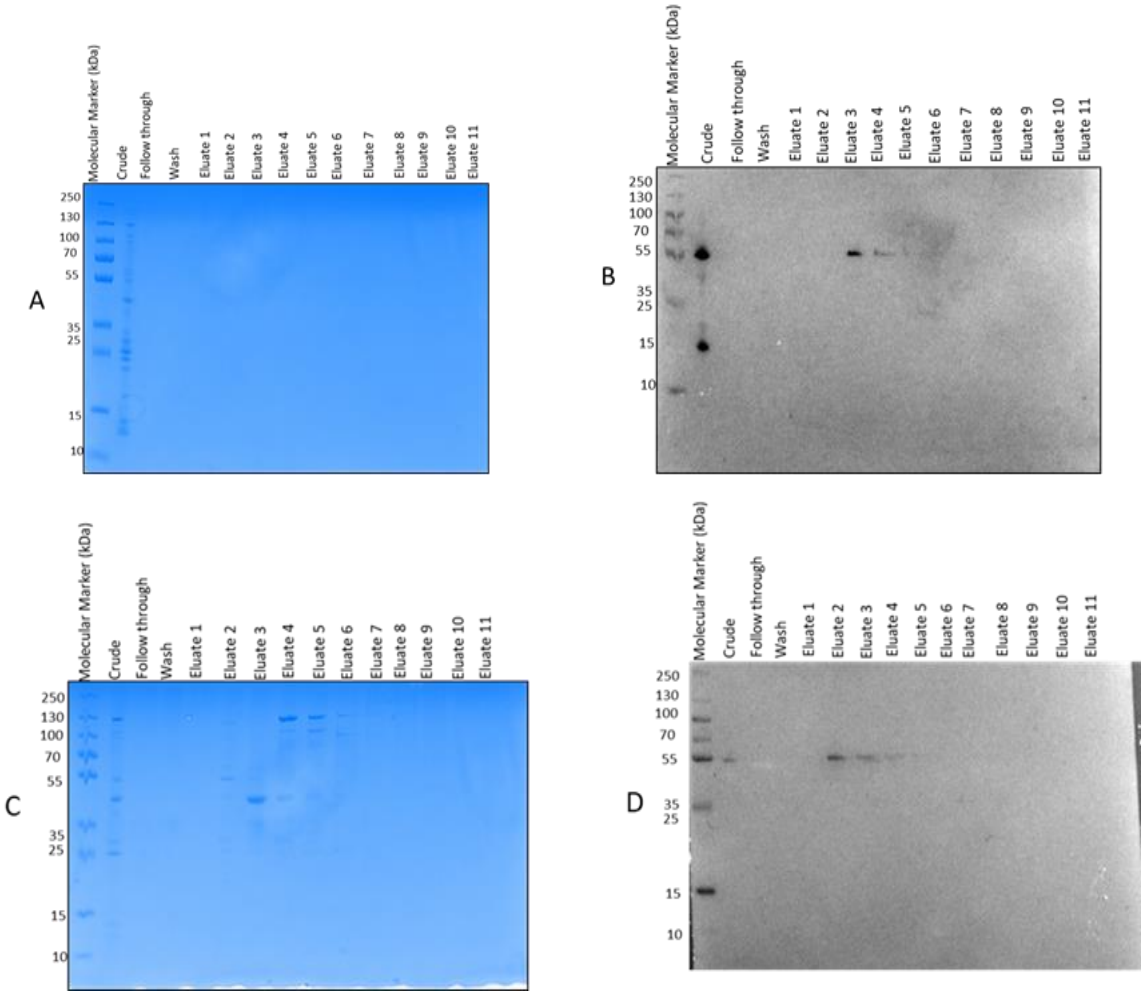


Figure 4.16: SDS-PAGE (Left) and western blot (Right) analysis of second-step purification using anion exchange.
CaptoDEAE is represented by A & B, and CaptoQ is represented by C & D.

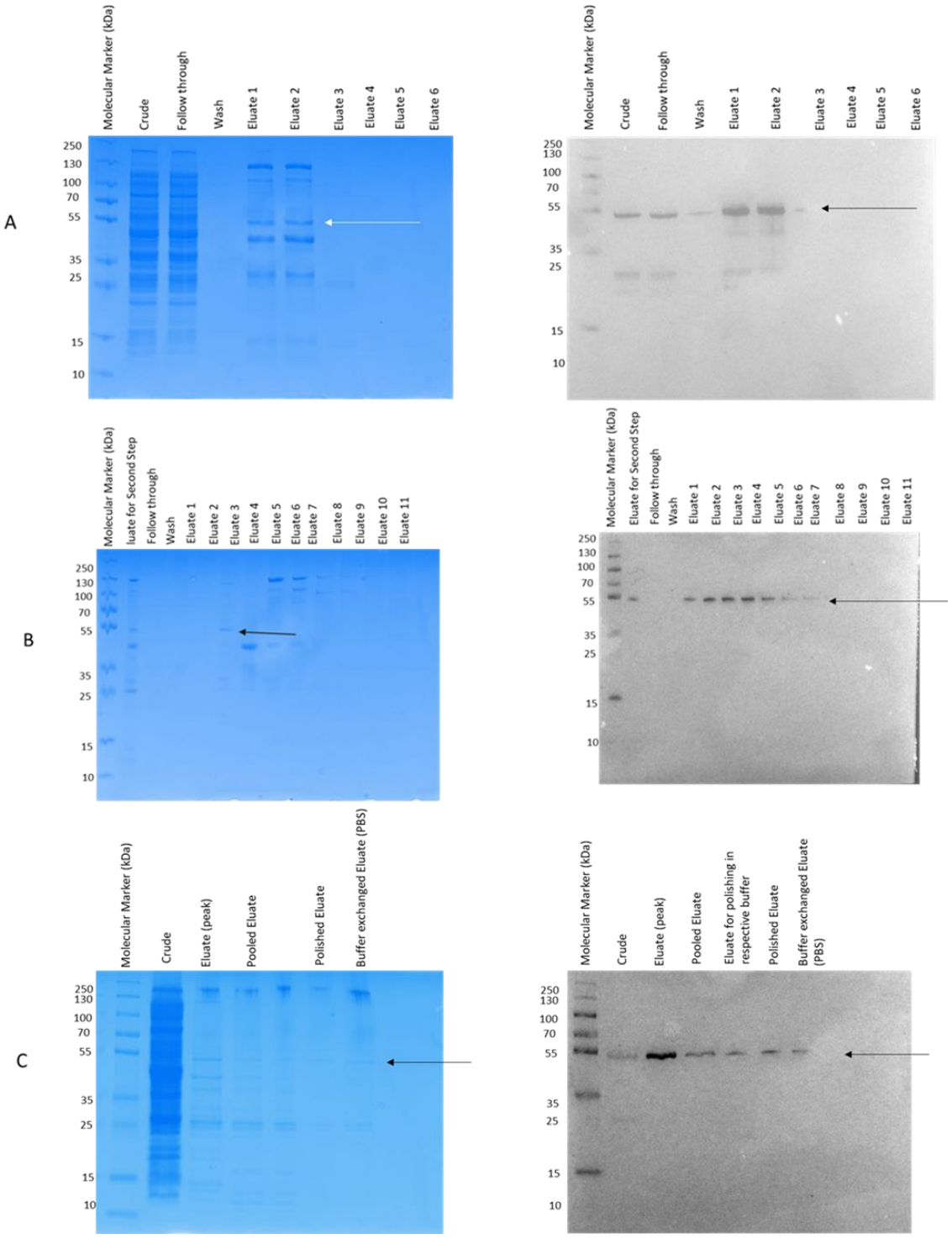


Figure 4.17: SDS-PAGE (Left) and western blot (Right) analysis of the purification of SM07 including polishing step.

(A) Shows the initial purification using IMAC Ni-TED resin, (B) indicates the purification of the eluate from IMAC in an appropriate buffer for the polishing step using CaptoQ anion exchange, and (C) shows the different fractions from IMAC to CaptoQ.

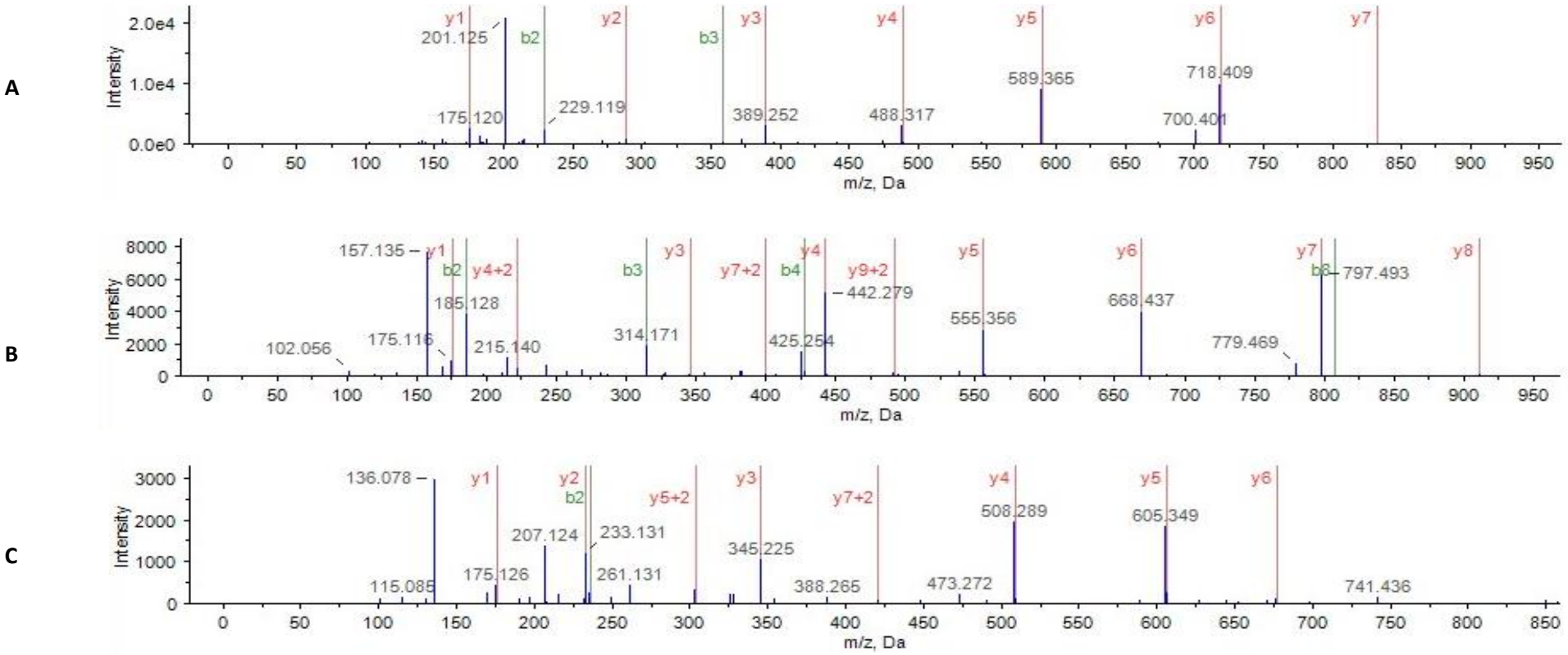


Figure 4.18: Peptide mass fingerprinting of SM07 using Protein Pilot from ionisation spectra.

Analysis revealed three peptides A (DLETVTLR), B (ALEILPAVR), and C (YAPYIGR) identifying SM07 as an endolysin.

4.4 Discussion

Exploring diverse methodologies for producing recombinant proteins encompasses many strategies, extending beyond transient expression in plants and bacterial expression in *E. coli* methods. In this study, two approaches were used to recombinantly produce phage lytic proteins to answer the call for alternative antimicrobials for antibiotic resistant bacterial strains. The utilization of transient plant expression to produce phage-derived proteins has shown promise in studies such as those conducted by Hammond et al., (2019) and Kovalskaya et al., (2019). This method has been instrumental in optimising expression within plant systems, presenting a valuable avenue for creating novel antimicrobial agents. However, our specific transient expression efforts yielded inconclusive results, resulting in the production of a truncated protein, as confirmed by western blot analysis. These unsatisfactory results could stem from various factors, including suboptimal codon usage and protein instability during expression. Inadequate codon optimisation could lead to minimal or no protein expression, as evidenced in previous studies by Komarova et al., (2010) and Mauro (2018). In contrast, protein instability during expression could result in degradation, leaving behind fragmented or truncated signals detectable by western blot analysis. These findings can be improved for better expression yields by exploring other expression vectors for plant expression tailored for optimised expression and stability.

To overcome limitations in transient expression, the bacterial expression system using *E. coli* was used; cloning the genes into a pET-30b(+) vector with modifications in tag placement. His_6 nucleotides were incorporated at the 5' and 3' end of the genes. BL-21 cells were used for protein expression studies due to their suitability for recombinant protein production, and their lack of Ion and ompT proteases (Rosano and Ceccarelli, 2014; Fathi-Roudsari *et al.*, 2016a). Initial expression level showed high levels of insoluble proteins in BL-21(DE3) cells. Different temperature conditions were tested thereafter with proteins still observed predominantly in the insoluble fraction. The temperatures were selected to optimise expression ensuring proper folding, stability and maintaining bacterial viability. Although urea solubilisation could mitigate this, considerable functional loss is a concern and the process is tedious (Singh *et al.*, 2015). Expression of the N-terminal His_6 tagged protein in BL-21(AI) cells achieved low-level solubility, which could potentially be related to phage-derived protein toxicity to the host cell, thus explaining the poor expression in BL-21(DE3) cells and the suspected leaky expression. Improved expression in BL-21(AI) cells can be attributed to the tight expression control preventing toxicity or aggregation issues and arabinose regulation of expression functions better. Due to limitations in detecting the target proteins within the crude lysate using SDS-PAGE throughout the time course, purification was done from samples

collected 4 hours after induction. The absence of bands corresponding to the expected molecular mass for the protein may be due to low expression and overshadowing by background proteins. Purification can increase the probability of isolating target proteins by purification.

In our study, amino acid sequences encoding for His_6 were introduced at either end of the genes via PCR amplification in pET-30b(+) to facilitate protein purification by immobilized metal affinity chromatography (IMAC). Despite the advantages of IMAC, challenges were encountered in achieving adequate purification, particularly with the C-terminal-tagged proteins, even under denaturing conditions. IMAC relies heavily on the strong interaction between histidine residues(Wong *et al.*, 2006; Gutiérrez *et al.*, 2007) and immobilised metal ions (Bornhorst and Falke, 2000), yet our observations revealed poor purification, with the protein predominantly located in the flow-through sample and wash fractions rather than the anticipated eluate. This issue might stem from the tag's potential inaccessibility within the folded protein structure hindering its binding to the immobilised metal. To address this, potential solutions included relocating the affinity tag from the C-terminus to the N-terminus of the protein or conducting the purification process under denaturing conditions (K Terpe, 2003a; Dutta and Bose, 2022). Although the N-terminal tagged protein showed low-levels of protein post purification, background contamination remained a notable challenge. This is of considerable concern particularly affecting low-expressing proteins, an inherent drawback of IMAC (Bornhorst and Falke, 2000) that may require additional optimisation or alternative purification strategies. To improve the isolation of our target proteins and purification efficiency while using the same resin (IMAC-Ni-TED), optimisation was done by introducing denaturing agents ad inhibitors. Our conditions included using dithiothreitol (DTT), Urea, and Triton-X 100. DTT acted as a reducing agent that stabilises proteins and was used within the limits in the manufacturers manual. Triton-X 100 minimised non-specific binding, thus increasing the possibility of our target proteins binding to the resin. Urea addition was incorporated to disrupt the protein structure thus aid in separating the target proteins from other components. Despite incorporation of the above-mentioned protein purification alterations, extraction with denaturing buffers yielded low levels of the target protein in both the crude and flow-through fractions Figure 4.9 and Figure 4.10, but these conditions did not notably improve protein purification. Consequently, further purification attempts were conducted using hydrophobic interaction chromatography coupled with ammonium precipitation (15%). SM07 demonstrated the most successful purification among the proteins tested. At the same time, SM03 showed poor purification, and some improvement was seen in SM29 (Figure 4.11), as revealed by western blot analysis.

Hydrophobic Interaction Chromatography (HIC) is a lengthy process. The N-terminal His_6-tagged lytic proteins were successfully expressed to minimise the purification steps, and purification was attempted with IMAC-NiTED. SM07 was partially purified successfully and subjected to a second purification step using ion exchange chromatography (IEC) to eliminate background protein contaminants in the final protein sample. IEC is a robust method for protein purification, separating proteins based on their net charge (Porath, 1992; Tripathi, 2016b; Cummins *et al.*, 2017). High protein purity is vital for therapeutic applications, allowing us to answer the call for discovery of alternative antimicrobial agents to combat multiple drug resistant bacterial strains. The findings indicate that anion exchange chromatography, particularly using CaptoQ resin, displayed better performance when compared to CaptoDEAE. These results highlight the efficiency of ion-exchange chromatography, particularly the potential advantages of using specific resins like CaptoQ for protein purification, paving the way for improved isolation and refinement of target proteins.

