

**BIOPHYSICAL AND BIOLOGICAL FUNCTIONAL ANALYSIS OF
RECOMBINANT PROTEINS EXPRESSED
IN *ESCHERICHIA COLI***

WISDOM JOHN NZUBE OKOYE

PhD

2023

BIOPHYSICAL AND BIOLOGICAL FUNCTIONAL ANALYSIS OF
RECOMBINANT PROTEINS EXPRESSED
IN *ESCHERICHIA COLI*

WISDOM JOHN NZUBE OKOYE

Department of Pharmacy and Pharmaceutical Sciences

A thesis submitted in partial fulfilment of the
requirements of the University of Sunderland for
the degree of Doctor of Philosophy

March 2023

Table of Contents

Abstract.....	viii
Declaration.....	x
Dedication.....	xi
Acknowledgments.....	xii
List of Abbreviations.....	xiii
Chapter One.....	1
Introduction.....	1
1.0. Protein-Protein and Protein-Ligand interactions.....	1
1.0.1 Specificity.....	3
1.0.2 Affinity.....	3
1.1.0 Protein Oligomerisation.....	5
1.2.0 Binding Interaction Techniques.....	6
1.2.1 Surface Plasmon Resonance.....	7
1.2.2 Principle of SPR.....	8
1.2.3 SPR Coupling Methods.....	9
1.2.4 Other Binding Interaction Strategies.....	12
1.2.4.1 Thermal Shift Assay.....	13
1.2.4.2 Isothermal Calorimetry.....	14
1.3.0 Protein Preparation for Binding Interaction Studies.....	15
1.4.0 Protein synthesis in prokaryotic and eukaryotic cells.....	16
1.4.1 Promoters.....	17
1.4.2 Positive inducible.....	18
1.4.3 Negative inducible.....	18
1.5.0 Cloning and Expression Vectors.....	23
1.6 Protein expression.....	24
1.7.0 Protein expression systems.....	26
1.8.0 <i>Escherichia coli</i> Expression System Advantages, and Limitations.....	27
1.9.0 Regulation of Gene Expression.....	33
1.9.1 The <i>lac</i> operon.....	34
1.9.2 The <i>lac</i> repressor.....	35
1.9.3 Catabolite activator protein (CAP).....	37
Research Questions (Aims of the Study).....	39
2. Materials and methods.....	42

2.1.0 Chemical and Reagents	42
2.1.1 Media	42
2.1.2 Solid Media	42
2.1.3 Liquid Media (Broth)	43
2.1.4 Selective Media.....	43
2.1.5 Bacterial Strains, Plasmids and Cloned Genes	43
2.2.0 DNA Methods.....	44
2.2.1 Extraction of Plasmid DNA.....	44
2.2.2 Plasmid Maps	44
2.2.3 Electrophoretic analyses of DNA.....	44
2.2.4 Restriction Analysis	45
2.2.5 Cloning of amplified DNA in Vector through Ligation.....	46
2.2.6 VAP Protein Plasmid-DNA Constructs.....	47
2.2.7 Transformation of chemically competent <i>Escherichia Coli</i> DH5 α	47
2.2.8 Transformation into BL21 (DE3) chemically competent <i>Escherichia Coli</i> Strain.....	48
2.2.9 Transformation into C43 (DE3) chemically competent <i>Escherichia Coli</i> Strain	48
2.3.0 Post-Overexpress Plasmid Stability Assay.....	49
2.4.0 Protein Method.....	49
2.4.1 Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE Analysis).....	49
2.4.2 Protein Overexpression	50
2.4.3 Large-scale gene expression	51
2.4.4 Production of Lysates.....	51
2.4.5 Protein Refolding and Solubilisation.....	52
2.4.6 Extraction of Soluble Protein by Further Mechanical Approach	52
2.5.0 Protein Purification Analysis	53
2.5.1 Batch Purification.....	53
2.5.2 IMAC Column Chromatography Protein Purification.....	54
2.5.3 Bradford Assay	55
2.5.4 Cleavage Restriction.....	55
2.6.0 Protein Isolation.....	56
2.6.1 Protein Isolation by MBP Reverse-Batch Purification.....	56
2.6.2 Protein Isolation by Heparin Reverse-Batch Purification	57
2.6.3 Isolation by Centrifugal Filtration	57
2.6.4 Buffer Exchange Optimised Centrifugal Filter Isolation.....	58
2.6.5 2 nd Buffer Exchange and Size Differential Optimised Centrifugal Filter Isolation.....	59
2.6.6 Size Differential Exclusion Based Centrifugal Filter Devices Isolation	60

2.7.0 Isolation of Disrupted Cellular Debris by Dialysis	61
2.7.1 Dialysis of Biotin-Heparin Sodium Salt.....	61
2.8.0 Surface Plasmon Resonance	62
2.8.1 SPR Capturing Methods	63
2.8.1.1 VapA-pET22a Immobilisation.....	63
2.8.1.2 Biotinylated-Heparin Immobilisation.....	64
2.9.0. SPR Binding Affinity and Kinetics Interaction Methods	65
2.9.1 Binding Interaction of <i>R. equi</i> Virulence Associated Proteins.....	65
2.9.2 Binding Interaction Chemokine - Biotinylated Heparin	65
2.10. Other Binding Interaction Assay Techniques.....	66
2.10.1 Trans Endothelium Flow Assay	66
2.10.1.1 Cell culture	66
2.10.1.2 Seeding the flow chamber	67
2.10.1.3 Performing the flow assays under flow condition.	67
2.10.2 Thermal Shift Assay.....	68
3.0 Case Study: Coronary Heart Disease.....	70
3.1.0 Pharmaceutical CHD Treatment Options.....	72
3.2 Mechanism of Action of CHD Drugs.....	73
3.2.1 Statins	73
3.2.2 Statin Binding Receptor	75
3.2. 2.1 Hydroxy-3-methylglutaryl coenzyme A reductase (HMG-CoA) Reductase Enzyme	75
3.2.2.2 Statin Interaction with HMGCR.....	76
3.3.0 HMGCR Gene Mutants and Polymorphisms.....	82
3.3.1 HMGCR Gene Exon 13 Splicing	83
3.4.0 AGTR1 Genes and Angiotensin II Receptor Blockers	85
3.4.1 AGTR1 Gene Mutants and Polymorphisms.....	89
Chapter Aim	91
3.5.0 Amplification of Full-length HMGCR Gene	93
3.5.1 DNA Extraction of Full-length WT- HMGCR Gene.....	94
3.5.2 Cloning of Amplified WT-HMGCR into pGEX-6p-1 Vector and Analysis	95
3.5.3 HMGCR gene cloning and over-expression challenges and troubleshooting attempts	96
3.5.4 Recombinant Overexpression of Full-length WT-HMGCR-GST Fusion Protein.	97
3.5.5 Comparative HMGCR-GST fusion protein Overexpression in DH5 α <i>E. coli</i> strain and HMGCR-GST fusion protein expression in C43 (DE3) <i>E. coli</i> strain	98
3.5.6 Solubilisation and Refolding of HMGCR-GST Fusion Protein.....	102
3.6.0 AGTRI Protein Synthesis.....	105

3.6.1 WT-AGTR1 Cloning into <i>E. coli</i> Vector pGEX-6p-1	106
3.6.2 Overexpression of recombinant WT-AGTR1-GST fusion Protein.....	108
3.7.0 AGTR1 over-expression challenges and troubleshooting attempts (problem-solving attempts)	109
3.7.1 Recombinant protein AGTR1 Overexpression using GST-fusion protein	109
3.7.2 Optimization of AGTR1 Overexpression Parameters.....	112
3.7.2.1 Vectors	112
3.7.2.2 Cloning of WT-AGTR1 into <i>E. coli</i> Vector pET32a	113
3.7.2.3 Cloning of WT- AGTR1 into <i>E. coli</i> Vector pMAL-p5x.....	115
3.7.3 AGTR1-Fusion Protein Overexpression Temperature Viability Assay.....	115
3.7.4 Comparative Expression of AGTR1-HIS tagged fusion protein at 37°C and 17°C.....	116
3.7.5 Optimizing Choice of Bacteria Host Cell Strain	119
3.7.5.1 AGTR1-HIS-Tagged Fusion Protein Expression in C43 (DE3) Strain	121
3.8.0 Other Troubleshooting Interventions and chapter conclusion	124
4.0. Chapter Four	126
4.1 Introduction	126
4.1.1 Chemokine and Chemokine Receptors.....	126
4.1.2 Chemokine Structures.....	127
4.1.3 Cellular Recruitment for Chemokine Activities.....	130
4.1.4 CCL5/RANTES	132
4.1.5 CCL5 and Inflammatory Diseases.....	133
4.1.6 Structure of CCL5	136
4.1.7 Glycosaminoglycan	136
4.1.8 Heparin and Heparan sulphate	138
4.1.9 CCL5 Interaction with Heparin and HSPGs during Inflammation	140
4.1.10 CCL5 Oligomerization in GAG-Binding Interaction	142
4.1.0 Research Aim	143
4.2.0 Results.....	144
4.2.1 Introduction	144
4.2.2 CCL5 DNA Extraction.....	147
4.2.3 Wild Type CCL5 Cloning in Different <i>E. coli</i> Vectors	149
4.2.4 Cloning of CCL5 Mutants into pMAL-c5x <i>E. coli</i> Vector	151
4.2.5 Protein Overexpression and Solubilisation of Wild-Type CCL5-MBP and CCL5 Mutants Fusion Protein.....	153
4.2.6 Solubilisation of WT-CCL5-MBP Protein	156
4.2.7 Batch Purification of WT-CCL5 and Mutant CCL5-MBP Fusion Protein.....	158