Phage-derived antibacterial proteins have great potential as antimicrobial agents because of their specificity and low-level toxicity in humans. However, the production of phage-derived proteins is challenging due to their low expression levels in both plant and bacterial host strains, observable from the data obtained in this study. Using IMAC followed by CaptoQ, unwanted background protein was eliminated; however, the concentration of the purified protein is still very low. The determination of the protein concentration was achieved using BSA standard quantification on an SDS-PAGE based on absolute values. Rather than using the colorimetric approach and that uses spectrophotometry. Using the spectrometry to ascertain the protein concentration might be higher in sensitivity, however, it is also broad and detects a range of proteins and not your target protein. SDS-PAGE quantification with BSA allows for more accurate protein concentration determination of target protein. SM07 purification confirmation was achieved through peptide mass fingerprint analysis, a valuable tool for protein identification and confirmation (Hamza *et al.*, 2021). The results revealed an endolysin with the three peptide stretches identified namely DLETVTLR, ALEILPAVR, and YAPYIGR. These confirmed that the purified SM07 is indeed an endolysin. However, the data demonstrated low confidence. This could be attributable to low level expression of the protein resulting in few peptides being extracted for identification. Peptides loss during protein fragmentation may also be of concern, however this could be improved with in-solution protein preparation. As low as the final peptide confidence results may be, the presence of SM07 was confirmed by western blot analysis using the His_6 tag for identification.

4.5 Conclusion

In conclusion, acquiring an effective endolysin as an antimicrobial agent is a promising avenue for combatting antibiotic-resistant bacteria. Transient expression was considered to be a favourable choice for endolysin production (using an alternate plant-based expression host system) due to the nature of the proteins. Nevertheless, when attempting transient expression of these proteins, our results showed that three proteins (SM07, SM03, and SM29) generated a truncated protein, with no SM31 protein observed even after optimisation of expression conditions. Having moved to a bacterial expression system with these three antimicrobial proteins, favourable expression was demonstrated especially in the case of SM07. The genes were successfully subcloned to incorporate His_6 tag at either ends of the genes under the T7 promotor of pET-30b(+) vector. Bacterial expression in BL-21(DE3) was poor, with most of the protein in the insoluble fraction. Challenges were encountered with purification of the C-terminal-tagged protein with IMAC; no improvements were observed in the purification even after introducing denaturing agents. Of the C-terminally tagged proteins, SM07 was purified successfully with HIC after 15% (w/v) ammonium sulfate precipitation. The proteins with the histidine tag on the N-terminal showed better results, a band corresponding with molecular mass for SM07 was observed under normal IMAC purification conditions. The same was not observed for SM03 and SM29. Following successful purification of SM07, a second purification "polishing" step was introduced using CaptoQ anion exchange. The final purified SM07 protein was confirmed via peptide mass fingerprinting analysis, with three peptide stretches identifying the protein as being an endolysin. By continuing to explore the potential of phage-derived proteins and developing effective purification techniques, we may discover new antimicrobial agents that can help address the growing threat of antibiotic resistance in the coming years.

4.6 References

Adrio, J.L. and Demain, A.L. (2014) Microbial enzymes: tools for biotechnological processes. Biomolecules 4: 117–39.

Bornhorst, J.A. and Falke, J.J. (2000) Purification of proteins using polyhistidine affinity tags. Methods Enzymol 326: 245–54.

Canto, T. (2016) Transient expression systems in plants: Potentialities and constraints. Adv Exp Med Biol 896: 287–301.

Coates, R.J., Young, M.T., and Scofield, S. (2022) Optimising expression and extraction of recombinant proteins in plants. Front Plant Sci 13: 1–15.

Cummins, P., Rochfort, K., and O'Connor Brendan (2017) Protein Chromatography. Protein Chromatography 1485: 423.

Dutta, S. and Bose, K. (2022) Protein Purification by Affinity Chromatography. Textbook on Cloning, Expression and Purification of Recombinant Proteins 141–171.

Fathi-Roudsari, M., Akhavian-Tehrani, A., and Maghsoudi, N. (2016) Comparison of Three Escherichia coli Strains in Recombinant Production of Reteplase. Avicenna J Med Biotechnol 8: 16–22.

Gutiérrez, R., Martín Del Valle, E.M., and Galán, M.A. (2007) Immobilized Metal-Ion Affinity Chromatography: Status and Trends. http://dx.doi.org/101080/15422110601166007 36: 71–111.

Hammond, R.W., Swift, S.M., Foster-Frey, J.A., Kovalskaya, N.Y., and Donovan, D.M. (2019) Optimized production of a biologically active Clostridium perfringens glycosyl hydrolase phage endolysin PlyCP41 in plants using virus-based systemic expression. BMC Biotechnol 19:.

Hamza, M., Ali, A., Khan, Suliman, Ahmed, S., Attique, Z., Ur Rehman, S., *et al*. (2021) nCOV-19 peptides mass fingerprinting identification, binding, and blocking of inhibitors flavonoids and anthraquinone of Moringa oleifera and hydroxychloroquine. J Biomol Struct Dyn 39: 4089–4099.

Hwang, H.-H., Yu, M., and Lai, E.-M. (2017) Agrobacterium -Mediated Plant Transformation: Biology and Applications . Arabidopsis Book 15: e0186.

Komarova, T. V, Baschieri, S., Donini, M., Marusic, C., Benvenuto, E., and Dorokhov, Y.L. (2010) Transient expression systems for plant-derived biopharmaceuticals.

Kovalskaya, N.Y., Herndon, E.E., Foster-Frey, J.A., Donovan, D.M., and Hammond, R.W. (2019) Antimicrobial activity of bacteriophage derived triple fusion protein against Staphylococcus aureus. AIMS Microbiol 5: 158–175.

Krenek, P., Samajova, O., Luptovciak, I., Doskocilova, A., Komis, G., and Samaj, J. (2015) Transient plant transformation mediated by Agrobacterium tumefaciens: Principles, methods and applications. Biotechnol Adv 33: 1024–1042.

Laemmli, U.K. (1970) Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. Nature 227: 680–685.

Mardanova, E.S., Blokhina, E.A., Tsybalova, L.M., Peyret, H., Lomonossoff, G.P., and Ravin, N. V (2017) Efficient Transient Expression of Recombinant Proteins in Plants by the Novel pEff Vector Based on the Genome of Potato Virus X. Front Plant Sci 8: 247.

Mauro, V.P. (2018) Codon Optimization in the Production of Recombinant Biotherapeutics: Potential Risks and Considerations. BioDrugs 2018 32:1 32: 69–81.

Merlin, M., Gecchele, E., Capaldi, S., Pezzotti, M., and Avesani, L. (2014) Comparative evaluation of recombinant protein production in different biofactories: the green perspective. Biomed Res Int 2014: 136419.

Mtimka, S., Pillay, P., Kwezi, L., Pooe, O.J., and Tsekoa, T.L. (2024) An Exploratory Review of the Potential of Lytic Proteins and Bacteriophages for the Treatment of Tuberculosis. Microorganisms 12:.

Munita, J.M. and Arias, C.A. (2016) Mechanisms of Antibiotic Resistance. Microbiol Spectr 4:.

Porath, J. (1992) Review: Immobilized Metal Ion Affinity Chromatography. Protein Expr Purif 3: 263–281.

C Reygaert, W. (2018) An overview of the antimicrobial resistance mechanisms of bacteria. AIMS Microbiol 4: 482–501.

Rosano, G.L. and Ceccarelli, E.A. (2014) Recombinant protein expression in Escherichia coli: advances and challenges. Front Microbiol 5: 172.

De Saeger, J., Park, J., Chung, H.S., Hernalsteens, J.P., Van Lijsebettens, M., Inzé, D., *et al*. (2021) Agrobacterium strains and strain improvement: Present and outlook. Biotechnol Adv 53:.

Seymour, S.L. and Hunter, C.L. (2015) ProteinPilotTM Software Overview High Quality, In-Depth Protein Identification and Protein Expression Analysis. Biomarkers and Omics 1.

Singh, A., Upadhyay, V., Upadhyay, A.K., Singh, S.M., and Panda, A.K. (2015) Protein recovery from inclusion bodies of Escherichia coli using mild solubilization process. Microb Cell Fact 14: 41.

Singh, A.A., Pillay, P., and Tsekoa, T.L. (2021) Engineering approaches in plant molecular farming for global health. Vaccines (Basel) 9:.

Starkevič, U., Bortesi, L., Virgailis, M., Ružauskas, M., Giritch, A., and Ražanskiene, A. (2015) High-yield production of a functional bacteriophage lysin with antipneumococcal activity using a plant virus-based expression system. J Biotechnol.

Terpe, K. (2006) Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. Appl Microbiol Biotechnol 72: 211–222.

Terpe, K. (2003) Overview of tag protein fusions: From molecular and biochemical fundamentals to commercial systems. Appl Microbiol Biotechnol 60: 523–533.

Tripathi, N.K. (2016) Production and Purification of Recombinant Proteins from Escherichia coli. ChemBioEng Reviews 3: 116–133.

Tsekoa, T.L., Singh, A.A., and Buthelezi, S.G. (2020) Molecular farming for therapies and vaccines in Africa. Curr Opin Biotechnol 61: 89–95.

Wong, J.W., Albright, R.L., and Wang, N.H.L. (2006) Immobilized Metal Ion Affinity Chromatography (IMAC) Chemistry and Bioseparation Applications. http://dx.doi.org/101080/03602549108021408 20: 49–106.

Xu, H., Chang, Q., Huang, L., Wei, P., Song, Y., Guo, Z., *et al*. (2023) An Agrobacterium-Mediated Transient Expression Method for Functional Assay of Genes Promoting Disease in Monocots. Int J Mol Sci 24: 7636.

5 CHAPTER FIVE: ANTIMICROBIAL ACTIVITY SCREENING OF SM07 AGAINST ESKAPE PATHOGENS AND DEMONSTRATING ITS APPLICATION

Abstract

The misuse of antibiotics and mutations in bacteria have led to the development of antibiotic-ineffectiveness and, subsequently, drug/multidrug-resistant bacterial strains. The global rise in antibiotic resistance has led to the growing interest in phages and their derivatives as antimicrobial alternatives to treat disease-causing bacteria. Phage-derived proteins, or endolysins, can hydrolyse peptidoglycan (PG), a cell wall component, resulting in cell lysis. An agar well diffusion assay was used to evaluate the efficacy of SM07 against ESKAPE pathogens for its antimicrobial properties. The application of SM07 was determined through a prototype spray-surface cleaner/disinfectant. Purified SM07 showed antimicrobial activity against *P. aeruginosa* without pre-treating the outer membrane (OM), at a minimum concentration of 0.004 µg/µl. Some activity was also observed against *Acinetobacter baumannii.* SM07 demonstrated a 30% toxicity effect against Vero cells at the highest concentration (0.0625 µg/µl) tested; a concentration that is almost 15 times more than MIC. The efficacy of the prototype spray developed with SM07 as an active ingredient (0.004µg/µl final concentration) showed a reduced bacterial count of approximately 98.70%. Our findings underscore the immense potential of endolysins as effective alternatives to conventional antibiotics with SM07 specifically treating *P. aeruginosa* infections. The results also demonstrate that endolysins can function effectively without having to pretreat the (OM) found in Gram-negative bacteria.

5.1 Introduction

The first antimicrobial to ever be used was to treat syphilis and was discovered by Paul Ehrlich in 1910 (Aminov, 2017). This was followed by the accidental discovery of penicillin by Alexandra Flemming in 1928 (Hare, 1983; Górski et al., 2019; Kumar et al., 2019). Following the discovery and official use of penicillin, new classes of antimicrobials were identified, with regular new developments (Aminov, 2017). The pace of discovery of new antibiotics has dropped significantly from the early 2000s until now, and infectious diseases have gradually grown resistance against the existing drugs (Luepke *et al.*, 2017). Antibiotic resistance is one of the most significant health threats worldwide, and the discovery and development of novel antimicrobial agents are crucial to combating this growing problem (Roca *et al.*, 2015; Miethke *et al.*, 2021).

Although this is a global health problem, its severity is worse in developing countries (Tadesse *et al.,* 2017; Ventola, 2015). The World Health Organisation (WHO) has identified a variety of disease-causing bacteria that demand immediate attention; particularly ESKAPE pathogens (*Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa*, and *Enterobacter* spp.) (Bhatia *et al.*, 2017; Mulani *et al.*, 2019). These pathogens account for various nosocomial infections, especially in infants, the elderly, and the immunosuppressed. The nosocomial infection rate in developing countries is almost double that in developed countries (Khan *et al.*, 2017; Abubakar *et al.*, 2022).

Based on WHO projections, multidrug-resistant infections deaths worldwide are expected to increase to ten million annually by 2050, with the danger of collapsing the global economy by 2050 (Criscuolo et al., 2017; Górski et al., 2019; Aslam et al., 2021). The WHO has recommended that attention should move towards advancing preexisting drugs and developing new antimicrobial agents to treat bacterial infections. Bacteriophages and lytic proteins have been proposed as potential antibacterial agents to tackle bacterial infections in humans, animals, and crops (Aslam et al., 2021b. Phages carry a range of essential proteins involved in their replication cycle, including receptor-binding proteins, VAPGHs, endolysins, and holins, which could be explored as a potential solution against antibiotic-resistant bacteria (Rodríguez-Rubio et al., 2013; Gutiérrez et al., 2018; Heselpoth et al., 2018; Witzany, 2020).

Characterisation of novel antimicrobials will also be vital for better understanding of their mode of action, efficacy, and safety. Understanding the mode of action can assist in optimising their use and also prevent resistance development. By characterising pre-existing, new, and improving existing ones, more could be developed further and design strategies to enhance their effectiveness and reduce their side effects. Unlike antibiotics, bacteriophage lysins

exclusively target specific bacteria and do not harm beneficial microflora, thus rendering them more advantageous over antibiotics (Aslam et al., 2021).

Phage-derived proteins (enzybiotics) have gained attention due to their nature and specificity (Fernandes and São-José, 2016; Torres-Barceló and Hochberg, 2016; Gerstmans et al., 2018). Their mode of action provides them with the advantage of little to no resistance when used. The aim of this chapter is to explore the potential antimicrobial activity of a novel protein (SM07) and to demonstrate its applicability for industrial use. Lysins have emerged as promising alternatives to conventional antibiotics due to their ability to target specific bacteria without harming beneficial microorganisms. Their ability to specifically target and disrupt bacterial cell walls makes them promising candidates for the development of new antimicrobial agents. In this chapter, we evaluate the efficacy of a newly discovered protein (SM07) against a wide range of pathogenic bacteria, including both Gram-positive and Gram-negative species. By demonstrating the effectiveness of the SM07, this chapter aims to establish lysins as a valuable tool for combating antimicrobial resistance and for industrial applications.

5.2 Methods and Material

The antimicrobial activity was performed under strict BioSafety Level 2 laboratory working regulations. The biological function of SM07 was tested against *Acinetobacter baumannii (*BAA747*),* ESCCO (ATCC 25922), *Enterococcus faecalis (*ATCC 51299*), Escherichia coli* (ATCC 35218*), Staphylococcus aureus* (ATCC 977BA) and *Pseudomonas aeruginosa (*ATCC 27853 and ATCC 15442).

5.2.1 Antimicrobial Susceptibility Test

The antibacterial activity of SM07 was performed using an agar well diffusion assay as Reeves (1989) described with slight modification using Mueller-Hinton (MH) agar. Each 100mm Petri dish was filled with ~20 ml of MH agar. Susceptibility studies were conducted against the ESKAPE pathogens. The Bauer-Kirby approach recommends the use of high concentrations of Bacto-sensitivity disks. Cefixime 30 µg was used as a control against *Acinetobacter baumannii* BAA747; Augmentin 30 µg against ESCCO ATCC 25922 *Enterococcus faecalis* ATCC 51299 and *Escherichia coli* ATCC 35218. While Cefotaxime 30 µg was used to inhibit both *Staphylococcus aureus* ATCC 977BA and *Pseudomonas aeruginosa* ATCC 27853. The recovered bacteria cultures were suspended in saline and adjusted to 0.5 McFarland. The suspensions were plated on MHA plates using sterile cotton swabs. The plates were inoculated in triplicate. A sterile borer created the wells on the plates and a volume of 90 µl of SM07 was added to the individual wells. Before incubation, each plate was numbered and a template for inserting the sterile disk and wells to allow identification and monitoring of the

progress. This was followed by an incubation step for 12 to 24 h at 37 °C. Following overnight incubation, plates were removed from the incubator and clearance zones around the well were observed.

5.2.1.1 Minimum Inhibitory Concentration (MIC)

The MIC was determined for SM07 using the same reference strain(s) according to the method proposed by Eloff (1998) with slight modification. Briefly, 150 µl of MH broth was pipetted into all the wells of the 96-well plate. Subsequently, 150 µl of each was added to the first well, followed by a two-fold serial dilution. Fifty microlitres (50 µl) of an overnight strain suspension prepared in MH broth and standardised to 0.5 MacFarhland was added to each well. The 96-well plate was then sealed and incubated for 24 h at 37 °C. After incubation, 20 µl of p-iodonitrotetrazolium chloride (0.2 mg/ml) was added to each well and incubated for 2 h under the same conditions. MICs were recorded as the lowest concentration that inhibited visible bacteria growth.