4.2.8 Cleave Removal of WT-CCL5-MBP and Mutant CCL5-MBP Fusion Protein	161
4.3.0 Isolation of Recombinant CCL5-MBP Protein from Cleaved Fusion Protein.....	163
4.3.1.1 Differential Size Exclusion Amicon Separation of CCL5/MBP Mutants	167
4.3.1.2 Heparin Resin Batch Purification	168
4.3.2 Sample Validation of Pure Recombinant CCL5-MBP Fusion Protein	169
4.3.3 Biotin-Heparin Immobilization.....	174
4.3.4 Binding Analysis of Recombinant CCL5-MBP Protein to GAG- Heparin.....	175
4.3.5 Endothelial Migration Flow assay videos.....	184
4.4.0 Discussion.....	187
4.4.1 Introduction	187
4.4.2 Overexpression of recombinant Chemokine MBP fusion proteins	190
4.4.3 Chemokine Protein Purification.....	192
4.4.4 Removal of MBP-Tag.....	192
4.4.5 Biotin-Heparin Sample Validation for Surface Plasmon Resonance	197
4.4.6 SPR interaction of Chemokine proteins in free solution with immobilized Biotin-Heparin ..	198
4.4.7 Endothelia Flow Chamber Assay.....	201
Chapter Five	204
5.0 Introduction	204
5.1 Genome and Virulence Associated Proteins.....	206
5.2 <i>Rhodococcus equi</i> infections: pathogenesis and pathology.....	209
5.3 Mechanism of <i>Rhodococcus equi</i> infectivity.....	210
5.4 <i>Rhodococcus equi</i> Virulent Factors	211
5.4.1 Chromosome-encoded Factors.....	212
5.4.2 Environmental Factors	214
5.4.3 Plasmid-encoded Factors.....	214
5.5 Immunological aspects of <i>Rhodococcus equi</i> infections.....	215
5.6 Aim of Study.....	221
5.7 Results.....	222
5.7.1 Cloning, Purification, and Characterization of Vap proteins	222
5.7.1.1 Introduction	222
5.8 Amplification of <i>R. equi</i> Vap Genes and Cloning.....	223
5.9 Recombinant <i>R. equi</i> VAP Protein Expression	223
5.10 Recombinant <i>R. equi</i> VAP Protein Purification	229
5.11 Isolation of Pure Recombinant Vap - His Fusion Protein.....	234
5.12 SPR Interaction of <i>VapA</i> with other <i>Vaps</i>	238
5.12.1 <i>VapA</i> -pET23 Immobilization.....	239

5.12.2 Binding Analysis of Recombinant <i>R. equi</i> Virulence Associated Proteins.....	240
5.13 Thermal Shift Assay.....	249
5.14 Discussion.....	251
5. 14. 1 Introduction	251
5.14.2 Overexpression of recombinant Vap-His Tag fusion proteins.....	253
5.14.3 Recombinant Vap Protein Purification	254
5.14.4 SPR interaction of Vap proteins in free solution with immobilized C-terminal His-tag VapA	255
5.14.5 Thermal Shift Assay interaction of <i>R.equi</i> Vap proteins	258
6.0 Chapter Six	260
6. 1 Final Thoughts and Reflections.....	260
6.1.1 Recombinant protein synthesis	260
6.1.2 Protein Purification Method Validation.....	261
6.1.3 Sample preparation and Optimization.....	262
6.1.4 Biophysical functional Assay	263
6.2 Conclusion.....	265
6.3 Recommendations for Future Work	266
References	268
Appendix	320
Number 1	320
Number 2	321
Number 3	325
Number 4	328
Number 5	332
Number 6	335
Number 7	338
Number 8	341

Abstract

Protein synthesis through bacterial overexpression using *Escherichia coli* can facilitate the production of high quantity and quality recombinant proteins for investigating biological function. The principal technique for studying protein function used in this study was surface plasmon resonance. Initially, the project was aligned to studying the human proteins AGTR1 and HMGCR, which are targets for coronary heart disease therapies, to create clinically relevant mutations associated with reduced efficacy. However, these proteins proved to be highly toxic to the host. As a result, two additional projects were initiated. The first to study the human chemokine CCL5 and to better characterize the low affinity binding domain for glycosaminoglycan. This would combine surface plasmon resonance with a novel trans-endothelial flow assay technique. The second was to investigate the interactions of the *Rhodococcus equi* virulence plasmid VapA protein and other related virulence proteins, again using surface plasmon resonance.

Wild type CCL5 and glycosaminoglycans binding domain mutants were cloned into the vector pMAL-p5X and successfully expressed and purified as a maltose binding fusion protein. The MBP tag could be removed using a protease but isolating the native protein from the tag proved problematic and potentially impacted on subsequent biological assays. Surface plasmon resonance showed that CCL5 binding interaction with the glycosaminoglycan heparin, was concentration dependent and that the recombinant wild type CCL5-MBP protein demonstrated the highest affinity for heparin. A novel assay to demonstrate the ability of CCL5 to promote mononuclear cell adherence to endothelial cells under flow condition demonstrated some positive initial data suggesting

that the low affinity glycosaminoglycan binding domain plays a significant under flow conditions.

In the second study a range of Vap proteins were successfully expressed and purified as “his-tag” fusion proteins. The use of surface plasmon resonance to study Vap protein – protein interaction demonstrated that VapA interacted with other members of the vap family. Interactions appear to be concentration dependent, and N or C terminal location of the His-tag did not appear to affect binding. However, VapA appears to rapidly degrade on the surface of a CM5 biosensor chip during surface plasmon resonance analysis meaning that an immobilised chip can only be used a single time making any further study very expensive. Protein Thermal shift assays were attempted to assess Vap protein interactions, but the hydrophobic nature of the Vap proteins meant that the standard dye is not suitable for this type of analysis.

In conclusion, protein expression workflows were successfully established and the use of surface plasmon resonance generated some interesting preliminary data. The surface plasmon resonance data also appeared to correlate well with the additional functional assays. Future work will refine these workflows to allow a better understanding of the proteins investigated within this study.

Declaration

This work in this thesis was performed by the author from May 2015- April 2020, under the supervision of Associate Professor. Noel Carter and Dr. Mark Carlile in the Faculty of Pharmacy Health and Wellbeing University of Sunderland. This thesis has been submitted for a degree of Doctor of Philosophy at the University of Sunderland and has not been submitted in full or part for another degree in any other institution.

This thesis may be made available for consultation, photocopying and for use through any other lending Libraries.

NAME: WISDOM JOHN NZUBE OKOYE

SIGNED:

DATE:

Dedication

This PhD thesis is dedicated to my parents Engr Madu & Mrs Edith Okoye for their unending support, prayer, love and believe in my dreams and to all who have lost their lives during the novel covid-19 pandemic crisis.