5.2.2 Cytotoxicity Studies of SM07 on Vero Kidney Cells

Vero cells were seeded at a density of 10,000 cells per well in a 96-well plate and incubated for 24 h under standard conditions (37 °C, 5% CO_2, 95% humidity). Following incubation, 2-fold serial dilutions of SM07, starting with a concentration of 0.0625 µg/µl, were prepared and 100 µl of the different concentrations were added to the corresponding wells of the 96-well plate containing vero cells, except in the cell control wells. The cells were then incubated with SM07 for 48 h. Subsequently, the media was replaced with 25 µl of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl-2H-tetrazolium bromide (MTT) reagent at 5 mg/ml (Sigma-Aldrich, St. Louis, MO). This was followed by another incubation of the cells at 37 °C for 3 h to facilitate formazan formation by viable cells. Subsequently, the MTT solution was discarded and 100 µl of dimethylsulfoxide (DMSO) (Sigma-Aldrich, St. Louis, MO) was added and incubated at room temperature for 15 min to dissolve the formazan product. The absorbance was then measured at 620 nm using a Tecan Infinite F500 luminometer (Tecan, Mannedorf, Switzerland).

5.2.3 Application of the Recombinant Produced SM07 as a Potential Surface Disinfectant

A surface disinfectant was formulated to determine the application of SM07. The ingredients in Table 5.1 were added to a beaker in the order shown and were mixed at 500 rpm using an overhead stirrer (IKA EUROSTAR 20 digital). Once the formulation was adequately mixed, the products were transferred to clean 250 ml bottles and packed at room temperature until testing. The bactericidal activity of the test product was evaluated using the membrane filtration method. The *P. aeruginosa* (ATCC 15442) was grown in nutrient broth overnight. The

culture was diluted to a concentration of 10^8 cfu/ml. An aliquot of the diluted culture was added to a membrane filter. The membrane filter was placed in a well containing the formulated surface disinfected with purified SM07 as the active ingredient at 3% of the final volume. The well was incubated at 20 °C for 5 min. The membrane filter was removed from the well and rinsed with the rinsing solution (Tryptone 1 g/l, NaCl 9 g/l and Tween80 5 g/l). The membrane filter was then placed on a plate of Tryptone Soy Agar and incubated at 37 °C for 24 h. The number of colonies that grew on the plates was used to determine the bactericidal activity of the test product.

Table 5.1: Prototype formulation of the disinfected surface.

	Prototype	**Control**
Water	91.95	94.95
Tween20	5	5
EDTA	0.05	0.05
Active Ingredient (SM07)	3 [0.12µg.µl]	0
Colourant (Blue)	q.s	q.s
Total Volume (ml)	100	100

5.3 Results

5.3.1 Microbial Screening of Recombinant Produced SM07

The antimicrobial activity of SM07 against the ESKAPE pathogens was tested (Table 5.2). Sensitivity was observed against *P. aeruginosa* with purified SM07, while some inhibition was also observed against *Acinetobacter baumannii*. SM07 was ineffective against Gram-positive bacteria such as *Staphylococcus aureus* and *Enterococcus faecalis*. However, the crude lysate did not show antimicrobial activity, which could be attributed to inhibition of the target protein by the presence of background proteins that are expressed in abundance or a very low concentration of the protein. Minimum inhibitory concentration studies were carried out against *P. aeruginosa* to determine the lowest concentration at which SM07 could inhibit growth which was found to be 0.004 µg/µl. Due to the inherently low initial antimicrobial susceptibility observed in the *Acinetobacter baumannii* strain, the conduct of minimum inhibitory concentration tests was deemed unnecessary.

5.3.2 Cytotoxicity Studies

SM07 was tested for cytotoxicity using the MTT assay against Vero kidney cells. Observations were made 48 h post exposure. At the highest concentration tested (0.0625 µg/µl), which is

almost 15 times the minimum inhibitory concentration, SM07 inhibits *P. aerigunosa* with no significant toxicity being observed, making the protein relatively safe even when used at high concentrations (Figure 5.1).

Table 5.2: Antibacterial activity screening on ESKAPE pathogens.

Strain	**Control Antibiotic**	**Test samples**	**Activity**	**MIC**
***Acinetobacter baumannii* BAA747**	Cefixime 30µg	Phosphate Buffer (PBS, negative control 1)	R'	ND
		Empty pET vector (negative control 2)	R'	ND
		Expression strain BL-21 (AI), (negative control 3)	R'	ND
		Clarified Crude Lysate	R	ND
		Purified SM07	Y	ND
ESCCO ATCC 25922	Augmentin 30µg	Phosphate Buffer (PBS, negative control 1)	R'	ND
		Empty pET vector (negative control 2)	R'	ND
		Expression strain BL-21 (AI), (negative control 3)	R'	ND
		Clarified Crude Lysate	R	ND
		Purified SM07	R	ND
***Staphylococcus aureus* ATCC 977BA**	Cefotaxime 30µg	Phosphate Buffer (PBS, negative control 1)	R'	ND
		Empty pET vector (negative control 2)	R'	ND
		Expression strain BL-21 (AI), (negative control 3)	R'	ND
		Clarified Crude Lysate	R	ND
		Purified SM07	R	ND
***Enterococcus faecalis* ATCC 51299**	Augmentin 30µg	Phosphate Buffer (PBS, negative control 1)	R'	ND
		Empty pET vector (negative control 2)	R'	ND
		Expression strain BL-21 (AI), (negative control 3)	R'	ND
		Clarified Crude Lysate	R	ND
		Purified SM07	R	ND
***Pseudomonas aeruginosa* ATCC 27853**	Cefotaxime 30µg	Phosphate Buffer (PBS, negative control 1)	R'	ND
		Empty pET vector (negative control 2)	R'	ND
		Expression strain BL-21 (AI), (negative control 3)	R'	ND
		Clarified Crude Lysate	R	ND
		Purified SM07	Y	0.004µg/µl
***Escherichia coli* ATCC 35218**	Augmentin 30µg	Phosphate Buffer (PBS, negative control 1)	R'	ND
		Empty pET vector (negative control 2)	R'	ND
		Expression strain BL-21 (AI), (negative control 3)	R'	ND
		Clarified Crude Lysate	R	ND
		Purified SM07	R	ND

R-Resistant

R'-Negative Control

ND- Not Determined

Y-Yes

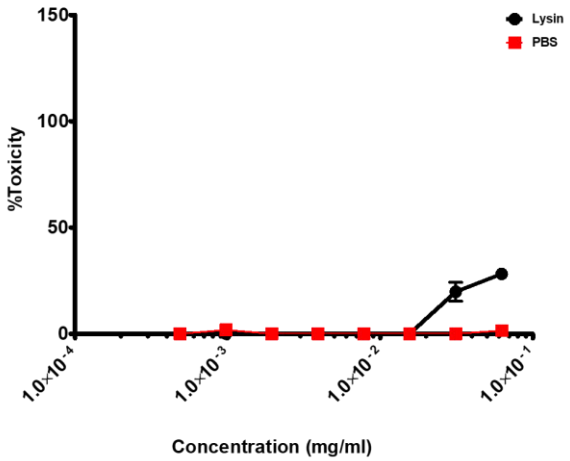


Figure 5.1: Purified SM07 toxicity evaluation. The data displayed are the means and standard deviations of three different experiments.

5.3.3 Application of the Recombinant Produced SM07 as a Potential Surface Disinfectant

The bactericidal efficacy of SM07 in a formulation was tested against *P. aeruginosa* (ATCC 15442) under dirty conditions. The product was applied in its neat form and assessed after 5-minute contact time including an interfering substance (3.0 g/l of bovine albumin) replicating dirty conditions. The initial bacterial count decreased to 5.57 (Log N_a) from 7.44 (Log N_0), indicating a reduced bacterial count of 98,70%. The concentration of SM07 in the formulation was 0.004 µg/µl. Detailed validation test observations, which show the experimental conditions, dilutions, and log reduction values against the *P. aeruginosa* strain with the purified SM07, can be seen in Table A.3.

5.4 Discussion

Bacteriophages have often been used as antimicrobial agents; however, reports of phage-derived proteins used as therapeutic agents only appeared in 2001 by Fischetti *et al.* where efficacy was demonstrated in mice. Phages and their derivatives have potential uses in several sectors (Clark and March, 2006; Vázquez *et al.*, 2018). SM07 was tested for its antimicrobial activity against ESKAPE pathogens and observed inhibition of growth with *P. aeruginosa*. No activity was observed from any of the Gram-positive pathogens, and this is a based on the specificity of SM07, the CBD has specificity for Gram-negative bacteria, hence the activity against *P. aeruginosa*. The inability to be functionally against Gram-bacteria might also be the additional barrier from the cell wall component of the bacteria like Teichoic acid which is not found in Gram-negative bacteria. The inhibition by SM07 against an organism known to cause

infections in the blood, lungs, urinary tract, and gastrointestinal tract (Yamada *et al.*, 2009) is promising. Fischetti (2018) suggests that endolysins are active against Gram-positive bacteria and less active against Gram-negative bacteria because of the presence of the OM. The OM makes it difficult for endolysins to reach the PG layer. SM07 did not only show efficacy against *P. aeruginosa,* but some inhibition was observed against *Acinetobacter baumannii.* It is worth noting that inhibition in both was observed in the absence of outer membrane destabilizers. This phenomenon has been observed in other studies (Guo *et al.*, 2017) and this, according to Park et al., (2014) and Murray et al., (2021), is because Gram-negative infecting endolysins have a broader range of targets.

One of the limitations of endolysins against Gram-negative bacteria is their inability to penetrate the OM (Yang *et al.*, 2014). However, by destabilising OM, thus exposing the PG, endolysins can be effective (Vázquez *et al.*, 2018). Surprisingly, SM07 was able to inhibit a Gram-negative bacterium without OM destabilizers as an indication of its potency. It exhibited activity at a low concentration of 0.004 µg/µl. The efficacy can potentially be enhanced by pretreatment of the OM with agents such as Triton-X or EDTA prior to use (Guo *et al.*, 2017; Vázquez *et al.*, 2018).

Cytotoxicity studies were conducted in Vero kidney cells as a representative for general cytotoxicity to mammalian cells (Perryman *et al.*, 2018). SM07 exhibited a 30% toxicity towards the cells, at the highest tested concentration (0.0625 µg/µl) which is almost 15 times the MIC for *P. aerigunosa* inhibition. Having observed antimicrobial activity and bacterial reduction of 98,70% at a concentration 0.004 µg/µl and no observed toxicity at that level in Vero cells, SM07 demonstrates potential for application towards combatting the global rise of antibiotic resistance specifically against *P. aerigunosa* using proteins and with no cytotoxic effects in normal mammalian cells.

SM07 can be used for its antimicrobial function in different ways and a prototype spray was developed. The formulation contained a concentration of 3% (v/v) SM07 (equivalent to our determined MIC) and exhibited a log reduction of 1.87 of *P. aeruginsa* (ATCC 15442) when applied at its neat concentration. A noteworthy reduction in bactericidal activity (98.70%) was achieved with the prototype. These findings show promise; it is crucial to note that for a substance to be regarded a high-level disinfectant, it typically needs to achieve a log reduction of at least 5, which is equivalent to a bacterial count reduction of 99,99%. The bactericidal activity displayed by SM07 under the tested conditions has potential in settings where complete sterilisation is not a prerequisite or in combination with other agents.

The discovery SM07 with an effective ability to inhibiting *Pseudomonas aeruginosa* at low concentrations with no harm to mammalian cells holds significant promise for future

therapeutics. It offers a potential solution to the growing threat of antibiotic-resistant *P. aeruginosa* infections. Its potency at low doses suggests reduced treatment costs and side effects, while its specificity minimizes damage to healthy gut bacteria. Furthermore, studying SM07 further can provide crucial insights into the cell wall vulnerabilities of *P. aeruginosa*, paving the way for the development of even more targeted and effective anti-*P. aeruginosa* therapies.

5.5 Conclusion

In conclusion, antimicrobial studies conducted on SM07 have demonstrated its remarkable ability to inhibit the growth of *Pseudomonas aeruginosa*, a pathogen known for its resistance to multiple antibiotics and for causing severe respiratory tract infections. These findings highlight the tremendous potential SM07 holds as a promising candidate for treating *P. aeruginosa* infections. Its ability to target and disrupt the specific mechanisms used by *P. aeruginosa* without having to destabilize the OM makes it an attractive alternative to conventional antibiotics. In addition, these proteins are known for their specificity and efficacy against bacteria, further strengthening their potential as therapeutic agents. SM07 showed 30% cytotoxicity towards the cells at the highest concentration tested which is 15 times more than the MIC. The application test as a surface disinfectant/cleaner containing SM07 exhibited significant bactericidal activity, achieving a log 1.87 reduction in *Pseudomonas aeruginosa* ATCC 15442. These findings underscore its effectiveness as a powerful disinfectant in combating bacterial contamination. Opportunities for further optimising of the formula and application conditions present avenues for future research. Further research and development efforts should be directed toward exploring the stability of the polished protein and further characterizing of it. Overall, the results of this study provide a strong foundation for the future development and utilisation of this phage-derived protein as a novel antimicrobial agent for use against *P. aeruginosa* infections. Moreover, it also affects some other Gram-negative bacteria, such as *Acinetobacter baumannii,* making it a promising candidate for Gram-negative bacteria as well.

5.6 References

Abubakar, U., Amir, O., and Rodríguez-Baño, J. (2022) Healthcare-associated infections in Africa: a systematic review and meta-analysis of point prevalence studies. J Pharm Policy Pract 15: 1–16.

Aminov, R. (2017) History of antimicrobial drug discovery: Major classes and health impact. Biochem Pharmacol 133: 4–19.

Aslam, B., Arshad, M.I., Aslam, M.A., Muzammil, S., Siddique, A.B., Yasmeen, N., *et al*. (2021) Bacteriophage Proteome: Insights and Potentials of an Alternate to Antibiotics. Infect Dis Ther.

Bhatia, P., Sharma, A., George, A.J., Anvitha, D., Kumar, P., Prakash Dwivedi, V., and Chandra, N.S. (2017) Antibacterial activity of medicinal plants against ESKAPE: An update.

Clark, J.R. and March, J.B. (2006) Bacteriophages and biotechnology: vaccines, gene therapy and antibacterials. Trends Biotechnol 24: 212–218.

Criscuolo, E., Spadini, S., Lamanna, J., Ferro, M., and Burioni, R. (2017) Bacteriophages and Their Immunological Applications against Infectious Threats. J Immunol Res 2017:.

Ehrlich Paul (1910) Nerve disorders and Salvarsan treatment.

Eloff, J.N. (1998) A sensitive and quick microplate method to determine the minimal inhibitory concentration of plant extracts for bacteria. Planta Med 64: 711–713.

Fernandes, S. and São-José, C. (2016) More than a hole: the holin lethal function may be required to fully sensitize bacteria to the lytic action of canonical endolysins. Mol Microbiol 102: 92–106.

Fischetti, V.A. (2018) Development of phage lysins as novel therapeutics: A historical perspective. Viruses 10:.

Gerstmans, H., Criel, B., and Briers, Y. (2018) Synthetic biology of modular endolysins. Biotechnol Adv 36: 624–640.

Górski, A., Miedzybrodzki, R., and Borysowski, J. (2019) Phage therapy: A practical approach, Górski, A., Międzybrodzki, R., and Borysowski, J. (eds) Springer.

Guo, M., Feng, C., Ren, J., Zhuang, X., Zhang, Y., Zhu, Y., *et al*. (2017) A novel antimicrobial endolysin, LysPA26, against *Pseudomonas aeruginosa.* Front Microbiol 8: 1–9.

Gutiérrez, D., Fernández, L., Rodríguez, A., and García, P. (2018) Are phage lytic proteins the secret weapon to kill staphylococcus aureus? mBio 9:.

Hare, R. (1983) The scientific activities of alexander fleming, other than the discovery of penicillin. Med Hist 27: 347–372.

Heselpoth, R.D., Swift, S.M., Linden, S.B., Mitchell, M.S., and Nelson, D.C. (2018) Enzybiotics: Endolysins and Bacteriocins.

Khan, H.A., Baig, F.K., and Mehboob, R. (2017) Nosocomial infections: Epidemiology, prevention, control and surveillance. Asian Pac J Trop Biomed 7: 478–482.

Kumar, M., Jaiswal, S., Sodhi, K.K., Shree, P., Singh, D.K., Agrawal, P.K., and Shukla, P. (2019) Antibiotics bioremediation: Perspectives on its ecotoxicity and resistance. Environ Int 124: 448–461.

Luepke, K.H., Suda, K.J., Boucher, H., Russo, R.L., Bonney, M.W., Hunt, T.D., and Mohr, J.F. (2017) Past, Present, and Future of Antibacterial Economics: Increasing Bacterial Resistance, Limited Antibiotic Pipeline, and Societal Implications. Pharmacotherapy 37: 71–84.

Miethke, M., Pieroni, M., Weber, T., Brönstrup, M., Hammann, P., Halby, L., *et al*. (2021) Towards the sustainable discovery and development of new antibiotics. Nature Reviews Chemistry 2021 5:10 5: 726–749.

Mulani, M.S., Kamble, E.E., Kumkar, S.N., Tawre, M.S., and Pardesi, K.R. (2019) Emerging Strategies to Combat ESKAPE Pathogens in the Era of Antimicrobial Resistance: A Review. Front Microbiol 10: 539.

Murray, E., Draper, L.A., Ross, R.P., and Hill, C. (2021) The Advantages and Challenges of Using Endolysins in a Clinical Setting. Viruses 13: 1–22.

Park, Y., Lim, J.A., Kong, M., Ryu, S., and Rhee, S. (2014) Structure of bacteriophage SPN1S endolysin reveals an unusual two-module fold for the peptidoglycan lytic and binding activity. Mol Microbiol 92: 316–325.

Perryman, A.L., Patel, J.S., Russo, R., Singleton, E., Connell, N., Ekins, S., and Freundlich, J.S. (2018) Naïve Bayesian Models for Vero Cell Cytotoxicity. Pharm Res 35: 170.

Roca, I., Akova, M., Baquero, F., Carlet, J., Cavaleri, M., Coenen, S., *et al*. (2015) The global threat of antimicrobial resistance: science for intervention. New Microbes New Infect 6: 22–29.

Rodríguez-Rubio, L., Martínez, B., Donovan, D.M., Rodríguez, A., and García, P. (2013) Bacteriophage virion-associated peptidoglycan hydrolases: Potential new enzybiotics. Crit Rev Microbiol 39: 427–434.

Torres-Barceló, C. and Hochberg, M.E. (2016) Evolutionary Rationale for Phages as Complements of Antibiotics. Trends Microbiol 24: 249–256.

Vázquez, R., García, E., and García, P. (2018) Phage lysins for fighting bacterial respiratory infections: A new generation of antimicrobials. Front Immunol 9: 2252.

Witzany, G. (2020) Biocommunication of phages, Witzany, G. (ed) Springer.

Yamada, N., Nishida, T., and Chikama, T.I. (2009) Pseudomonas aeruginosa Infections. Japanese Journal of Clinical Ophthalmology 63: 305–326.

Yang, H., Yu, J., and Wei, H. (2014) Engineered bacteriophage lysins as novel anti-infectives. Front Microbiol 5: 542.

6 CHAPTER SIX: *IN SILICO* MODELLING OF SM07 TO ELUCIDATE ITS FUNCTION

Abstract

Recent advances in computational approaches and structural biology techniques have provided valuable tools for predicting and elucidating various molecular structures, including endolysins. Endolysins are enzymes that hydrolyse a component of bacterial cell walls. Thus, they are considered to be a potential therapeutic tool for treating bacterial infections. Understanding the structure-function relationships of endolysins is critical to facilitate their optimisation as antimicrobial agents. To build a model for SM07, four tools were used, based on homology, protein folding, and *ab initio* modelling. AlphaFold produced the most reliable and accurate 3D structure model. The verification of the model was carried out using various structural analysis tools, Ramachandran plots, ERRAT, Verify-3D, ProCheck, G-factor, and Z scores of ProSA and AlphaFold outperformed each other, with Modeller showing the poorest structure and accuracy. The predicted 3D structure of SM07 revealed a cell-wall binding domain (CBD) and a catalytic/enzymatic domain that possesses a typical catalytic dyad (Glu228 and Glu237). An interesting observation was the discovery of a CBD; it is unusual for Gram-negative infecting endolysins. These findings provide valuable insight into the structure-function relationship. The unusual presence of a putative CBD warrants further research to elucidate the underlying mechanism and potential implications for antimicrobial therapies.

6.1 Introduction

Despite the use of antibiotics worldwide, antibiotic resistance continues to threaten the health system. Attention has been focused on phage-derived proteins due to their ability to hydrolyse the peptidoglycan from within cell components (Lin et al., 2017; Fischetti, 2018; Cernooka et al., 2022). These proteins typically contain cell wall-binding domains and/or a catalytic domain (Low *et al.*, 2005; Cernooka *et al.*, 2022). These domains can be found at the protein's N- or C- terminal based on the type of bacteria they target. Gram-negative targeting phage-derived proteins typically do not have a CBD but have an N-terminal transmembrane domain (Cernooka *et al.*, 2022). Tools to determine protein structure is critical as they provide valuable insights into function and behaviour (Haas *et al.*, 2013).

There are two main approaches used to determine the structure of a protein. One of the approaches is the experimental method, which uses physical techniques such as X-ray crystallography, NMR spectroscopy, and cryoelectron microscopy to determine the three-dimensional structure of a protein and is more accurate than the computational method (Baker and Sali, 2001; Masrati *et al.*, 2021). The second approach is the computational method which predicts the structure of a protein based on its amino acid sequence (Kortemme and Baker, 2004; Farhadi and Hashemian, 2018). While experimental methods are essential to obtain accurate and high-resolution protein structures, advances in computational modelling have made it an attractive alternative due to its speed, cost effectiveness, and broad applicability (Masrati *et al.*, 2021). These approaches can be used separately or coupled. Throughout the years, protein structure modelling has significantly advanced biomedical research and drug development.

In the early 1990s, the prediction of the HIV-1 protease structure led to the design of drugs that inhibit the protease, effectively treating HIV/AIDS. Subsequently, protein structure modelling was used to engineer a more heat-resistant version of glucose-6-phosphate dehydrogenase, which is beneficial in treating the genetic disorder G6PD deficiency. In the early 2000s, this approach provided insights into the interaction between the proteins prion and tau, guiding the development of drugs to target these interactions for treating Alzheimer's and prion diseases. In the late 2000s, protein structure modelling enabled the design of a novel protein, the albumin-binding domain, capable of delivering drugs to the brain by attaching to the abundant blood protein, albumin.

Predicting protein structures poses a challenge, yet the significance of identifying protein structure continues to grow, particularly within the realm of biomedical research. The ability to predict protein structures has helped many understand the functionality of proteins,

interactions, and how some proteins are involved in diseases. This chapter will be dedicated to exploring structural modelling. Structural modelling is vital to predicting and understanding the three-dimensional arrangements of proteins and other molecules. It is also as important to determine the stability of a protein's 3D structure is crucial to ensure its functionality and reliability in biological processes. Unstable structures can lead to misfolding, which is often associated with diseases and loss of protein activity. The aim is to predict the overall structure of SM07 from its amino acid sequence. Throughout this chapter, we will examine the different levels of protein structure and decipher the arrangements of the protein.

This chapter presents an overview of the structural modelling techniques used to explore SM07 to elucidate its function. We will discuss the application of computational tools, including AlphaFold, homology modelling, and other structure analysis methods, to predict and validate the 3D structures. Moreover, we will highlight the importance of structural insights in identifying catalytic residues and catalytic domains in elucidating the function of SM07.

6.2 Methodology

6.2.1 Sequence Retrieval, Alignment and Validation

The protein sequence for SM07 was aligned with homologous crystallised proteins in the Protein Data Bank (PDB) database for the building of 3D structures. Subsequently, four different computational modelling tools, namely Modeller (Eswar *et al.*, 2008), I-TASSER (Yang and Zhang, 2015), SWISSMODEL (Bordoli *et al.*, 2009) and AlphaFold (David *et al.*, 2022a) were used to generate predicted protein models. The quality and accuracy of each predicted structure was validated using SAVESv6.0 (available at https://saves.mbi.ucla.edu/). ProCheck (Laskowski *et al.*, 1996) was used to validate the best and most reliable model through the generated Ramachandran plot (Ramachandran *et al.*, 1963). Further validation was performed through Verify 3D (Eisenberg *et al.*, 1997), ERRAT (Colovos and Yeates, 1993), and Protein Structure Analysis (ProSA) (Wiederstein and Sippl, 2007) to determine the overall quality of the protein as to well as determine the problematic regions of the predicted models.

6.2.2 Analysis of protein structure and molecular graphics

3D structure visualisation and manipulation were carried out using the molecular visualization tool PyMol (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC). The identification of putative domains was achieved through the utilisation of InterPro, a valuable resource for protein sequence analysis (https://www.ebi.ac.uk/interpro/). For the identification of the catalytic residues, the GraFSS tool was employed and was accessed through mfrlab.org. Additionally, to recognise distant homologues through sequence-structure

comparison, the FUGUE programme was utilized (Miguel *et al.*, 2002). The output was further enhanced by coupling it with BioEdit (Hall, 2011), which provided graphic visualisation of the results.

6.3 Results

6.3.1 Building SM07 Model and Evaluation

A BLAST query search for SM07 sequence using the Protein Database revealed four templates to model our structure. Those templates are 4OK7, 2CJL, 5H7T, and 4MCK, and their descriptions include total score, sequence similarity, query cover, and E-values can be seen in Table 6.1. Four tools were used to build the 3D structure of our protein. Modeller generated five models and the model evaluation was done by observing the discrete optimized protein energy (DOPE) and genetic algorithm 341 (GA341) score. The best model produced 0.68327 and -31256.20313 scores, respectively. Alphafold predicted five models with the best model with sequence identity and E-value of 62.6% and 1.1 x $10^{-74,}$ respectively. Swiss-Model used a template from AlphaFold with UniProtKB ID A0A1A0XGG3.1.A with a sequence identity of 70.72%, the model had a GMQE score of 0.85. I-Tasser generated ten templates to predict the most accurate model, producing C-, TM-, and RMSD scores of -1.33, 0.55, and 10.2 Å, respectively (Figure 6.1).

Table 6.1: Templates IDs and analysis of protein produced from NCBI-PDB BLAST against SM07.

Template IDs	**Protein Name**	**Total score**	**Sequence similarity (%)**	**Query cover (%)**	**E-value**
4OK7	SPN1S Endolysin from *Salmonella typhimurium*	47.4	25.24	34	2 x 10^{5}
2CJL	GH19 Chitinase	40.4	27.27	32	0.003
5H7T	GH19 chitinase domain	38.9	25.83	29	0.009
4MCK	GH19, Class IV Chitinase	38.5	30.77	14	0.012

Rank	PDB Hit	Iden1	Iden2	Cov	Norm. Z-score	Download Align.
1	4ok7A	0.32	0.15	0.35	2.29	Download
2	7vq1A	0.08	0.32	0.93	0.69	Download
3	4ok7	0.20	0.15	0.37	2.95	Download
4	6lnr	0.17	0.16	0.53	2.50	Download
5	5voxB	0.06	0.22	1.00	1.04	Download
6	4ok7	0.32	0.15	0.33	3.89	Download
7	4ok7A	0.31	0.15	0.37	2.02	Download
8	1wvuB	0.19	0.17	0.49	0.74	Download
9	7sz4K	0.07	0.21	0.89	1.33	Download
10	3hbdA	0.16	0.12	0.40	2.58	Download

Figure 6.1: I-TASSER templated with their scores used for the homology modelling of SM07.

6.3.2 Validation of Predicted Models

To obtain a reasonably accurate 3D structure for SM07, the validation and quality assessment of each model predicted by the four homology modelling techniques were carried out using ProCheck. Table 6.2 displays the corresponding data, highlighting the performance of each approach in the analysis. The validation of the models predicted by the four tools was evaluated using four evaluation metrics. Those metrics included ERRAT, Verify-3D, G-factor, and Z-score. The purpose of using these metrics was to access the accuracy and reliability of the models. As reflected in Table 6.3 it is evident that the model predicted by AlphaFold demonstrated the most reliable model with an overall quality of 94.86%. Additionally, 84.68% of its residues averaged a 3D-1D score of 0.1, indicating a pass. Despite it having four errors as seen in Table 6.3, it remained impressive with three passes in the ProCheck validation while others had zero passes or just one or two.

Table 6.2: ProCheck analysis for the models predicted by AlphaFold, Swiss-Model, Modeller, and I-Tasser.

Regions	***AlphaFold***	***SwissModel***	***I-Tasser***	***Modeller***
Residues in most favoured regions	357	345	269	306
Residues in additional allowed regions	27	30	100	63
Residues in generously allowed regions	4	2	14	10
Residues in disallowed regions	2	3	7	11
Number of non-glycine and non-proline	390	380	39	390
Number of end-residues	2	2	2	2
Number of glycine residues	44	42	44	44
Number of proline residues	21	21	21	21
Total coverage	457	445	457	457

Table 6.3 ERRAT, Verify3D, G- factor, and Z-scores of the predicted models for SM07.

	ERRAT (overall quality)	Verify3D	ProCheck	Overall G-factor	Z-scores from ProSA
Alphafold	94.8661	84.68% of the residues have averaged 3D-1D score >= 0.1 Pass	Errors = 4 Warning = 2 Pass = 3	-0.16	-9.88
Swiss model	93.9675	82.47% of the residues have averaged 3D-1D score >= 0.1) Pass	Error = 4 Warning = 3 Pass = 2	- 0.08	-9.62
I-TASSER	86.1048	65.86% of the residues have averaged 3D-1D score >= 0.1 Fail	Error = 6 Warning = 3 Pass = 0	- 0.66	-5.77
Modeller	13.5714	56.89% of the residues have averaged 3D-1D score >= 0.1) Fail	Error = 4 Warning = 3 Pass = 1	- 0.12	0.75

6.3.3 Overall structure of SM07

The structure of SM07 shows the N-terminal arm followed by a short beta sheet. The domain following the N-terminal starting with a coil comprises 22 alpha helices and eight beta sheets. The protein is a monomer (Figure 6.2). The generated structure appeared as a monomeric structure and a helical protein with two domains, the large domain at 181-344 and the small domain at 10-50. The Ramachandran plot for the best and most reliable model is presented in Figure 6.3. The plot illustrates the distribution of the psi and phi dihedral angles, confirming the high-quality and favourable conformational properties of the structure.

The N-terminal domain of residue 10-50 comprises of a putative CBD based on sequence homology with other enzymes. Figure 6.4 illustrates the domain highlighted from the 3D structure of SM07 in the previous figure. The C-terminal domain harbours the catalytic/enzymatic domain from residues 183-314 (Figure 6.4B). Figure 6.4C shows the catalytic residues within the catalytic domain. The catalytic resides Glu228 and Glu237 within our sequence form part of a putative dyad within a triad. The third residue Asn260 is located outside the cleft.

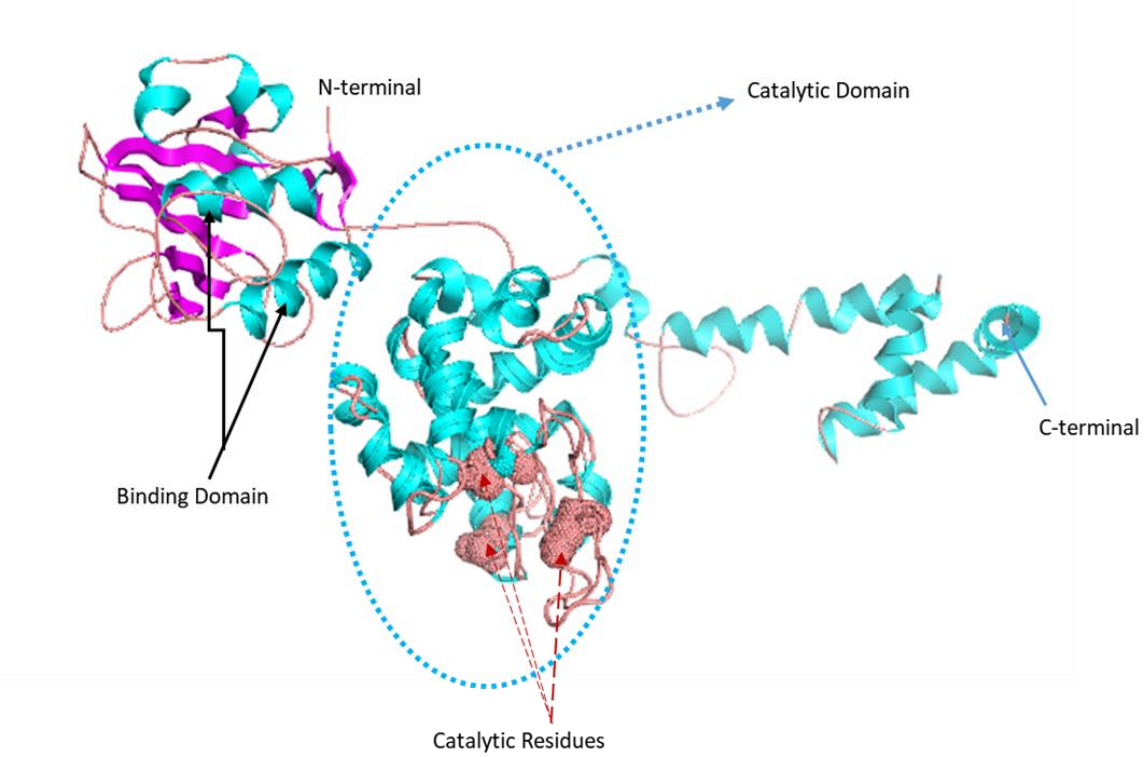
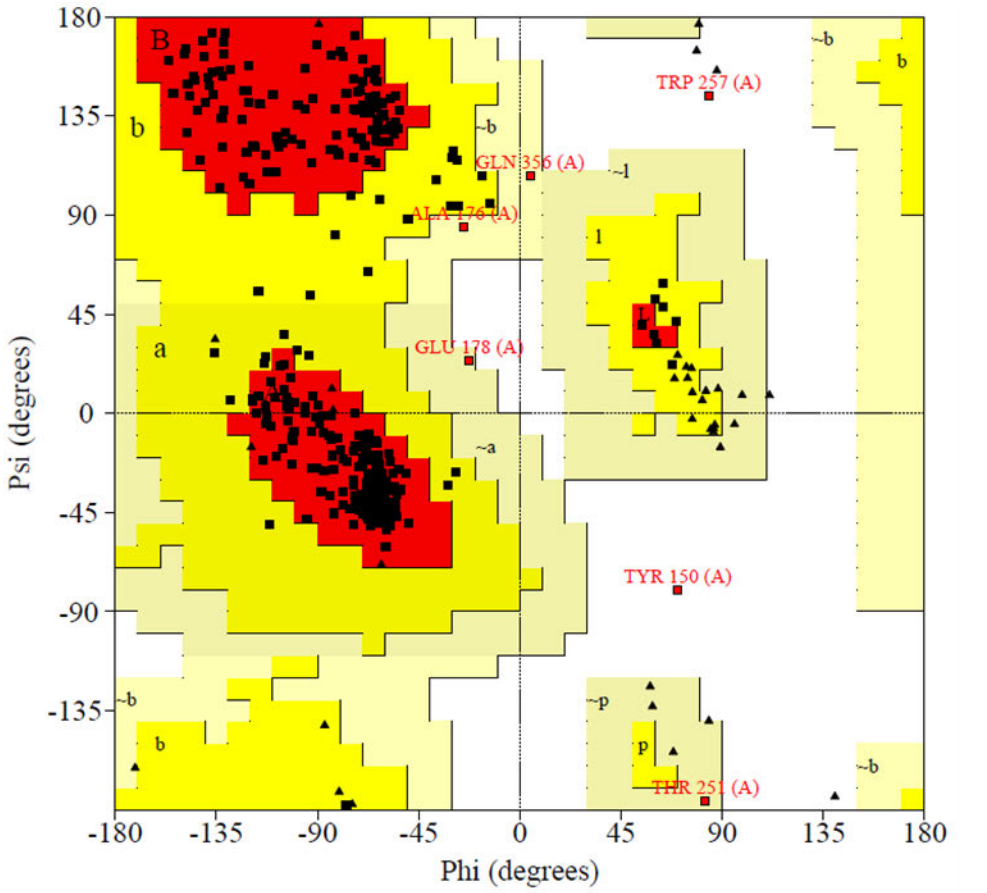


Figure 6.2: Cartoon 3D structure of SM07 illustrating its arrangement and molecular architecture.