Acknowledgments

Firstly, my gratitude is to God almighty for the gift of life, strength, health, and knowledge, for without which I would not have had the ability and opportunity to undertake and complete this research PhD. My deepest gratitude goes to my indispensable director of studies Associate Professor. Noel Carter for his immeasurable encouragements and indescribable support, towards me as a person and through my PhD journey, I am certain without him, I would not have gotten this far. I thank in no small measure my second supervisor Dr. Mark Carlile, for all his timely interventions and extremely brilliant moral and technical support, I am immensely grateful sir. I also like to show appreciation to my first director of studies, Professor Roz Anderson of blessed memory, who accepted me as her PhD student in 2015. I acknowledge and remain grateful for the financial assistance of my parents and the faculty through Professor Tony Alabaster for covering vital expenses and waivers during my research. Special thanks to my darling wife Geraldine Orji for all her love and support and to my colleague turned friends and brother's Dr Chidi Ahamuefula, Dr Joseph Sunday, Mr Jude Akuta and the rest of my PhD colleague for their gift of friendship, shared motivation, and support. I sincerely thank staff members of my faculty for their shared support. It really takes a village to raise a child. Gratitude to the tireless technician in the university of Sunderland microbiology laboratory Dr Linda Henry and Kayleigh Ironside for their constant help and assistance around the lab. I acknowledge and appreciate the assistance of Dr Helen Waller from Newcastle University and Shweta and Claire Shepherd of GE Healthcare for all their help and support in getting me started with SPR technique. I am indebted to Mr Sam Erakewa for his prayers and to my family and siblings and relations, whose value for me only grows with age.

List of Abbreviations

Abbreviation	Meaning
%	Percent
μ	Micro
μg	Microgram
μL	Micro Litre
μM	Micromolar
<i>α</i> :	directional orientation of the system
ATP	Adenosine Triphosphate
BLAST	Basic Local alignment search tool.
BP	Base pair(s)
BSA	Bovine serum albumin
C: -	COOH terminal
CDS	Coding sequences
C-terminal	Carboxy-terminal
<i>D</i> :	Dimer
DNA	Deoxyribonucleic acid
<i>EC</i> :	Enzyme code
EDC	<i>N</i> -3-dimethylaminopropyl)- <i>N</i> '-ethylcarbodiimide hydrochloride
EDTA	Ethylene diamine tetraacetic acid, disodium salt
ESI	Electrostatic induction
ExpASY	Expert protein analysis system
<i>g</i>	Gram(s)
h	Hour(s)
HBS	HEPES buffer saline

HEPES	<i>N</i> -[2-Hydroxyethyl] piperazine- <i>N'</i> -[2-ethanesulphonic acid]
<i>Ig</i> :	Immunoglobulin
IL	Interleukin
IMAC	Immobilised metal affinity chromatography
INF- γ	Interferon-gamma
IPTG	Isopropyl- β -D-thiogalactopyranoside
<i>K</i> :	<i>Equilibrium</i> constant
<i>k</i> :	<i>Reaction</i> rate constant
K_a	Association constant
K_a	Association rate constant ($M^{-1}s^{-1}$)
<i>Kb</i>	Kilobase pair(s)
K_d	Dissociation rate constant (S^{-1})
<i>Kd</i>	Dissociation constant
K_D	Equilibrium dissociation constant
kDa	Kilodalton
L	Litre(s)
LC-MS/MS	Liquid chromatography Tandem mass production
mg	milligram
min	minute(s)
<i>N</i> -	NH ₂ terminal
nt	Nucleotide(s)
<i>N</i> -terminal	Amino terminal
°C	Degree Celsius
OD	Optical density
ORF	Open reading frame

<i>P</i>	Protomer or protein monomer
PAGE	Polyacrylamide gel electrophoresis
PAI	Pathogenicity Island
PBS	Phosphate buffered saline.
PBST	Phosphate buffered saline containing 0.05% Tween-80
PCR	Polymerase glycol
PEG	Polyethylene glycol
<i>R. equi</i>	<i>Rhodococcus equi</i>
RI	Bulk refractive index contribution in sample
Rmax	Analyte binding capacity of the surface (RU)
rpm	Revolution per min
<i>RT</i>	Ideal gas constant times absolute temperature
RT-PCR	Reverse transcriptase- polymerase chain reaction
SDM	Site directed mutagenesis.
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
sec	Second(s)
sp	Species
SPR	surface plasmon resonance
T7 <i>lac</i>	tris(hydroxymethyl)aminomethane
TAE	Tris-Acetate-EDTA
tc	Flow rate-independent part of mass transfer constant
Tm	melting temperature.
TSA	Thermal shift assay
UV	Ultraviolet

V	Volt(s)
v/v	Volume per volume
VAP	Virulence association protein
w/v	Weight per volume
WT	Wild type
<i>x g</i>	Times gravity (centrifugal force)
α	Alpha
β	Beta
Δ	Mutant

Chapter One

Introduction

1.0. Protein-Protein and Protein-Ligand interactions

Protein-protein or protein-substrate interactions (PPIs and PSIs) could also be referred to as molecular recognition (Du *et al.*, 2016). This area of study helps in understanding biochemical features and the manner through which biological macromolecules interact among themselves and with various other substrates. High specificity and binding affinity are often fundamental to the concept of a biological complex formation. PPIs and PSIs are also a key part of most cellular processes. The ability to detect and analyse these interacting proteins or substrates has aided the understanding of cellular machinery dynamics and therefore offered insight into one of the major challenges of modern biology (Rivas and Fontanillo, 2010). To perform their essential biological functions, proteins interface, interact, and binds to both them and other substrates in a controlled manner. This phenomenon can occur both naturally and engineered artificially to achieve a range of therapeutic potentials (Janin, 1995). Biomolecules can have an affinity for one another because of electrostatic forces, ranging from hydrogen bonding, hydrophobic binding forces, and dispersion forces (Viswanathan *et al.*, 2019). Most genes similarly encode proteins that bind to specific receptors to exert certain biological functions (Harrow *et al.*, 2009). The resultant non-covalent interaction has several biological significances, for example enzyme catalysis of chemical reactions, antibody-mediated neutralization of antigens, activation, and inhibition of chemical activities by hormones (Tummino and Copeland, 2008).

For these processes to occur, receptors and ligand interact by exchange of solutes and solvent molecules. The interaction of proteins with their ligands shares several unique properties, which are explained below concisely. Understanding the crucial role of these several unique properties (specificity and affinity) is essential in maximizing the concept of PPIs and PSIs studies (Kevin, 2018).

The biochemical events that constitute molecular PPIs and PSIs require physical contacts of high specificity to be established between two or more protein molecules. These molecular recognitions are facilitated by different kinds of bonding interactions. Additionally, proteins bind to each other via a combination of hydrophobic bonding, salt bridges, and van der Waals forces at specific binding domains on each protein (Phizicky and Fields, 1995; Golemis, 2002). These domains range from binding clefts or large surfaces of a few peptides long to a binding domain of hundreds of amino acids. The strength of the binding interaction behaviors is therefore influenced by the specificity, affinity of the binding molecule, and the size of the binding domain (Kastritis and Bonvin, 2013; Akiba *et al.*, 2019). Interestingly, our understanding of molecular recognition continues to advance in new ways to manipulate the PPIs and PSIs interfaces, towards enhancing drug design applications (Larsen *et al.*, 2005).

1.0.1 Specificity

Any biological receptor typically has an affinity and binds to its natural ligand. Sometimes this binding is irreversible and absolute, while other times, the binding is loose and reversible. For example, some natural toxins like the α -bungarotoxin affect the body by binding irreversibly to the acetylcholine receptor (Nirthanan *et al.*, 2016) representing the dynamic characteristics of a specificity. The specificity of a ligand can be monitored via a competitive binding assay where the amount of ligand that binds to a receptor is measured in the presence of other recognized ligands (Song *et al.*, 2016). In this case, if the receptor is highly specific to the target ligand, the presence of other ligands will not affect its binding affinity (Chen *et al.*, 2017).

1.0.2 Affinity

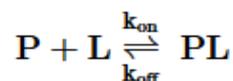
Molecules interact noncovalently with other molecules. For instance, proteins are found to stick to glass, partly due to polar surfaces interacting with one another (Georgakilas *et al.*, 2016). The surface of a cell is polar, primarily due to the extensive amount of carbohydrate that extends from membrane proteins and membrane lipids. Consequently, all proteins have some affinity for cell surfaces (Chen *et al.*, 2017). Receptor–ligand interactions are distinguished from other noncovalent interactions between molecules by their high affinity (Du *et al.*, 2016). It has been estimated that a protein on average interacts with about 3–10 other proteins (Bork *et al.*, 2004).

The relevant concepts that are critical for investigating molecular associations are binding kinetics, free energy, enthalpy, and entropy. Additionally, binding affinity

factors are known to influence the change in entropy – enthalpy, which helped rationalize protein-ligand and protein-protein binding interaction (Martin and Clements, 2013).