Plot statistics

Residues in most favoured regions [A,B,L]	357	91.5%
Residues in additional allowed regions [a,b,l,p]	27	6.9%
Residues in generously allowed regions [~a,~b,~l,~p]	4	1.0%
Residues in disallowed regions	2	0.5%
Number of non-glycine and non-proline residues	390	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	44	
Number of proline residues	21	
Total number of residues	457	

Figure 6.3: Ramachandran plot of SM07 from ProCheck output.

.

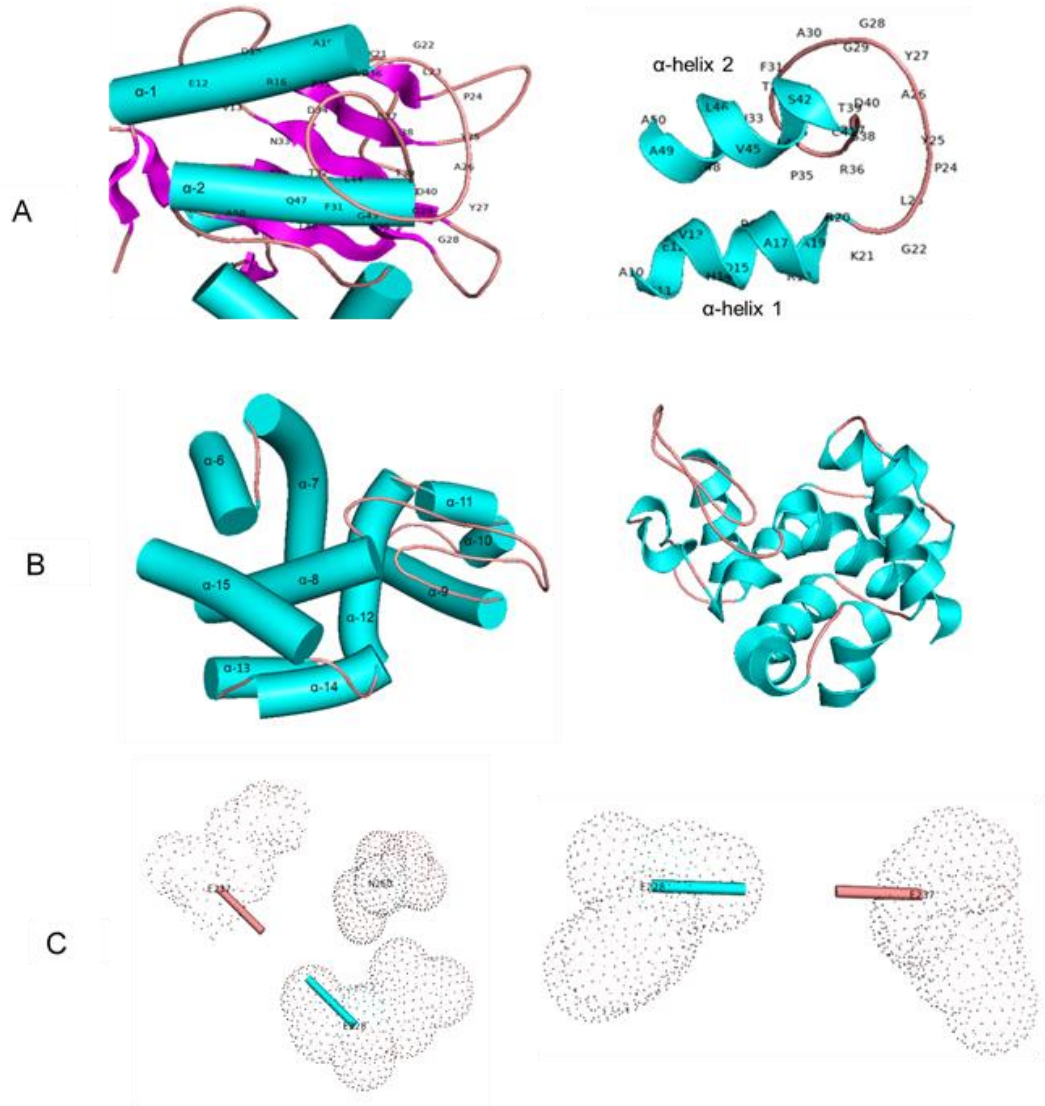


Figure 6.4: Highlighted domains from the 3D structure of SM07.
A) CBD, (B) Catalytic/Enzymatic Active Domain and (C) Putative catalytic residues are presented in blue for Glu228 and Pink for Glu237.

Using the 3D protein structure sequence output from the best model (AlphaFold), secondary structure analysis identified the regions of the protein corresponding to alpha-helices, beta sheets, and coil regions. Table 6.4 shows the detailed outcomes of the analysis. Through bioinformatics analysis, two putative functional domains were identified within SM07s sequence: a CBD, which is responsible for recognizing specific components of the bacterial cell wall components, and a catalytic domain, which is essential for enzyme activity and subsequent bacterial lysis. Figure 6.5 shows the alignment of the sequence with a PDB template that exhibited the best homology to our protein and highlights putative domains and binding residues.

Table 6.4: Detailed description of SM07 topology.

SM07			
Secondary Structure	Residues	**Secondary Structure**	Residues
β sheet	3-5	**α helix**	215-229
α helix	7-20	**α helix**	258-271
α helix	41-53	**α helix**	279-282
α helix	66-71	**α helix**	284-288
α helix	74-79	**α helix**	290-304
β sheet	81-83	**α helix**	309-314
α helix	89-92	**α helix**	317-325
β sheet	98-103	**α helix**	331-344
β sheet	112-117	**α helix**	346-352
β sheet	134-138	**α helix**	385-395
β sheet	147-149	**α helix**	397-408
β sheet	162-167	**α helix**	411-422
β sheet	169-172	**α helix**	431-447
α helix	181-189	**α helix**	449-456
α helix	193-210		

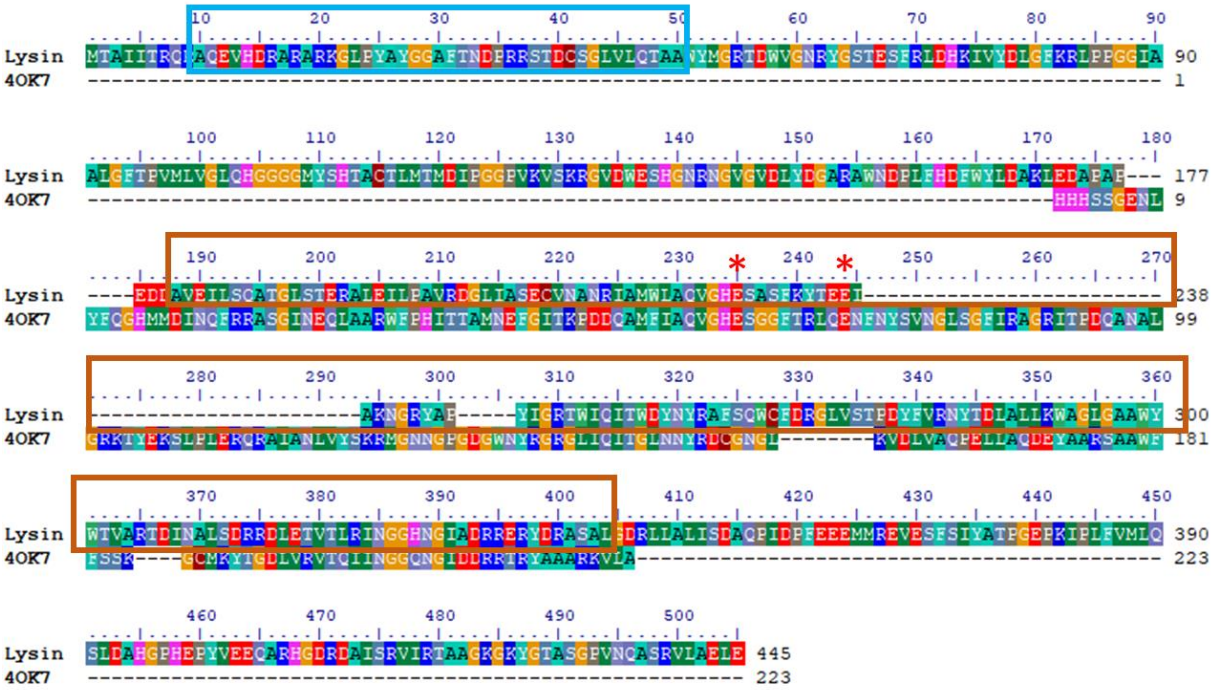


Figure 6.5: Modular composition of the phage-derived protein (SM07).

Full-length sequence of SM07 aligned with amino acid sequence of SPN1S endolysins (4OK7) with the catalytic domains highlighted with red asterisks and putative N-terminal CBD (Blue) and C-terminal catalytic domain (Orange).

6.4 Discussion

The 3D structure of a protein provides critical information needed to better understand the structure/function relationship and helps to identify catalytic binding regions (Deng *et al.*, 2018). In search of reliable sequence homologs and thorough structural characterization of SM07. SM07 was subjected to comparative analysis against existing protein crystal structures located in the PDB. The basic principle of homology modelling is the selection of a template and sequence alignment between the target and the template (Fiser, 2010). The PDB search revealed that SM07 was more like an endolysin from *Salmonella typhimurium* (PDB:4OK7) with sequence identity of 25.24% from query cover of 34%, followed by GH19 chitinases (PDB: 2CJL, 5H7T and 4MCK) (Table 6.1). The sequence identity between the query and template structures influences the quality of a model built using a homology modelling technique (Šali and Blundell, 1993). Four tools (AlphaFold, Swiss-Model, Modeller, and I-Tasser) were used for comparative modelling of SM07. These tools are based on sequence homology evaluation, fold recognition, and ab initio (Xu et al., 2005; Yang and Zhang, 2015) with all of these producing high-quality models when the sequence identity between the query and the template is greater than 30% (Kryshtafovych and Fidelis, 2009). However, for low sequence identities, it becomes difficult for the modelling algorithms to produce high-quality structure predictions. Therefore, different modelling techniques must be used to build protein structures with relatively low sequence identities to their templates.

Template 4OK7 from PDB was selected as a basis for modelling SM07, based on the E-value, sequence identity, and coverage outperforming the others despite its lower similarity. With that, AlphaFold exhibited the highest validity structure with a 94.86% overall quality, followed by Swiss Model at 93.96%, I-TASSER at 86.10% and with Modeller's structure prediction quality being the lowest (13.57%). AlphaFold produces more accurate models than its counterparts because it uses homology modelling and multiple sequence alignments to generate 3D structures and is a deep learning-based approach (David *et al.*, 2022b). The Swiss-Model model had a high score of 0.85, indicating accuracy. The accuracy of Swiss-Model can be attributed to the software integration of AlphaFold into its system. A significant observation from the I-Tasser output was the 7sz4K which highlights its origin as a *Pseudomonas* phage. One other critical observation made based on the homology search using the PDB output was the confirmation of the primary function of the protein as a peptidoglycan-targeting hydrolase with antimicrobial activity against Gram-negative bacteria. Having SM07 as a potential candidate to combat one of the six pathogens known for nosocomial infections is a good find as endolysins are known to be highly specific to genus species (Fischetti, 2008; Fenton et al., 2010; Nelson et al., 2012).

The AlphaFold-generated model was selected as the model for SM07. The model further demonstrated its quality through Verify3D and G-factor scores, and its nativeness is affirmed by SA's Z-score. Critical domains essential for the functionality of an endolysin were identified. The catalytic domain (enzymatic activity domain) and the binding domain are visible in Figure 6.2. Similar observations have been made with SPN1S endolysins (Park *et al.*, 2014b) which was the basis for the model from the PDB homology searched and KZ144 (Briers *et al.*, 2007). The SPN1S endolysin has a structure like SM07 described in this study, despite the sequence similarity coming in at just above 25%. Another similarity was observed with the 2DBT and 1WVU structures. These; however, are Gram-positive infecting endolysins (Kezuka *et al.*, 2006). This makes for an interesting observation; the SM07 identified in our study possesses a CBD unexpected for lytic protein targeting a Gram-negative bacterium. Gram-positive infecting endolysins typically contain two domains (CBD and EAD), while Gram-negative infecting endolysins usually have EAD only (Schmelcher *et al.*, 2012).

Using secondary structure analysis, we pinpointed the catalytic domain location at (α-helix) six to (α-helix) 15 with the cell wall binding domain located between the first and second α-helices. SM07 structure examination led to the identification of a catalytic triad (Glu228, Glu237 and Asn260). The catalytic residues in comparison to similar structures (PDB: 4OK7 and 2DBT), Glu228 and Glu237 were found to form a catalytic dyad synergistically functioning to facilitate the catalytic function of an enzyme. There is a misconception about the involvement of Asn260 as part of a triad. The Asn260 residue may still play a critical role in the structure or function, but its location outside of the ligand binding region makes this unlikely. Work conducted by Park et al., (2014) strongly suggests that the catalytic dyad contributes to the catalytic activity. The discovery of these distinct functional domains provides valuable insight into the molecular basis of endolysin activity, paving the way for further investigations into its application as a targeted therapeutic agent.

6.5 Conclusion

The focus of this study was to use a computational approach for the structural modelling of SM07. Four computational tools were used, and AlphaFold produced the most favourable and reliable structure, followed by SwissModel. The quality and integrity of the predicted structure was assessed using various software packages. Pymol enabled us to scrutinize the protein's intricate features visually. With deeper investigation exploring various other bioinformatics tools, intriguing findings were revealed. A CBD was identified which is unusual for Gram-negative infecting endolysin and a catalytic/enzymatic active domain (EAD) was aso identified in the N-terminal of the structure. Moreover, a catalytic dyad (Glu228 and Glu237) was also identified. Nonetheless, additional laboratory studies are required to better understand and

validate the presence and functional roles of the identified domains. Comprehensive approach yielded a highly accurate model and gained an unusual insight into the protein's structure and function, contributing to the broader understanding of SM07.

6.6 Reference

Baker, D. and Sali, A. (2001) Protein structure prediction and structural genomics. Science (1979) 294: 93–96.

Bordoli, L., Kiefer, F., Arnold, K., Benkert, P., Battey, J., and Schwede, T. (2009) Protein structure homology modeling using SWISS-MODEL workspace. Nat Protoc 4: 1–13.

Briers, Y., Volckaert, G., Cornelissen, A., Lagaert, S., Michiels, C.W., Hertveldt, K., and Lavigne, R. (2007) Muralytic activity and modular structure of the endolysins of Pseudomonas aeruginosa bacteriophages φKZ and EL. Mol Microbiol 65: 1334–1344.

Cernooka, E., Rumnieks, J., Zrelovs, N., Tars, K., and Kazaks, A. (2022) Diversity of the lysozyme fold: structure of the catalytic domain from an unusual endolysin encoded by phage Enc34. Scientific Reports 2022 12:1 12: 1–11.

Colovos, C. and Yeates, T.O. (1993) Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci 2: 1511–1519.

David, A., Islam, S., Tankhilevich, E., and Sternberg, M.J.E. (2022) The AlphaFold Database of Protein Structures: A Biologist's Guide. J Mol Biol 434:.

Deng, H., Jia, Y., and Zhang, Y. (2018) Protein structure prediction. Int J Mod Phys B 32: 1840009.

Eisenberg, D., Lüthy, R., and Bowie, J.U. (1997) VERIFY3D: Assessment of protein models with three-dimensional profiles. Methods Enzymol 277: 396–404.

Eswar, N., Eramian, D., Webb, B., Shen, M.Y., and Sali, A. (2008) Protein structure modeling with MODELLER. Methods Mol Biol 426: 145–159.

Farhadi, T. and Hashemian, S.M.R. (2018) Computer-aided design of amino acid-based therapeutics: A review. Drug Des Devel Ther 12: 1239–1254.

Fenton, M., Mcauliffe, O., O'mahony, J., and Coffey, A. (2010) Recombinant bacteriophage lysins as antibacterials.

Fischetti, V.A. (2008) Bacteriophage lysins as effective antibacterials. Curr Opin Microbiol 11: 393–400.

Fischetti, V.A. (2018) Development of phage lysins as novel therapeutics: A historical perspective. Viruses 10:.

Fiser, A. (2010) Template-Based Protein Structure Modeling. Methods Mol Biol 673: 73.

Haas, J., Roth, S., Arnold, K., Kiefer, F., Schmidt, T., Bordoli, L., and Schwede, T. (2013) The Protein Model Portal—a comprehensive resource for protein structure and model information. Database 2013:.

Hall, T. (2011) BioEdit: An important software for molecular biology Software Review. GERF Bulletin of Biosciences 2: 60–61.

Kezuka, Y., Ohishi, M., Itoh, Y., Watanabe, J., Mitsutomi, M., Watanabe, T., and Nonaka, T. (2006) Structural Studies of a Two-domain Chitinase from Streptomyces griseus HUT6037. J Mol Biol 358: 472–484.

Kortemme, T. and Baker, D. (2004) Computational design of protein–protein interactions. Curr Opin Chem Biol 8: 91–97.

Kryshtafovych, A. and Fidelis, K. (2009) Protein structure prediction and model quality assessment. Drug Discov Today 14: 386–393.

Laskowski, R.A., Rullmann, J.A.C., MacArthur, M.W., Kaptein, R., and Thornton, J.M. (1996) AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. J Biomol NMR 8: 477–486.

Lin, D.M., Koskella, B., and Lin, H.C. (2017) Phage therapy: An alternative to antibiotics in the age of multi-drug resistance. World J Gastrointest Pharmacol Ther 8: 162–173.

Low, L.Y., Yang, C., Perego, M., Osterman, A., and Liddington, R.C. (2005) Structure and lytic activity of a Bacillus anthracis prophage endolysin. Journal of Biological Chemistry 280: 35433–35439.

Masrati, G., Landau, M., Ben-Tal, N., Lupas, A., Kosloff, M., and Kosinski, J. (2021) Integrative Structural Biology in the Era of Accurate Structure Prediction. J Mol Biol 433: 167127.

Miguel, R.N.R., Shi, J., and Mizuguchi, K. (2002) Protein fold recognition and comparative modeling using HOMSTRAD, JOY and FUGUE. IUL Biotechnology Series 3: 143–169.

Nelson, D.C., Schmelcher, M., Rodriguez-Rubio, L., Klumpp, J., Pritchard, D.G., Dong, S., and Donovan, D.M. (2012) Endolysins as Antimicrobials. Adv Virus Res 83: 299–365.

Park, Y., Lim, J.A., Kong, M., Ryu, S., and Rhee, S. (2014) Structure of bacteriophage SPN1S endolysin reveals an unusual two-module fold for the peptidoglycan lytic and binding activity. Mol Microbiol 92: 316–325.

Ramachandran, G.N., Ramakrishnan, C., and Sasisekharan, V. (1963) Stereochemistry of polypeptide chain configurations. J Mol Biol 7: 95–99.

Šali, A. and Blundell, T.L. (1993) Comparative Protein Modelling by Satisfaction of Spatial Restraints. J Mol Biol 234: 779–815.

Schmelcher, M., Donovan, D.M., and Loessner, M.J. (2012) Bacteriophage endolysins as novel antimicrobials. Future Microbiol 7: 1147.

Wiederstein, M. and Sippl, M.J. (2007) ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res 35: W407–W410.

Xu, D., Xu, Y., and Uberbacher, E. (2005) Computational Tools For Protein Modeling. Curr Protein Pept Sci 1: 1–21.

Yang, J. and Zhang, Y. (2015) I-TASSER server: New development for protein structure and function predictions. Nucleic Acids Res 43: W174–W181.

7 CHAPTER SEVEN: GENERAL DISCUSSION, CONCLUSION AND FUTURE STUDIES ON THE DISCOVERY OF NOVEL PHAGE-DERIVED LYTIC PROTEINS

7.1 General discussion and conclusion

The discovery of novel enzymes through biotechnological approaches has significantly contributed to addressing challenges across diverse scientific fields. Bacteriophages have historically been used to combat bacterial infections; and their derivatives including phage lytic proteins have since gained attraction based on the prior knowledge of phage therapy. In this study, a public, online phage database from the University of Kwa-Zulu Natal was successfully screened for phage lytic proteins. Nowadays, sequencing techniques have made it possible to sequence hundreds of viral and bacterial genomes at a cost-effective means per year (Klumpp et al., 2012). Using a sequence-based approach, the database containing genome sequences was screened and four uncharacterized genes with potential lytic function were identified using NCBI blast tool. One drawback of a sequence-based approach is that the discovery of novel genes is largely affected by the accuracy of genome annotation and the completeness of the data available (Ngara and Zhang, 2018) and is based on existing data.

The uncharacterised genes selected had catalytic domains such as lysozyme, and amidase and PG binding domain, domains which are essential for the functionality of lytic proteins. The domains play different roles, such as cleaving specific bonds within the peptidoglycan (PG) component (Heselpoth et al., 2018; Shrivastava et al., 2019), and binding to a bacterial cell wall, all of which aid in the specificity of the protein (Heselpoth et al., 2018). Exploring these uncharacterized sequences with these domains provides valuable insights into developing alternatives in the fight against antibiotic-resistant bacteria. Recombinant production of the proteins initially was conducted using transient expression system using the pEAQ-HT vector, but yielded inconclusive results as the only expression observed was of truncated proteins, particularly for SM03, SM29, and SM07; SM31 did not have expression. This poor expression could be because of suboptimal codon usage or protein instability during expression. With these proteins being high value proteins, the three truncated proteins were sub-cloned for bacterial expression into pET-30b(+) vector under the T7 promoter for expression. All the genes were cloned with a His_6 tag to facilitated easy purification with Ni-TED resin (Strickler, 2008).

A subset of the proteins fused with the His_6 tag on the C-terminal were initially expressed in BL-21(DE3) under different conditions and at 0.1mM IPTG. No visible soluble protein was observed on the SDS-PAGE; however, the vast majority of the protein was observed in the

insoluble fraction Figure A.6 to Figure A.8. The process of the inclusion body is not fully understood, but proteins being misfolded during expression or expression in the inclusion body is not uncommon for bacterial expression (Ramón et al., 2014; Fathi-Roudsari et al., 2016). One time point (T_4) post-induction was taken and purified (Figure 4.8) showed that the protein was not binding to the column as no corresponding band was observed in the eluates. Denaturing conditions were explored to improve purification, however, the challenges remained. Further purification was attempted using HIC after precipitation with 15% of ammonium sulphate. Lysin (SM07) purification improved while the purification remained limited for the other two (Figure 4.11).

To minimize the need for extensive purification steps required for HIC, the proteins were then expressed with the His6 tag on the N-terminal. The expression was done also using a different cell line [BL-21(AI)] at 37 °C induced with 0.1 mM IPTG and 0.2% arabinose (Figure 4.12 to Figure 4.14). SM07 was successfully purified with Ni-TED under normal conditions. A secondary purification step utilizing ion exchange chromatography (IEC) showed promise in eliminating contaminants; however, the cleaned protein showed some instability. Further characterization of the SM07 using peptide mass fingerprinting revealed three peptides (DLETVTLR, ALEILPAVR, and YAPYIGR), revealing the protein to be an endolysin (Figure 4.18).

Phage lytic proteins have the potential to be antimicrobial, the successfully purified SM07 was tested for its antimicrobial affects against ESKAPE pathogens, and it showed efficacy *P. aeruginosa*, without the need for OM destabilizers. Interestingly, it also exhibited inhibition against *Acinetobacter baumannii*, a Gram-negative bacterium. No inhibition was observed in Gram-positive bacteria. No significant toxicity was observed; this is a positive characteristic, as the goal is to combat the global rise in antibiotic resistance using proteins that have antimicrobial activity. A surface disinfectant using SM07 as the sole active ingredient demonstrated promising bactericidal activity (98,7%) against *P. aeruginosa.* This finding holds significance, especially in settings where complete sterilization is not mandatory, potentially benefiting industries requiring effective surface cleaners to combat specific pathogens like *P. aeruginosa*.

Additionally, a 3D structure was predicted to investigate SM07 function using computational methods. The four techniques used leveraged on homology, fold recognition, and ab initio techniques (Xu et al., 2005; Yang and Zhang, 2015). AlphaFold stood out, followed by Swiss-Model, which incorporates Alphafold in its systems. The model unveiled critical domains such as the binding domain, catalytic domains and catalytic residues, shedding light on the mechanisms of action of SM07. I-Tasser added another layer to the analysis by highlighting

the origin of 7sz4K as a *Pseudomonas* phage, supporting the antimicrobial activity observed against *P. aeruginosa*, as endolysin are known to be highly specific (Shrivastava *et al.*, 2019). Notably, catalytic residues (Glu228, Glu237, and Asn260) were also identified, based on SPN1S endolysins although the precise role of Asn260 remains uncertain. The structural insight opens doors to potential therapeutic applications.

In conclusion, this study underscores the significance of biotechnological approaches in discovering novel enzymes, particularly focusing on phage lytic proteins and their potential as antimicrobial agents. The exploration of a phage database led to the identification of four uncharacterized phage lytic proteins. Despite challenges in protein expression and purification, functional characterization of SM07 revealed the promising antimicrobial efficacy against ESKAPE pathogens, particularly *P. aeruginosa.* Computational methods in protein modelling provided structural insights into SM07s function, highlighting critical domains essential for its antimicrobial activity. This research holds promise for developing effective surface cleaners and therapeutic agents targeting specific pathogens, emphasizing the importance of interdisciplinary approaches in addressing contemporary challenges in microbiology and biotechnology.

7.2 Future Studies

7.2.1 Optimization of both bacterial and transient expression of the lytic proteins

Further investigations will be needed to optimize the expression of the lytic proteins in plants by codon optimization of the genes. Gene codon optimized might increase the expression yields in the system, providing in-depth knowledge of phage-derived lytic protein expression behaviour in plant expression systems.

Optimize bacterial expression of the N-terminal tagged SM03 and SM29 proteins and optimize their purification. Clone the Holin (SM31) into a bacterial expression vector for recombinant protein production and purification. Continued development of lytic proteins is critical for discovering next-generation antimicrobials.

7.2.2 Explore Alternative Expression Systems

Consider improving protein production by exploring other expression systems such as yeast. Yeast expression strains can be optimised for expression of one's target protein, thus improving expression yields of the protein.

7.2.3 Optimise Stability of Secondary Step Purified SM07

Consider exploring different buffer formulations and varying storage conditions. Additionally, conducting a comprehensive study on the influence of various chromatographic parameters and resin types could contribute to refining the purification process for enhanced stability.

7.2.4 Increase the Efficacy of the Surface Disinfected

Improve the efficacy of the surface disinfectant to increase bactericidal efficacy against *P. aeruginosa* from 98.4% to 99.99%. This can be achieved by increasing the concentration of SM07 in the formulation.

The efficacy can also be improved by introducing OM -disrupting agents to aid the functionality of SM07. Disrupting the OM will increase the potency of the SM07 by disrupting the OM and allowing SM07 to get to the PG more quickly.

7.2.5 Functional Analysis of catalytic residues and binding domain

Validate the functionality of the catalytic residues as a dyad/triad by performing an antimicrobial activity assay and using an outer membrane-permeabilized *E. coli* as a substrate.

Characterize the binding activity of SM07 to outer-membrane-permeabilized cells using fluorescence microscopy. This can be achieved by infusing GFP into the binding domain of SM07. Engineer out the binding domain of the gene and determine the functionality of the protein against the same pathogenic strains.

7.2.6 Functional analysis and stability of SM07's structure

Validate the functionality of the structure and stability as this will further highlight the efficacy of the potential as an antimicrobial. The functional analysis can b don though enzyme assays, binding studies, and mutagenesis experiments to identify active sites to support those already identified though computational approach. Additionally, validating the stability of the protein structure is essential to confirm its reliability and functionality. Techniques such as molecular dynamics simulations, thermal stability assays, and crystallography should be explored to assess structural integrity and resilience.

7.3 References

Fathi-Roudsari, M., Akhavian-Tehrani, A., and Maghsoudi, N. (2016) Comparison of Three Escherichia coli Strains in Recombinant Production of Reteplase. Avicenna J Med Biotechnol 8: 16–22.

Heselpoth, R.D., Swift, S.M., Linden, S.B., Mitchell, M.S., and Nelson, D.C. (2018) Enzybiotics: Endolysins and Bacteriocins.

Klumpp, J., Fouts, D.E., and Sozhamannan, S. (2012) Next generation sequencing technologies and the changing landscape of phage genomics. Bacteriophage 2: 190–199.

Ngara, T.R. and Zhang, H. (2018) Recent Advances in Function-based Metagenomic Screening. Genomics Proteomics Bioinformatics 16: 405–415.

Ramón, A., Señorale-Pose, M., and Marín, M. (2014) Inclusion bodies: not that bad…. Front Microbiol 5: 56.

Shrivastava, A., Shrivastava, N., and Singh, P.K. (2019) Enzymes in Pharmaceutical Industry. Enzymes in Food Biotechnology 591–602.

Strickler, O.U.M.S.D.C.J. (2008) Fusion Tags for Protein Expression and Purification. Biopharm Int.

Xu, D., Xu, Y., and Uberbacher, E. (2005) Computational Tools For Protein Modeling. Curr Protein Pept Sci 1: 1–21.

Yang, J. and Zhang, Y. (2015) I-TASSER server: New development for protein structure and function predictions. Nucleic Acids Res 43: W174–W181.

APPENDIX A: SUPPLEMENTARY DATA

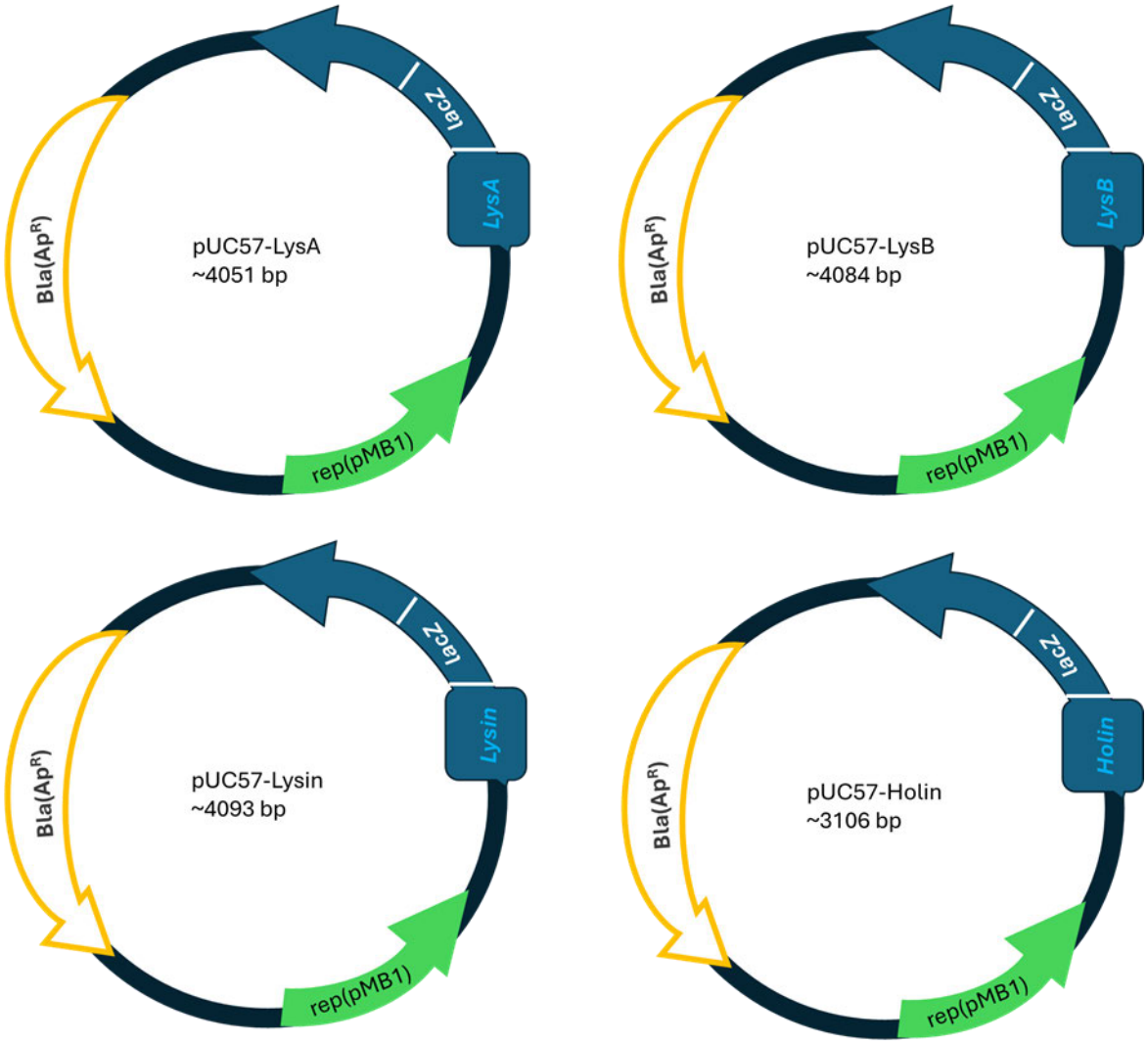


Figure A.1: Schematic representation of the genes synthesised in pUC57.