Binding via PPIs and PSIs are essentially based on the principle of electrostatic induction (ESI), perfectly illustrated as a mechanism of enzyme action through the lock and key theory, induced-fit theory, and conformational selection (Du *et al.*, 2016). The lock and key model postulated in 1894 by Emil Fischer and the most effective binding model based on the induced fit theory, captures the original rationale behind most PPIs and PSIs and studies. The induced-fit model of enzyme – substrate interaction is described by an interaction where the substrate can induce the proper alignment or conformational change on the binding site of the enzyme, causing the substrate interaction to perform a range of reactions of both catalytic and biological functions (Singh *et al.*, 2022).

The first significant step towards the many advances of protein-ligand interaction, in discovery, design, and drug development, is achieved when the mechanism responsible for the molecular recognition is understood (Steinbrecher and Labahn, 2010). Mathematically, it could be represented as:



Where PL means the protein-ligand complex and k_{on} and k_{off} represent the kinetic rate constants, which account for binding and unbinding reaction, respectively. At

equilibrium, the association binding reaction $P + L \rightarrow PL$ should be balanced by the dissociation unbinding reaction $PL \rightarrow P + L$, and this is represented as:

$$k_{\text{on}}[P][L] = k_{\text{off}}[PL]$$

Where the square brackets denote the equilibrium concentration of any molecular species. The binding constant K_b (in a unit of M^{-1}) is therefore defined by

$$K_b = \frac{k_{\text{on}}}{k_{\text{off}}} = \frac{[PL]}{[P][L]} = \frac{1}{K_d}$$

Where K_d is called the dissociation constant, therefore, the fast-binding rate complemented by a slow dissociation rate results in a high/low binding/dissociation constant and, hence, a high binding affinity (Xing *et al.*, 2016).

1.1.0 Protein Oligomerisation

Oligomeric proteins are proteins that consist of more than one subunit; meaning, they can form supramolecular structure by natural or artificial means (Garratt *et al.*, 2013). Oligomeric proteins are also made up of either several copies of identical polypeptides (homo-oligomers) or several copies of different polypeptides (hetero-oligomers). These homo- or hetero-oligomeric proteins form through covalent linkages either a nearly irreversible stabilized, or a reversible association of electrostatic, hydrophobic interactions and hydrogen bonds (Gotte and Libonati 2014). Protein oligomerization is the process through which functional tertiary protein dimers, trimers, tetramers and pentamers are formed from single unit

monomers, for example viral capsids. This phenomenon is crucial in the activation of one or multiple cascades of physiological pathways. Interestingly, Hashimoto and Panchenko (2010) believe the concept of protein-protein and protein-ligand recognition are interpreted best, by the study of homo- oligomers which in turn mediate and regulate the behaviour of binding receptors, and cell-cell adhesion processes (Dayhoff *et al.*, 2010).

1.2.0 Binding Interaction Techniques

To achieve the needed understanding of the potential novel effects of these protein-protein and protein-ligand binding interactions investigated under this research study, the use of a range of biophysical binding interaction techniques, are to be applied to this study. These techniques include thermal shift assays, surface plasmon resonance, and isothermal calorimetry (Du *et al.*, 2016). Usually, a variety of different techniques are necessary to validate, characterize, and confirm protein interactions. There are instances when proteins are unknown, their association with one or more protein may, help in identifying those unknown proteins. Protein interaction analysis may also unveil unique functional, and non-functional roles of known proteins. To discover or verify an interaction is the first step on the road to understanding where, how, and under what conditions these protein-protein interactions take place *in vivo* and the overall biology and functional implications of these interactions.

1.2.1 Surface Plasmon Resonance

The X100 Biacore surface plasmon resonance (SPR) is a specialized, reproducible high throughput microfluidic optical detection technique (Baron and Pauron, 2014). X100 SPR is used in real-time, to investigate, the binding interaction, specificity, concentration, kinetics, and affinity between non-labelled biomolecules and substrates. The two interacting molecules could be ions, viruses, proteins, chemical compounds, or genes, where one binding partner is referred to as the ligand and the other partner the analyte (Patching, 2014). A wide variety of binding analyses can be performed using the SPR technique such as protein-protein, protein-membrane (lipids), protein-carbohydrate, protein-nucleic acid, and even protein-small molecules (Baron and Pauron, 2014). The SPR instrument used in this thesis is a Biacore system X100, this instrument works by immobilizing a molecule often referred to as a “ligand” onto the surface of a sensor chip, after which the “analyte” molecule is flown over the immobilized ligand. Real-time intramolecular or analyte:ligand interactions on the surface of the sensor chip, generates a response, proportional to the ligand-analyte binding mass, and reflects sensitive variations to the nearest picograms per mm². More information on the principle of SPR is presented in section 1.2.2. The SPR Biacore sensor chips are made of about 50 nm gold-coated glass slide, upon which are dextran of various binding characteristics, depending on the type of sensor chips (Fee *et al.*, 2010). The gold coated slide acts as electrical conducting surfaces that generate resonance upon being hit by polarised light.

Electron charge density waves act by modulating the reflected light at a specific angle called resonance angle, relative to the ligand-analyte binding mass on the surface of the sensor chip (see Figure 1.0). Real-time intramolecular or analyte: ligand interactions on the surface of the sensor chip, generates a response, proportional to the ligand-analyte binding mass. All SPR interface occurs in conditions of absolute internal reflection, using resonance units (RU) to read all alterations in binding interactions.

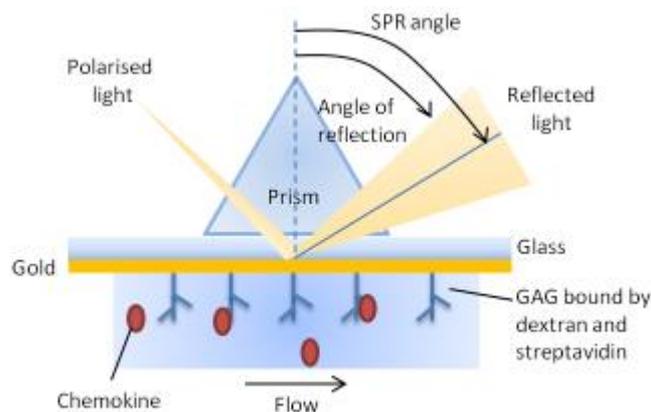


Figure 1.0: Surface Plasmon Resonance Schematics

Illustration of surface plasmon resonance. A substrate is immobilized to the gold-plated sensor surface using dextran. As ligand molecule (analyte) binds to the substrate, the refractive index shifts, and the SPR angle is altered. The movement is the signal monitored, and the change over time are represented as sensograms.

1.2.2 Principle of SPR

The principle of SPR is based on sample presentation through the flow cell of a microfluidic system at a small volume (20 – 60 nL), and low concentration (5 -30 µg/mL) to a glass slide coated with a thin gold layer on the surface known as the sensor chip. The gold surface of the sensor chip is covered with the dextran matrix acting as the substrate, to which molecules (ligand binding partners) can attach. These dextran matrixes provide the hydrophilic environment for the binding interaction. The analytes are after that flown over the surface of the matrix in a

continuous flow ($1-100 \mu\text{L} \cdot \text{min}^{-1}$) of solution (buffer) despite the characteristics of the molecule involved. The SPR applies the phenomenon of resonance detection to monitor biomolecular interaction, which causes a reduction in the intensity of light reflected at a specific angle from the glass compartment of the sensor chip (as described in figure 1.1). As the biomolecules bind to the SPR chip surface, the refractive index close to the sensor surface changes (within $\sim 300\text{nm}$). These changes alter the electron charge density waves and the angle of minimum intensity. The changes observed in the SPR refractive angle are proportional to the mass of the ligand molecule bound on the surface of the sensor chip.

All SPR interface occurs in conditions of absolute internal reflection, using resonance units (RU) to read all alterations and kinetics in binding interactions in real-time. The results are displayed and presented as a sensorgram, which is plotted as response signal/units (RUs) against time (seconds). The sensorgram can be evaluated using evaluation software and offer insight into the following parameters binding (Yes/No), specificity, affinity, kinetics, and sample concentration.

1.2.3 SPR Coupling Methods

There are different ways through which the ligand molecules are attached to the surface of the sensor chip. The most common procedure is by direct or covalent coupling, where the reactive groups on the ligand molecule are utilized for immobilization on the sensor chip surface (see Figure 1.2a), while the other procedure is by indirect or non-covalent coupling, in this case, the ligands are immobilized using an already covalently coupled molecule (see Figure 1.2b).

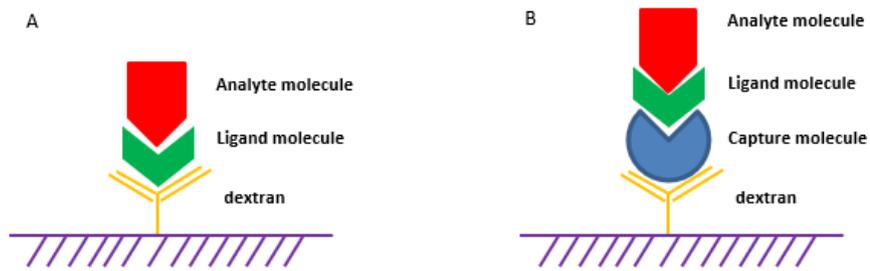


Figure 1.1: Illustration of most common Biacore SPR Immobilization or Capturing Procedure
 (a) Indicating a direct or covalent capturing procedure, (b) showing an indirect or non-covalent capturing procedure on the biosensor surface.

The covalent or direct immobilization fixes the ligand molecule permanently on the surface of the chip, often with a heterogeneous orientation. Interestingly, the capture-based (indirect) or non-covalent immobilization procedure is applied when a ligand molecule is difficult to regenerate, analyze, or are unstable in nature to the extent they are unable to bind directly. Other forms of immobilization procedure include hydrophobic attachment which is applied when immobilizing lysosome or lipid bilayers unto an SPR biosensor chip.



Figure 1.2: Picture of Biacore X100 Surface Plasmon Resonance (Live picture)

The BIAcore Surface Plasmon Resonance offers a range of general-purpose sensor chips applied towards achieving the different types of coupling chemistry described above. Examples of sensor chips applied during covalent coupling chemistry are the CM5, CM7, CM4, CM3, NTA and C1 BIAcore chip etc. The main differences among these sensor chips are the extent and peculiarity of the ligand type applicable to these sensor chip surfaces (example: lipids, his-tag, etc). The overall processes of immobilizing ligand molecules are essentially similar for all covalent coupling methods. The most common covalent coupling procedure is the amine coupling. The amine-coupling activates the carboxymethyl group of dextrans using the introduction of N-hydroxyl succinimide (NHS), which creates highly reactive succinimide esters to the surface matrix of the sensor chips by the modification of the carboxyl methyl groups, with the help of NHS and 1-ethyl-3-(3-dimethylamino propyl) carbodiimide (EDC). These esters interact spontaneously with the amines and other nucleophilic groups of the proteins to form covalent linkages with the matrix. The amine coupling reaction is illustrated in the diagram below (figure 1.4).

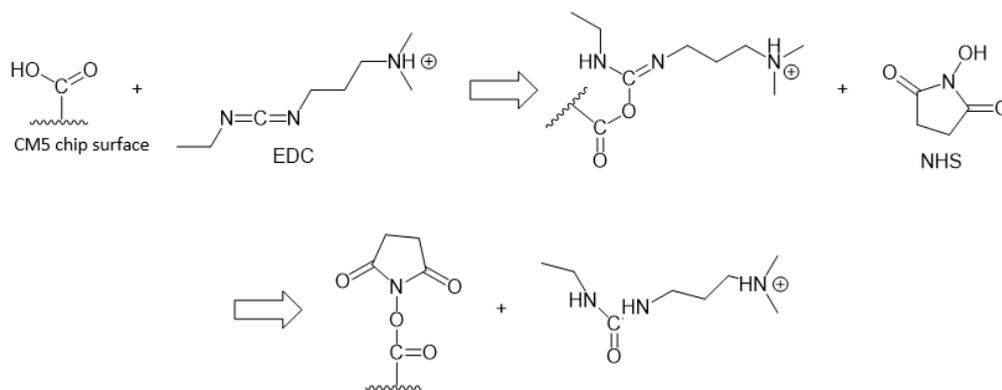


Figure 1.3: Chemical Reaction of Amine Coupling Equation (ChemDraw)

Further to the NHS and EDC activation of the sensor chip surface, during an amine coupling reaction. The proteins are injected in a pre-concentration buffer, which helps in achieving high protein concentration, thereby driving the coupling reaction. The final step is known as the blocking step, where ethanolamine is injected to block the remaining activated carboxymethyl group, this step elutes any non-covalent bound molecule and reduces the occurrence of non-reactive binding interactions. The final step is the regeneration of the immobilized sensor chip, by eluting any non-covalently bound analyte without necessarily interfering with the immobilized activities of the ligand.

1.2.4 Other Binding Interaction Strategies

To maximize the potential of the range of vital cellular activity in play between proteins, molecules, or substrates of interest. Attempts must be made to ensure that the tool employed to investigate binding interaction is not only useful but most ideal for the nature of the samples under investigation. Other methods presently available that confirm protein-protein interactions, are isothermal titration calorimetry (ITC), Forster resonance energy transfer (FRET) and luminescent oxygen channelling assay (LOCI), reflectometric interference spectroscopy (RIFS), surface plasmon resonance (SPR), circular dichroism spectroscopy (CD), thermal shift assay, Raman optical activity (ROA), nuclear magnetic resonance (NMR) and cryo-electron microscopy (cryo-EM) (Weiss, 2000; Graft *et al.*, 2009; Miura, 2018).

However, the main method of investigating binding protein interaction in this thesis is the surface plasmon resonance. Other techniques, such as protein affinity

chromatography, sedimentation, gel filtration columns, and fluorescence methods, can also be used to study these interactions. However, most of these alternative techniques require complex labelling steps and/or complex instrumentation. Additionally, most of these alternative techniques only give endpoint evaluation of yes / no binding, to binding affinity interaction experiments. Although the alternative methods briefly highlighted below, produce robust binding affinity data in previous recent studies (Ladbury, 2015; Duff and Howell, 2014; Layton and Hellinga, 2011, Choudhary *et al.*, 2017).

1.2.4.1 Thermal Shift Assay

Thermal stability shift analysis (TSSA) or thermal shift assay (TSA) is a method for examining binding interactions between PPIs and PSIs (Lo *et al.*, 2004, Matulis *et al.*, 2005). The technique utilizes thermodynamic models, to monitor PPIs and PLIs via evaluation of shift in thermal stability. Protein stability is determined here by the application of fluorescent dyes, example: SYPRO orange (Redhead *et al.*, 2017). The dye marks and reports protein denaturation during thermal heat, therefore, unfolding binding interactions. The TSA data are collected via a machine equally used for real-time qualitative thermocycler machine, in a rapid manner, using robust data application.

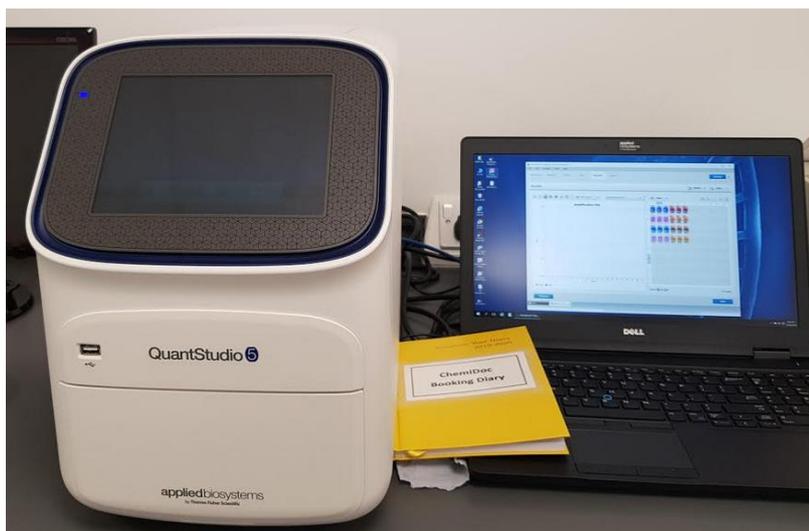


Figure 1.4: Picture of Thermal Shift Assay machine

TSA is a ubiquitous biophysical technique that determines the free energy of the binding ligand, by assessing protein stability on ligand dependence (Hubbard and Murray, 2011). The TSA technique has the advantage of measuring a wide range of PPI affinities and has been applied in drug discovery (Redhead *et al.*, 2015) and other structural studies like protein isolation and purification (Huynh and Partch, 2015). A new version of the TSA however is the fluorescent thermal shift assay (FTSA) (Redhead *et al.*, 2017). This technique is generally considered easier to develop (Jafari *et al.*, 2014) and can screen hundreds of compounds via a standard real-time quantitative polymerase chain reaction (RT-qPCR) and thermal cycler instruments (Huynh and Partch, 2015).