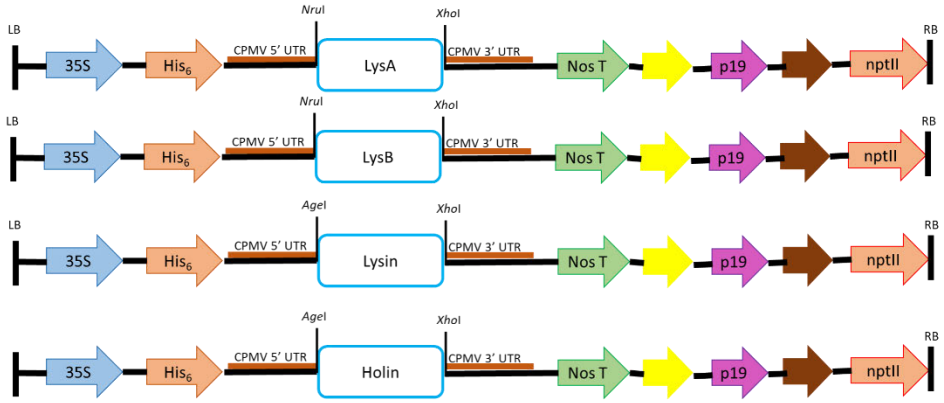


Figure A.2: Schematic representation of the four genes inserted into pEAQ-HT vector.

Table A.1: Infiltration and extraction buffer compositions for transient expression.

	Buffer	**Composition**	pH
1.	Infiltration Buffer 1	10 mM MES, 21.4 MgCl	5.6
2.	Infiltration Buffer 2	10 mM MES, 20 mM $MgSO_4$, 58.4 mM Sucrose	
3.	Extraction Buffer 1	0.137 M NaCl, 0.0027 M KCl, 0.01 M Na_2HPO_4, 0.0018 M KH_2PO_4	7.4
4.	Extraction Buffer 2	20 mM Na_3PO_4, 20 mM EDTA, 1 mM Protease Inhibitor, 20 mM $C_6H_8O_6$, 1 mM DTT, 1% Triton X-100	7.2

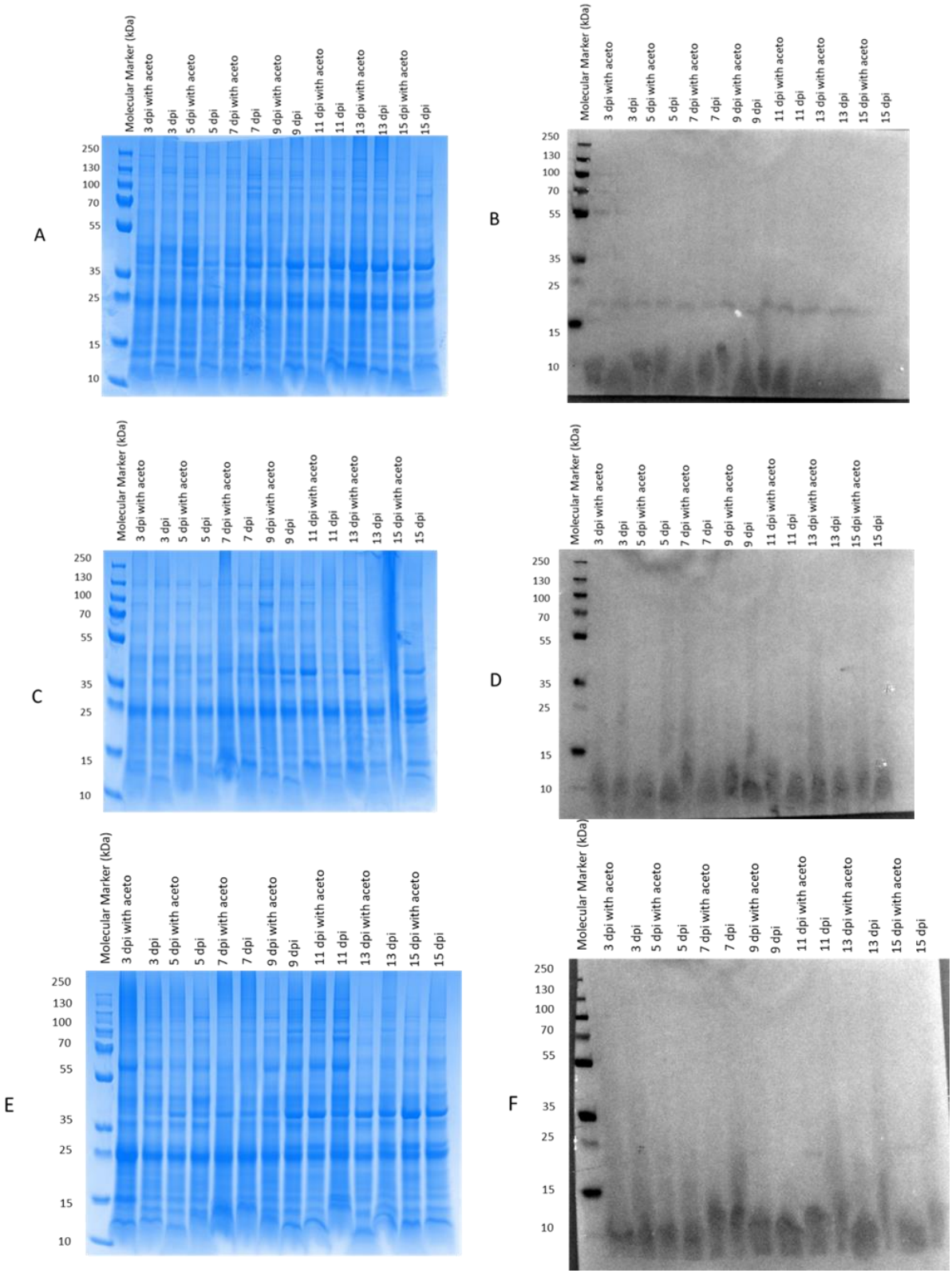


Figure A.3: SDS-PAGE (left) and western blot (right) analysis of (A-B) SM03, (C-D) SM29 and (E-F) SM07.
Time profile expression studies with LBA OD_{600} = 0.4, infiltrated with buffer 2 with and without aceto. Extraction buffer two was used for extractions.

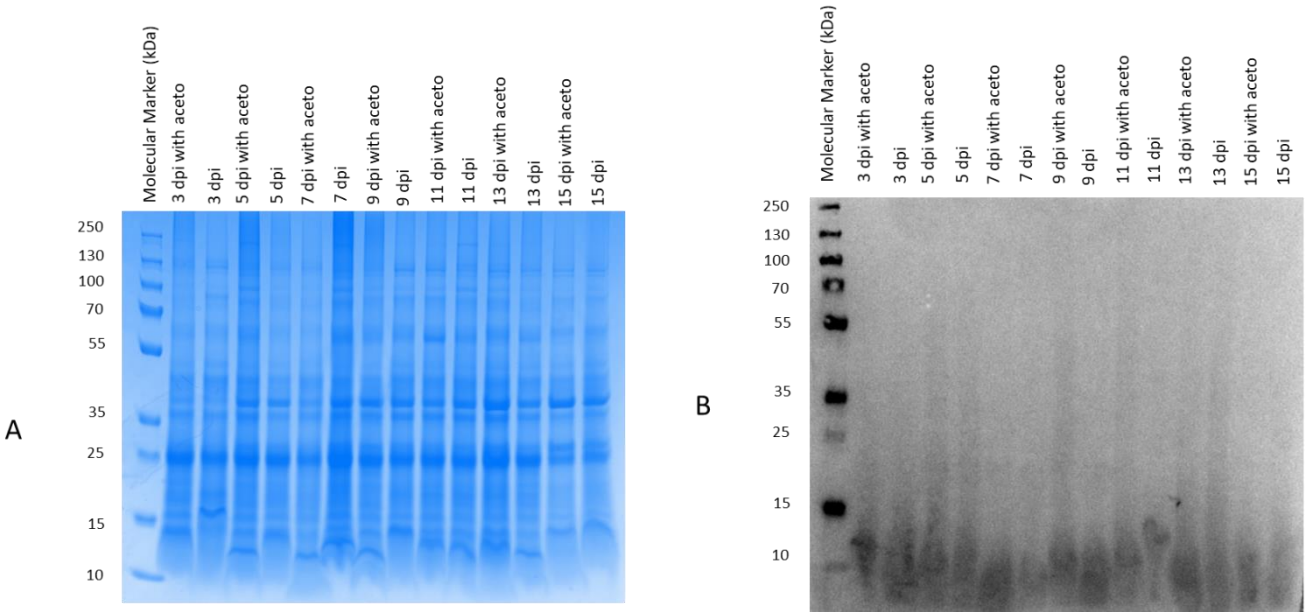


Figure A.4: SDS-PAGE (left) and western blot (right) analysis.
(A-B) SM31 time profile with LBA OD600 = 0.4, infiltrated with buffer 2 with and without aceto. Extraction buffer two was used for extractions.

Figure A.5: Illustration of SM03, SM29 and SM07 in pET-30b(+).

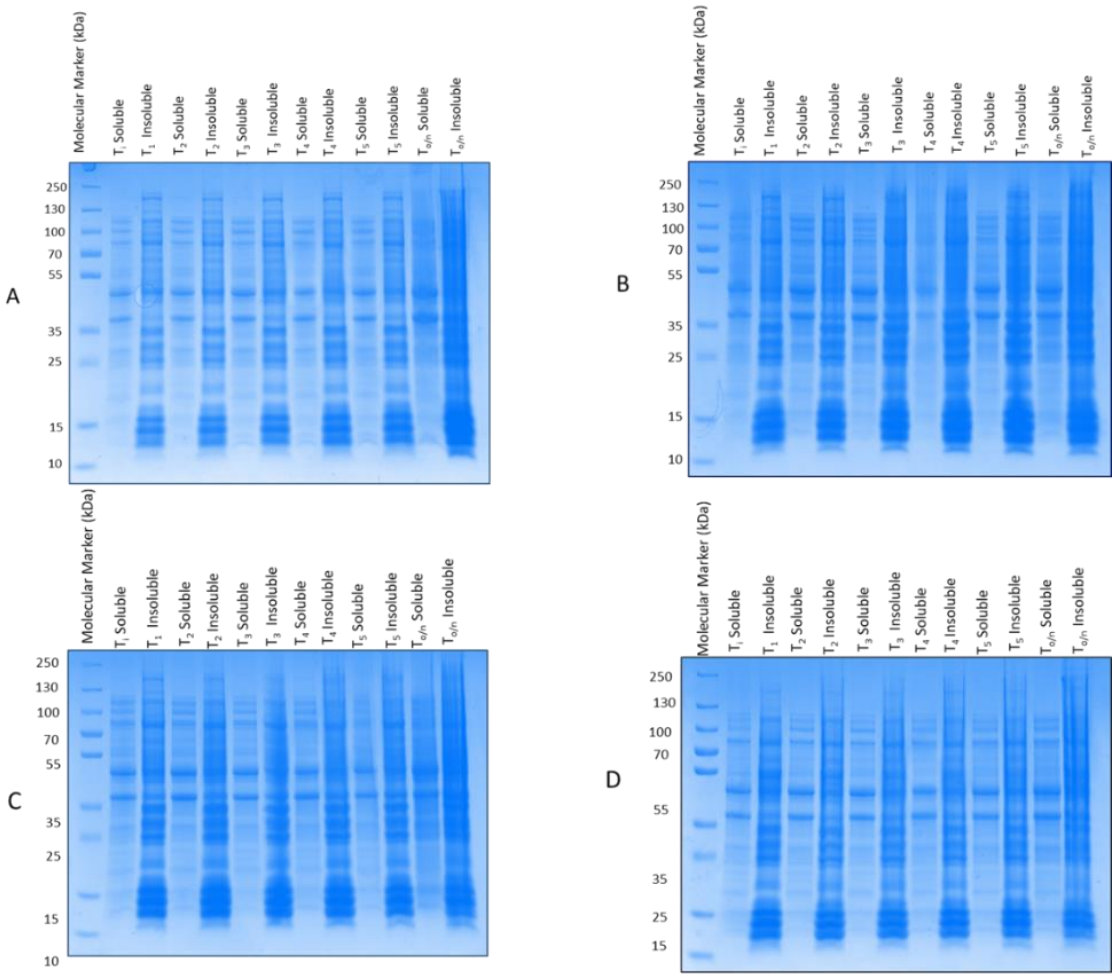


Figure A.6:SDS-PAGE analysis of the time profile of SM03 using LB in BL-21 (DE3) cells at four different temperature points.
(A) 17 °C, (B) 25 °C, (C) 30 °C, and (D) 37 °C, respectively. The figures depict the soluble and insoluble fractions.

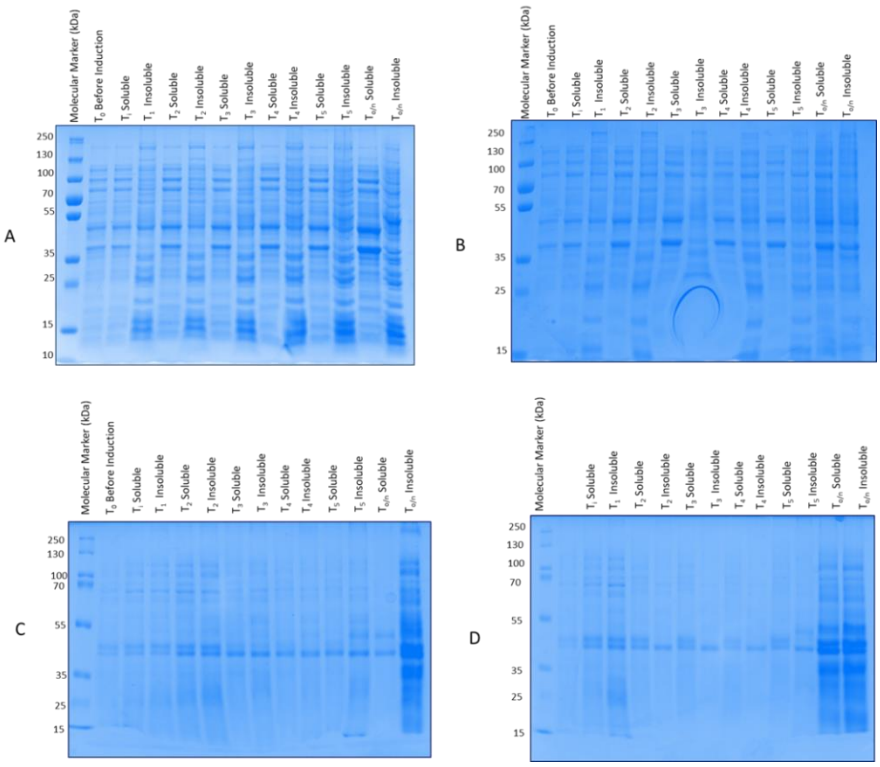


Figure A.7:SDS-PAGE analysis of the time profile of SM29 using LB in BL-21 (DE3) cells at four different temperature points.
(A) 17 °C, (B) 25 °C, (C) 30 °C, and (D) 37 °C, respectively. The figures depict the soluble and insoluble fractions.