1.2.4.2 Isothermal Calorimetry

Isothermal Calorimetry (ITC) is a quantitative technique that can be defined as the measurement of heat either absorbed or released during a biomolecular binding event. It is often applied as a technique that provides a range of data concerning thermodynamics (process enthalpy) and kinetics (process rates) parameters.

Thermodynamics deals with the end states of the process during which interaction occurs, whereas kinetics deals with the nature of the interaction and the time rate at which it occurs.

In summary, protein-protein, or protein-ligand parameters, such as binding free energy, entropy, enthalpy, and binding affinity, can be generated via the application of ITC (Chaires, 2008; Du *et al.*, 2016). In a single ITC PPI experiment, the following set of data could be generated (i) affinity (K_D): strength of binding (ii) stoichiometry (n): rate of binding sites (iii) heat of binding (ΔH) and entropy (ΔS) and (iv) enzyme kinetics. The ITC technique validly measures the direct heat exchange of molecular interaction, during complex formation at a constant temperature, and has become a major technique in determining the forces that drive the binding process or intermolecular binding interactions (Perozzo *et al.*, 2004; Bronowskwa, 2011).

1.3.0 Protein Preparation for Binding Interaction Studies

Proteins are the main elements that facilitate most biological processes in a cell, and this would include gene expression, motility, proliferation, nutrient uptake, morphology, cell growth, intercellular signalling, and apoptosis. The availability of correctly expressed recombinant proteins facilitate functional studies such as protein self-interactions and PSI (Golemis, 2002; Wingfield, 2015). The phenotypic behavior of a cell depends on how proteins that are produced from gene expression interact with one another.

Chapters 3, 4 and 5 of this work focuses on investigating the protein-protein and protein- substrate interactions of selected proteins and molecules. Therefore, enough emphasis is laid here in the introduction, to understand the processes (stages), vital factors and potential limitation of obtaining recombinant proteins through protein synthesis. Being that only the production of the desired recombinant fusion-protein via protein synthesis could necessitate and progress towards other functional studies of these proteins via PPIs and PSIs. Interestingly, protein expression thrives under variable and different expression conditions hence the sections below attempt to systematically demonstrate the fundamental essence of protein biosynthesis, from gene cloning, protein overexpression, and purification. These insights are especially vital when trying to understand protein function both in a broader biological context and in the content of specific functional biological investigation of protein-protein or protein – substrate binding interaction.

1.4.0 Protein synthesis in prokaryotic and eukaryotic cells

In other to make proteins for biological and functional studies, protein molecules are created via a biological system known as protein synthesis. Protein synthesis involves amino acid synthesis, transcription, translation, and post-translational modification events depending on if a prokaryotic or eukaryotic cell is being used. *Esherichia coli* as a host is the most studied and utilised organism for protein synthesis towards the production of recombinant proteins (Rosano and Ceccarelli, 2014).

Due to the lack of membrane-enclosed nuclei in prokaryotic cells (bacteria), the process of DNA transcription and mRNA translation take place at the same time (Fan *et al.*, 2017). The field of recombinant protein expression requires a protein purification process, these two concepts has continued to evolve and advance over time towards the production of heterologous proteins as summarised in the workflow below.

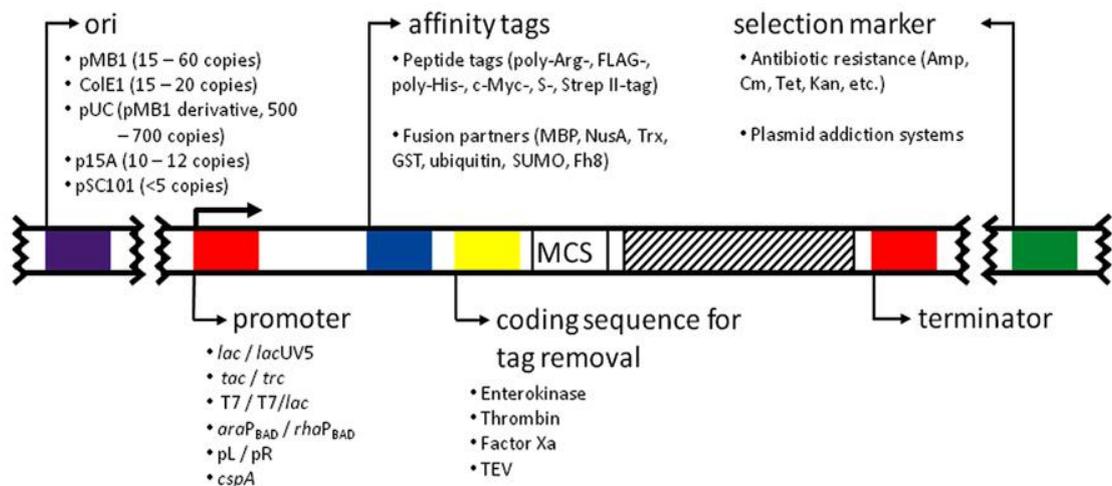


Figure 1.5: Anatomy of an expression Vector

The figure illustrates the major features present in common expression vectors as described in the text. The origin of replication (ori), affinity tags and coding sequences for their removal were positioned arbitrarily at the N-terminus for simplicity. MCS, multiple cloning sites. Striped patterned box: coding sequence for the desired protein and selection marker. (Rosano and Ceccarelli, 2014).

1.4.1 Promoters

Promoters control the binding of RNA polymerase and transcription factors. Promoter elements could be upstream (-24 to -31) or downstream (+28 to +34) of the transcription start site. Since the promoter region drives transcription of a target gene, it therefore determines the timing of gene expression and largely defines the amount of recombinant protein that will be produced. There are two types of promoters, those that are always active like the CMV, EF1A, and SV40

promoters, thus referred to as constitutive promoters and others which are only active under specific circumstances and commonly referred to as inducible promoter. An example of inducible promoter is Lac Z. The inducible promoters which can be switched from an OFF to an ON state, can be regulated by positive or negative control. Chemical agents, temperature, and light are all examples of factors that can trigger the induction of a promoter.

1.4.2 Positive inducible

In the OFF state, the promoter is inactive because the activator protein cannot bind. After an inducer binds to the activator protein, the activator protein can bind to the promoter, turning it ON and initiating transcription. For example, the chemically regulated promoters are among the most common inducible promoters. The positive inducible tetracycline ON (Tet-On) system, a versatile tool developed for use in prokaryotes and eukaryotes, works via direct activation. In this system, the activator rtTA (reverse tetracycline-controlled transactivator) is normally inactive and cannot bind the tetracycline response elements (TRE) in a promoter. Tetracycline and its derivatives serve as inducing agents to allow promoter activation.

1.4.3 Negative inducible

In the OFF state, the promoter is inactive because a bound repressor protein actively prevents transcription. Once an inducer binds the repressor protein, the repressor protein is removed from the DNA. With the repressor protein absent, transcription is turned ON. For example, pLac promoter is one of the most used prokaryotic promoters and is a negative inducible promoter. The pLac promoter

requires removal of the lac repressor (*lacI* protein) for transcription to be activated. In the presence of lactose or lactose analogy IPTG, the lac repressor undergoes a conformational change that removes it from lacO sites within the promoter and ceases repression of the target gene. A simplified lac inducible system is found in many bacterial expression vectors. Another negative inducible prokaryotic promoter is the pBAD promoters often used for bacteria purification.

A basal expression control for protein expression can be achieved in this instance by the introduction of a mutated promoter of the *lacI* gene, called *lacIQ*, that leads to higher levels of expression (almost 10-fold) of LacI (Calos, 1978). The lac promoter and its derivative lacUV5 are found to be weak and thus not very useful for recombinant protein production (Makoff and Oser, 1991). Synthetic hybrids such as the *tac* promoter demonstrate a combination of the optimised strength of other promoters and the advantages of the lac promoter. The *tac* promoter consists of the -35 region of the trp (tryptophan) promoter and the -10 region of the lac promoter and approximately 10 times stronger than lacUV5 (De Boer *et al.*, 1983). Commercial plasmids of these promoters that use the lac or *tac* promoters to drive protein expression are the pUC series (lacUV5 promoter, Thermo Scientific) and the pMAL series of vectors (*tac* promoter, NEB) which was utilised in this work.

Similarly, the T7 promoter system present in the pET vectors (pMB1 ori, medium copy number, Novagen) is extremely utilized for recombinant protein expression. In this system, the gene of interest is cloned behind a promoter recognized by the phage T7 RNA polymerase (T7 RNAP). This highly active polymerase is often provided in another plasmid or, most commonly, placed in the bacterial genome in

a prophage (λ DE3) encoding for the T7 RNAP under the transcriptional control of a *lacUV5* promoter (Studier and Moffatt, 1986). Thus, the system can be induced by lactose or its non-hydrolyzable analogy isopropyl β -D-1-thiogalactopyranoside (IPTG) and the basal expression can be controlled by *lacIQ* and also by T7 lysozyme co-expression (Moffatt and Studier, 1987). T7 lysozyme binds to T7 RNAP and inhibits transcription initiation from the T7 promoter (Stano and Patel, 2004). Therefore, in the event of a leaky expression due to production of small amount of T7 RNAP, the T7 lysozyme will effectively control unintended expression of heterologous genes placed under the T7 promoter. T7 lysozyme is provided by a compatible plasmid (pLysS or pLysE). Another level of control lies in the insertion of a *lacO* operator downstream of the T7 promoter, making a hybrid T7/*lac* promoter (Dubendorff and Studier, 1991). All three mechanisms (tight repression of the *lac* inducible T7 RNAP gene by *lacIQ*, T7 RNAP inhibition by T7 lysozyme and presence of a *lacO* operator after the T7 promoter) make the system ideal for avoiding basal expression.

1.4.4 Selection Marker

To deter the growth of plasmid-free cells, a resistance marker is added to the plasmid backbone. In the *E. coli* system, antibiotic resistance genes are habitually used for this purpose. Resistance to ampicillin is conferred by the *bla* gene whose product is a periplasmic enzyme that inactivates the β -lactam ring of β -lactam antibiotics. However, as the β -lactamase is continuously secreted, degradation of the antibiotic ensues and in a couple of hours, ampicillin is almost depleted (Korpimaki *et al.*, 2003). Under this situation, cells not carrying the plasmid are

observed to increase in number during cultivation. Tetracycline has been shown to be highly stable during cultivation (Korpimaki *et al.*, 2003), because resistance is based on active efflux of the antibiotic from resistant cells (Roberts, 1996). The cost of antibiotics and the dissemination of antibiotic resistance are major concerns in projects dealing with large-scale cultures.

1.4.0.3. Affinity Tags

When developing a project where a purified soluble recombinant protein is to be recovered (as is often the case), it is essential to incorporate means to (i) detect the protein along the expression and purification scheme, (ii) attain maximal solubility, and (iii) easily purify it from the *E. coli* cellular milieu. The expression of a section of amino acids (peptide tag) or a large polypeptide (fusion partner) in tandem with the desired protein to form a chimeric protein may allow these three goals to be straightforwardly reached (Rosano and Ceccarelli, 2014).

Being small, peptide affinity tags are less likely to interfere when fused to the protein of interest. However, in some cases they may trigger negative effects on the tertiary structure or biological activity of the fused chimeric protein (Chant *et al.*, 2005; Khan *et al.*, 2012). Common examples of small peptide tags are the poly-Arg-, FLAG-, poly-His-, c-Myc-, S-, and Strep II-tags (Terpe, 2003). Since commercial antibodies are available for all of them, the tagged recombinant protein can be detected by Western blot along expression trials, which is helpful when the levels of the desired proteins are not high enough to be detected by SDS-PAGE. Also, tags allow for one-step affinity purification, as resins that tightly and specifically bind the tags are available. For example, His-tagged proteins can be

recovered by immobilized metal ion affinity chromatography using Ni²⁺ or Co²⁺-loaded nitrilotriacetic acid-agarose resins (Bornhorst and Falke, 2000), while anti-FLAG affinity gels (Sigma-Aldrich) are used for capturing FLAG fusion proteins (Hopp *et al.*, 1988). On the other hand, adding a non-peptide fusion partner has the extra advantage of working as solubility enhancers (Hammarstrom *et al.*, 2002). The most popular fusion tags are the maltose-binding protein (MBP; Kapust and Waugh, 1999), thioredoxin (Trx; LaVallie *et al.*, 1993), glutathione S-transferase (GST; Smith and Johnson, 1988), ubiquitin (Baker, 1996) and SUMO (Butt *et al.*, 2005). The reasons why these fusion partners act as solubility enhancers are unclear and several hypotheses have been proposed (reviewed in Raran-Kurussi and Waugh, 2012). The MBP, has been shown to possess an intrinsic chaperone activity (Raran-Kurussi and Waugh, 2012). In a comparison studies, GST showed the poorest solubility enhancement capabilities (Hammarstrom *et al.*, 2006; Bird, 2011). The MBP, and Trx display the best solubility enhancing properties, but their large size may lead to the erroneous assessment of protein solubility after the recovery of recombinant protein (Costa *et al.*, 2013). Therefore, fusion tags of smaller sizes with strong solubility enhancing effects are desirable. Moreover, the recombinant proteins maintained their solubility after tag removal (Costa *et al.*, 2013). The MBP and GST later used in this work, can be used to purify the fused protein by affinity chromatography, as MBP binds to reduced amylose-agarose and GST to reduced glutathione-agarose. The MBP (NEB UK), is present in the pMAL series of vectors and GST (GE) is in the pGEX series. As demonstrated in the workflow particularly in chapter 4 and 5, peptide tag must be added to the fusion partner-containing protein if an affinity chromatography step is to be applied

during protein purification scheme. MBP and GST bind to their substrates non-covalently.

1.5.0 Cloning and Expression Vectors

An expression vector, also called an expression construct, is often a plasmid or a virus intended for gene expression in cells (Sharon and Kamen, 2017). In synthesizing recombinant protein, the protein-coding sequence is copied into an appropriate expression vector and then transferred into the host cell. In this case, the vector serves as a conveyor of specific genes to a target cell and can hijack the cell's metabolic mechanism for protein synthesis to manufacture the protein encoded by the gene. Expression vectors are standard tools in biotechnology and proteomics for heterologous protein expression. Because of their role in protein synthesis, the expression vectors must contain materials that facilitate gene expression (Rosano and Ceccarelli, 2014). These elements include a promoter, a translation initiation sequence like a ribosomal binding site (RBS), initiation and termination codon, and a transcription termination sequence (Porowińska *et al.*, 2013). In addition to these, an expression vector may also have a purification tag, which is usually added to the target protein (Norouzi, Hojati, and Badr, 2016). Finally, some vectors contain unique elements that are important for transformation and transfection. Both the cloning vector and the expression vectors are circular DNA constructs which are different from the chromosomes. However, an expression vector is more complicated. A cloning vector is a DNA molecule that carries and imported DNA into a host cell, replicates, and manufactures copies of the desired gene. Cloning vectors are engineered for self-

replication (Chen *et al.*, 2016), but they do not contain regulatory materials, unlike the expression vectors which contain both regulatory and self-replication elements (Zhang *et al.*, 2014).

1.6 Protein expression

Protein expression and purification play a principal role in biochemistry and protein studies (Jia and Jeon, 2016). The production of recombinant proteins is one of the most robust tools utilised in functional biology studies and a vital component of major discoveries in biologics and small molecule drug development (Sanchez-Garcia *et al.*, 2016). Target proteins undergo structural and mechanistic studies that generate essential data hence driving chemical designs (Hung and Chen, 2014). Following the establishment of recombinant gene/protein technology between the 1970 - 1980s methodologies and systems have evolved with the development of other modalities that enable high-throughput manufacturing of desired recombinant genes and their variants (Jia and Jeon, 2016).

In eukaryotic cells, protein synthesis occurs separately in a sequential manner, with transcription occurring in the nucleus and translation occurring in the cytoplasm (Chatterjee *et al.*, 2018). After translation into proteins, the polypeptides are altered in different ways to complete their structure, optimize their activity, and assign a location for them in a process referred to as post-translational modification (Duan and Walther, 2015). Post-translational modifications involve the various structural editing of the newly produced protein and are essential to their overall function. To be able to carry out biologic investigations, the scientist usually needs a process of manufacturing functional target proteins. Hence living

cells and their metabolic mechanisms can be exploited as workshops for the construction of desired proteins given the required building materials, also known as genetic templates (Dudley, Karim, and Jewett, 2014). Unlike proteins, DNA is less unstable to construct in vitro using effective recombinant DNA technologies. Hence, the DNA arrangements of genes can be constructed as templates for future protein expression (Jia and Jeon, 2016). These proteins manufactured from these templates are called recombinant proteins.