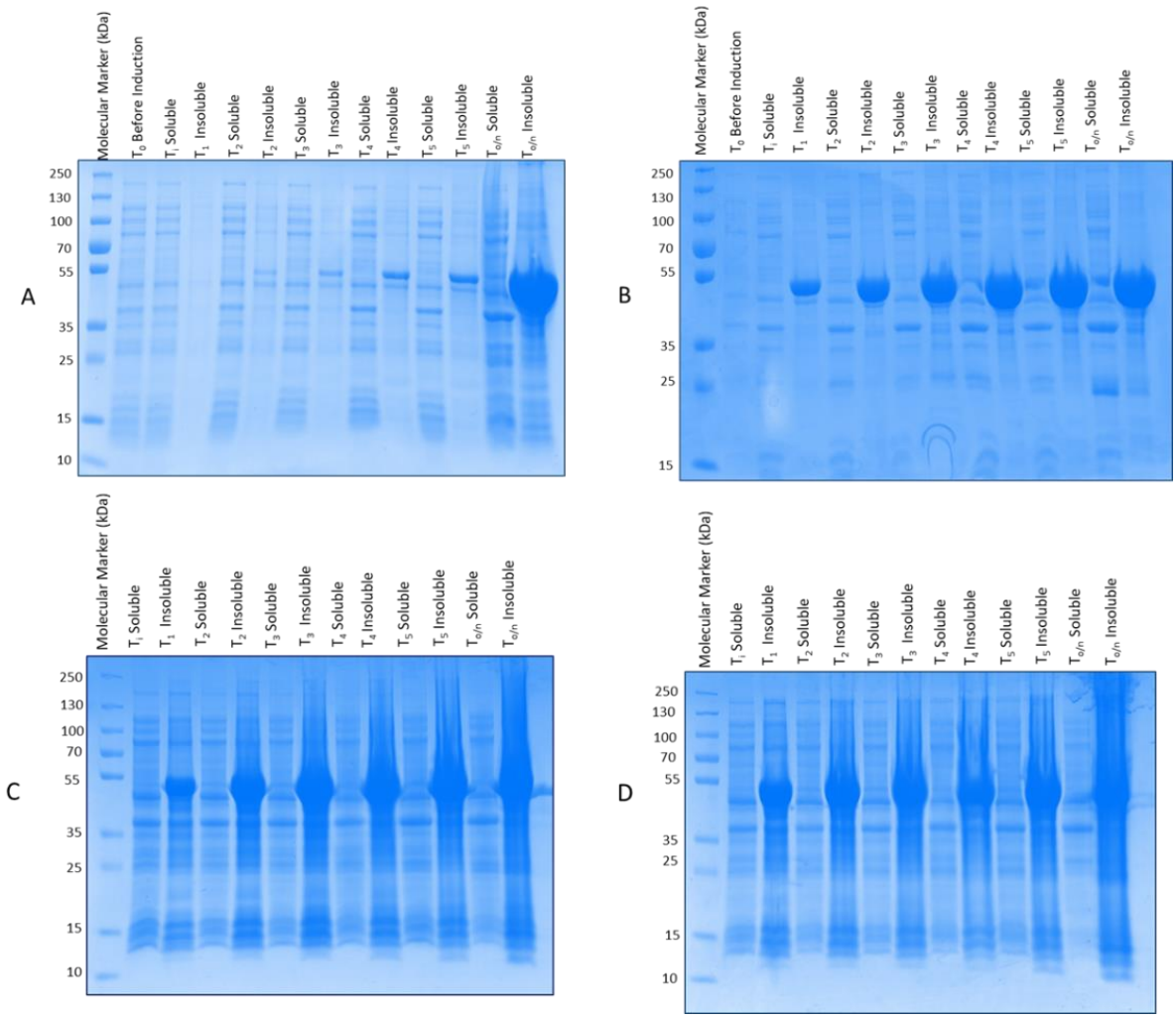


Figure A.8:SDS-PAGE analysis of the time profile of SM07 using LB in BL-21 (DE3) cells at four different temperature points.
(A) 17 °C, (B) 25 °C, (C) 30 °C, and (D) 37 °C, respectively. The figures depict the soluble and insoluble fractions.

Table A.2: Modification of the buffer used in purification for optimization of the purification.

#	**Composition**	
1	300 mM NaCl, 50 mM NaH_2PO_4 * 2 H_2O and 100 µg/µl DNase Inhibitor	pH 8
2	300 mM NaCl, 50 mM NaCl, 50 mM NaH_2PO_4 * 2 H_2O and 0.5 mM Dithiothreitol	
3	300 mM NaCl, 50 mM NaCl, 50 mM NaH_2PO_4 * 2 H_2O and Triton-X	
4	300 mM NaCl, 50 mM NaCl, 50 mM NaH_2PO_4 * 2 H_2O and 8 M Urea	

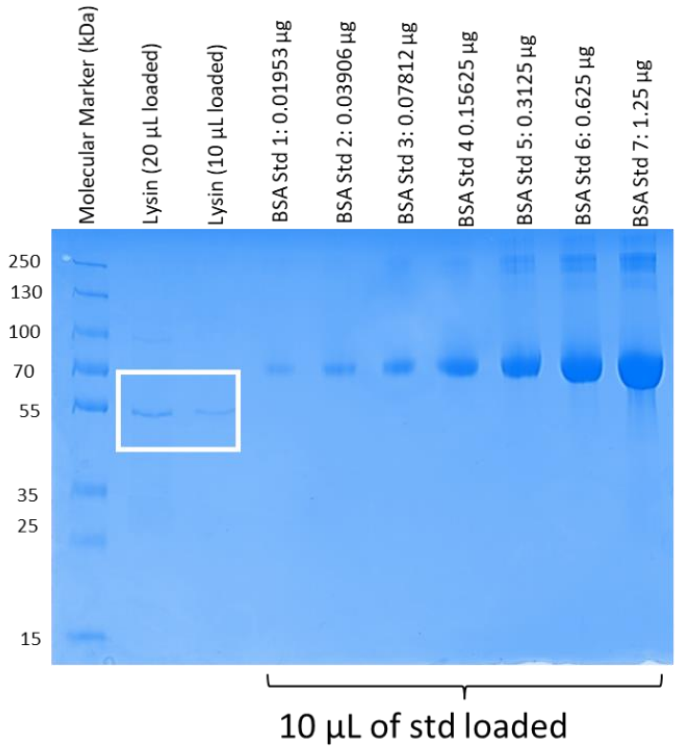


Figure A.9: SDS-AGE Bradford quantification of the polished SM07.

Table A.3: Validation tests.

Experimental Conditions Validation suspension (Nv_0)			**Experimental Conditions Control (A) = NvA**			**Filtration Control (B) = NvB**			**Method Validation (C)** **(Neat Product Concentration) = NvC**		
Ave			**Ave**			**Ave**			**Ave**		
Vc1	30	31	Vc1	26	23.5	Vc1	44	40	Vc1	37	39
Vc2	32		Vc2	21		Vc2	36		Vc2	41	
Acceptance limits	Nv_0 = 30 -160		**Acceptance limits**	≥ 0.5x Nv_0		**Acceptance limits**	≥ 0.5x Nv_0		**Acceptance limits**	≥ 0.5x Nv_0	
Complies	Yes		**Complies**	Yes		**Complies**	Yes		**Complies**	Yes	
0.5 x Nv_0 = 15.5											

Dilution **(Test suspension)**	**Vc1**	**Vc2**	**Average N (wm)**	**Log N**	**N0**	**Log N0**
10^{-6}	261	286	$2.77X10^8$	8.44	$2.77X10^7$	7.44
10^{-7}		30			32	

Acceptance limits	Log N is between 8.17 and 8.70	Complies	Yes
Acceptance limits:	Log N_0 is between 7.17 and 7.70	Complies	Yes
Acceptance limits	Control of weighted mean (wm) counts: 8.82	Complies	Yes

Product Concentration	**Vc1**	**Vc2**	**Na** **(Ave Vc1 & Vc2 x 10 000)**	**Log Na**	**Log Reduction** **(N0: 7.44)**	**Contact time**
Neat	31	44	375 000	5.57	1.87	5 min

VC = Viable Count, N = Test suspension, N0 = Test suspension at the beginning of contact time (t=0), Na = Test suspension (survivors) before membrane filtration, Nv = Validation suspension, Nv0= Validation suspension at the beginning of contact time, A = number of cfu/ml of the experimental conditions control, B = number of cfu/ml of the filtration control, C = number of cfu/ml of the method validation

APPENDIX B.1: PUBLICATION 1

Heliyon 9 (2023) e16723

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Review article

South Africa's indigenous microbial diversity for industrial applications: A review of the current status and opportunities

Varsha Chhiba [a], Priyen Pillay [a], Sibongile Mtimka [a,b], Ghaneshree Moonsamy [a], Lusisizwe Kwezi [a], Ofentse J. Pooe [b], Tsepo L. Tsekoa [a,*]

[a] Future Production: Chemicals Cluster, Council for Scientific and Industrial Research (CSIR), Pretoria, South Africa
[b] School of Life Sciences, Discipline of Biochemistry, University of KwaZulu-Natal, Durban, South Africa

ARTICLE INFO

Keywords:
Africa
Microbial diversity
Culture collection
Metagenomics
Metaviromics
Research networks
Biocatalysis
Industrial applications

ABSTRACT

The unique metagenomic, metaviromic libraries and indigenous micro diversity within Southern Africa have the potential for global beneficiation in academia and industry. Microorganisms that flourish at high temperatures, adverse pH conditions, and high salinity are likely to have enzyme systems that function efficiently under those conditions. These attributes afford researchers and industries alternative approaches that could replace existing chemical processes. Thus, a better understanding of African microbial/genetic diversity is crucial for the development of "greener" industries. A concerted drive to exploit the potential locked in biological resources has been previously seen with companies such as Diversa Incorporated and Verenium (Badische Anilin-und SodaFabrik-BASF) both building business models that pioneered the production of high-performance specialty enzymes for a variety of different industrial applications. The market potential and accompanying industry offerings have not been fully exploited in South Africa, nor in Africa at large. Utilization of the continent's indigenous microbial repositories could create long-lasting, sustainable growth in various production sectors, providing economic growth in resource-poor regions. By bolstering local manufacture of high-value bio-based products, scientific and engineering discoveries have the potential to generate new industries which in turn would provide employment avenues for many skilled and unskilled laborers. The positive implications of this could play a role in altering the face of business markets on the continent from costly import-driven markets to income-generating export markets. This review focuses on identifying microbially diverse areas located in South Africa while providing a profile for all associated microbial/genetically derived libraries in this country. A comprehensive list of all the relevant researchers and potential key players is presented, mapping out existing research networks for the facilitation of collaboration. The overall aim of this review is to facilitate a coordinated journey of exploration, one which will hopefully realize the value that South Africa's microbial diversity has to offer.

1. Introduction

Microbes are some of the most ubiquitous and abundant organisms that reside on planet Earth and play a crucial role in creating the

* Corresponding author. P.O. Box 395, Pretoria, 0001, South Africa.
E-mail address: TTsekoa@csir.co.za (T.L. Tsekoa).

https://doi.org/10.1016/j.heliyon.2023.e16723
Received 19 September 2022; Received in revised form 5 May 2023; Accepted 25 May 2023
Available online 1 June 2023

APPENDIX B.2 : PUBLICATION 2

microorganisms

MDPI

Review

An Exploratory Review of the Potential of Lytic Proteins and Bacteriophages for the Treatment of Tuberculosis

Sibongile Mtimka [1,2], Priyen Pillay [1], Lusisizwe Kwezi [1], Ofentse Jacob Pooe [2] and Tsepo Lebiletsa Tsekoa [1,*]

[1] Future Production: Chemicals Cluster, Council for Scientific and Industrial Research, Pretoria 0001, South Africa; 219094519@stu.ukzn.ac.za (S.M.); ppillay3@csir.co.za (P.P.); lkwezi@csir.co.za (L.K.)

[2] Discipline of Biochemistry, School of Life Sciences, University of KwaZulu-Natal, Durban 4000, South Africa; pooeo@ukzn.ac.za

* Correspondence: ttsekoa@csir.co.za; Tel.: +27-(0)12-841-3274

Abstract: Tuberculosis (TB) is a highly prevalent infectious disease that causes more than 1.5 million deaths a year. More than 25% of TB deaths occur in Africa, and TB is South Africa's leading cause of death, with about 89,000 people dying of it yearly. The emergence of multidrug-resistant TB (MDR-TB) poses a significant threat to health security and could reverse the positive gains already made in the fight against TB. Antibiotic treatments are available, but side effects and the alarming increase in the prevalence of drug-resistant strains of *Mycobacterium tuberculosis* (*Mtb*) will compromise the control of the spread and treatment of the disease. A promising option is to employ specialized enzymes encoded by bacteriophages, which destroy bacterial cell membranes and walls to treat tuberculosis. Phage therapy against bacteria is a known treatment that is now reemerging with lytic proteins. These proteins provide an alternative means to treat infectious diseases where conventional antibiotic regimens do not meet the requirements. This review explores and discusses the potential of lytic protein therapy as an antimicrobial strategy against *M. tuberculosis* and multidrug-resistant tuberculosis.

Keywords: phage-derived lytic proteins; *Mycobacterium tuberculosis*; antibiotic resistance

Citation: Mtimka, S.; Pillay, P.; Kwezi, L.; Pooe, O.J.; Tsekoa, T.L. An Exploratory Review of the Potential of Lytic Proteins and Bacteriophages for the Treatment of Tuberculosis. *Microorganisms* **2024**, *12*, 570. https://doi.org/10.3390/microorganisms12030570

Academic Editors: Grzegorz Wegrzyn, Mariateresa Vitiello, Roberta Colicchio and Chiara Pagliuca

Received: 9 January 2024
Revised: 5 March 2024
Accepted: 8 March 2024
Published: 12 March 2024

1. Introduction

The current antibiotic resistance crisis poses a significant global threat to public health. The emergence of antibiotic resistance is due to (i) overuse, (ii) inappropriate prescribing, (iii) extensive agricultural use, and (iv) the scarcity of new antibiotics/drugs. These factors have rendered many conventional antibiotics ineffective against infections. Multidrug resistance is caused by multiple biochemical factors and presents a serious challenge in fighting infectious diseases such as *Mycobacterium tuberculosis* (*Mtb*). The growing resistance to antibiotics is alarming, as it limits treatment options, prolongs illnesses, increases healthcare costs, and increases mortality rates. Infections that were once easily treatable are now life-threatening due to antibiotic resistance. If left unchecked, this trend could lead to a future in which common infections and routine medical procedures become dangerous and threaten lives. Furthermore, microbes can employ more than one mechanism, making it even more challenging to combat multidrug resistance.

Drug resistance develops in a number of ways, but the five major ones are: (i) mutation or enzymatic alteration of the bacterial target, (ii) inactivation of the antibiotic by enzymatic degradation and modifications, (iii) bypass by the bacterium of the inhibition mechanism, (iv) overexpression of the drug target, (v) inhibition or slowing down of drug penetration by certain groups of proteins and the cell wall, and (vi) efflux of the antibiotic efflux pumps [1,2]. The development of new antimicrobials, such as endolysins, is crucial to combating this crisis. Endolysins offer a promising solution by precisely targeting bacterial pathogens while potentially evading existing resistance mechanisms. Their ability to disrupt the walls of bacterial cells presents a unique approach that might help overcome the

APPENDIX B.3: ORAL PRESENTATION 1

CSIR Human Capital Development

PO Box 395 Pretoria 0001 South Africa
Tel: +27 12 841 2911
Email: Enquiries@csir.co.za

LETTER OF ACCEPTANCE

Dear Sibongile

Abstract: **Discovery of novel enzymes with potential for antimicrobial activity.**

You had submitted an abstract for **oral presentation** for the 7th CSIR Emerging Researchers Symposium. I am pleased to inform you that it has been accepted for **an oral presentation.**

Thank you for your interest and cooperation. We hope that you will find your participation in the symposium rewarding and professionally stimulating.

The final programme will be published on the conference website shortly.

Tentative Title: Discovery of novel enzymes with potential for antimicrobial activity.
Presenter: Ms Sibongile Mtimka

Kindly confirm your acceptance by sending an email indicating your interest to Mr Justice Komane at emergingresearchersymposium1@csir.co.za by close of business **(16:30 pm)** on **Thursday, 02 June 2022**.

Kind regards,

Justice Komane
Project Administrator
7th CSIR Emerging Researchers Symposium
[redacted]

APPENDIX B.4: ORAL PRESENTATION 2

SASBMB 2022
BIOCHEMISTRY MATTERS
originally SASBMB 2020

13 January 2022

Dear Sibongile Mtimka,

We are pleased to inform you that your submission entitled "DISCOVERY OF NOVEL ENZYMES WITH POTENTIAL OF ANTIMICROBIAL ACTIVITY", Submission ID: 153, has been accepted as an **ORAL PRESENTATION** at the SASBMB 2022 Congress.

Below, please find your provisional presentation details. Please note that these may still change. The programme will be published on the website asap. Please check back regularly for updates and changes.

Presentation Date:	Monday, 24 January
Presentation Time:	14:00 - 14:15
Session Name:	Environmental Biochemistry and Phytomedicine
Full Session Time:	13:30 - 15:15
Session Joining Time:	1:20:00 PM

Remember that your "presentation time" includes 5 minutes for questions and answers. Please keep this in mind when preparing your presentation. The session chairs will be asked to ensure speakers do not extend beyond their allotted slot.

Where possible, prepare slides in 16:9 format for optimal viewing.

Presenters must join their sessions 20 min before the session start time to allow the platform technical host to brief them. A dedicated calendar invite will be sent to you in this regard.

By default, all presentations will be live. Nevertheless, we request you submit a video recording as a backup. Use any software you are familiar with such as PowerPoint, Zoom, Google Meet, etc. for the recoding. Guidelines for PowerPoint are available through the conference website https://sasbmb2022.org.za. See "scientific programme", then "oral presentation guidelines".