The DNA encoding a protein of interest is replicated downstream via an expression vector, which, when incorporated in a host cell, using the protein synthesis mechanisms of the host to produce the protein of interest. Technically, the expression of proteins can be complicated as it is affected by a lot of different biological factors. For example, the folding and structural orientation of a protein is tailored to its function and physiological role. Therefore, these proteins require post-translational modifications to perfect their complementarity and function (Audagnotto and Dal Peraro, 2017; Kuan, Bergamini and Weil, 2018). Also, some proteins might have biological activity that impacts negatively on their host (Martínez-Alonso *et al.*, 2010).

1.7.0 Protein expression systems

The production of recombinant proteins is a complex, multidisciplinary, and costly process. The demand for recombinant proteins for biological research remains on the rise. As well as the demand for novel and improved bioprocessing strategies that are cost-effective and timesaving. The continuous improvement in biopharmaceutical expression systems has led to the synthesis of quality products. Modern molecular biology techniques are at the forefront of the production of recombinant proteins using prokaryote or eukaryote expression systems. Various optimised techniques, such as systems biology, metabolic engineering, and CRISPR systems, are currently being applied for strain engineering to improve bioprocess performance and to synthesize biologically active and stable proteins. With the help of recombinant protein technology, expression of recombinant protein has successfully been achieved using bacteria (*E. coli*), mammalian cells, yeast, insect cells, transgenic plants, and transgenic animals (Ahmad *et al.*, 2014; Merlin *et al.*, 2014; Gupta S.K. *et al.*, 2019; Owczarek *et al.*, 2019). Although other hosts have peculiar advantages and limitations, this work especially focuses on synthesis of protein via bacterial (*Escherichia coli*) protein expression. Table 1.1 below shows a concise summary of the four different recombinant protein expression systems highlighted in this thesis, while identifying their advantages and disadvantages across a range of different protein overexpression parameters. Although other hosts have been mentioned, this work focuses on protein expression via *Escherichia coli*.

Table 1.1: Comparison of different recombinant protein expression systems (Ma *et al.*, 2003)

Parameters	Organism: Expression System			
Characteristics	<i>E. coli</i>	Yeast	Insect cells	Mammalian Cells
Cell growth	Rapid (30 Min)	Rapid (90 Min)	Slow (18 – 24 h)	Slow (24 h)
The complexity of growth medium	Minimum	Minimum	Complex	Complex
Cost of growth	Low	Low	High	High
Expression level	High	Low- High	Low-High	Low-moderate
Extracellular Expression	Secretion to periplasm	Secretion to mentions	Secretion to medium	Secretion to medium
Post-translational modification	No	Yes	Yes	Yes
Protein folding	Refolding usually required	Refolding may be required	Proper folding	Proper folding
N-linked glycosylation	None	High mannose	Simple no sialic acid	complex
O-linked glycosylation	No	Yes	Yes	Yes
Phosphorylation	No	Yes	Yes	Yes
Acetylation	No	Yes	Yes	Yes
Acylation	No	Yes	Yes	Yes
Gamma-carboxylation	No	No	No	Yes

1.8.0 *Escherichia coli* Expression System Advantages, and Limitations

The bacterial expression systems are often used in recombinant protein expression experiments, particularly in large scale protein synthesis. The rate of bacteria reproduction leads to high production yield (Baeshen *et al.*, 2015; Gupta and

Shukla, 2016) and overall low production cost. The bacterial expression system is efficient for most applications such as the generation of antibodies (De Marco, 2015; Beck and Liu, 2019), protein interaction research, and functional assays (Xu and Noyes, 2014; Caufield *et al.*, 2017). There are, however, concerns, eukaryotic proteins when expressed in bacteria, maybe non-functional because the cells do not have the mechanisms required for post-translational modification (PMT) (Sahdev, Khattar, and Saini, 2007; Jia and Jeon, 2016). Most *E. coli* expression systems are not compatible with N- and O- linked glycosylation, amidation, sulfation or palmitation, hydroxylation. These concerns are often encountered during protein synthesis of protein molecules greater than 50kDa (De Marco, 2009), intricate protein structures like several domains and integral membrane protein within the cell membrane, which require post-translational modifications such as disulphide bonds and glycosylation. PTMs characterised by glycosylation, disulfide bond formation, phosphorylation, or proteolytic processing are involved in folding processes, stability, and biological activity (Ferrer-Miralles *et al.*, 2009). *E. coli* has been optimised to allow for PTM towards production of recombinant proteins, as well as modified to produce glycosylated antibodies (Wacker *et al.*, 2002; Valderrama-Rincon *et al.*, 2012; Gupta and Shukla, 2016).

Similarly, *E. coli* has been engineered to allow glycosylation of proteins by transferring the N-glycosylation system of *Campylobacter jejuni* into it. However, further studies are required to establish it for the industrial production of commercial therapeutic proteins (Gupta and Shukla, 2017c).

Other limitations of the use of bacterial expression systems are the collection and formation of recombinant disulphide-bonded proteins in *E. coli*, often referred to as inclusion bodies (IB), elementary bodies, or cytoplasmic aggregates (Singh *et al.*, 2015). The presence of these aggregates actively prevents the formation of strong disulphide bonds in proteins (Rinas *et al.*, 2017). About 70%-80% of recombinant proteins expressed in *E. coli* are localized in cytoplasmic aggregates (Baumgarten *et al.*, 2018).

To mitigate the formation of inclusion bodies in *E. coli*, inducing protein expression by lowering the temperature after induction of culture leads to soluble protein. This strategy also increases protein stability and proper folding. Gupta and Shukla (2016) reported that the application of novel promoters and glycoengineering *E. coli* cells has also led to increased expression of a recombinant protein.

Protein purification is time-consuming as it involves two significant steps: extraction of the cytoplasmic aggregates from the bacteria and subsequent solubilization of the pure cytoplasmic aggregates (Norouzi, Hojati and Badr, 2016). A technique such as the incorporation of fusion tags (Liu M. *et al.*, 2019) to the sequence of the gene of interest, cofactor supplementation, and co-expression of the protein with molecular or chemical chaperones have shown to mitigate the concerns of IB formation (Gupta S. K. *et al.*, 2019) and improve soluble expression (Malekian *et al.*, 2019). Different tags such as Fh8, SUMO, His, TRX, and MBP at the N- or C-terminal enhance protein solubility and help in affinity purification (Paraskevopoulou and Falcone, 2018).

Another major limitation of the *E. coli* expression system is the occurrence of incomplete codon configuration known as codon bias or low usage codons. In the *E. coli* system, the absence of essential codon families has shown to cause problems during protein translation. This phenomenon is due to the under-representation of some prominent codon families that code for amino acid: Arg (AGA, AGG, and CGA), Ile (AUA), and Leu (CUA).

	Codon	Amino acid ²	% ³	Ratio ⁴	Codon	Amino acid	%	Ratio	Codon	Amino acid	%	Ratio	Codon	Amino acid	%	Ratio	
U	UUU	Phe (F)	1.9	0.51	UCU	Ser (S)	1.1	0.19	UAU	Tyr (Y)	1.6	0.59	UGU	Cys (C)	0.4	0.43	U C A G
	UUC	Phe (F)	1.8	0.49	UCC	Ser (S)	1.0	0.17	UAC	Tyr (Y)	1.4	0.47	UGC	Cys (C)	0.6	0.57	
	UUA	Leu (L)	1.0	0.11	UCA	Ser (S)	0.7	0.12	UAA	STOP	0.2	0.62	UGA	STOP	0.1	0.30	
	UUG	Leu (L)	1.1	0.11	UCG	Ser (S)	0.8	0.13	UAG	STOP	0.03	0.09	UGG	Trp (W)	1.4	1.00	
C	CUU	Leu (L)	1.0	0.10	CCU	Pro (P)	0.7	0.16	CAU	His (H)	1.2	0.52	CGU	Arg (R)	2.4	0.42	U C A G
	CUC	Leu (L)	0.9	0.10	CCC	Pro (P)	0.4	0.10	CAC	His (H)	1.1	0.48	CGC	Arg (R)	2.2	0.37	
	CUA	Leu (L)	0.3	0.03	CCA	Pro (P)	0.8	0.20	CAA	Gln (Q)	1.3	0.31	CGA	Arg (R)	0.3	0.05	
	CUG	Leu (L)	3.2	0.55	CCG	Pro (P)	2.4	0.55	CAG	Gln (Q)	2.9	0.69	CGG	Arg (R)	0.5	0.08	
A	AUU	Ile (I)	2.7	0.47	ACU	Thr (T)	1.2	0.21	AAU	Asn (N)	1.6	0.39	AGU	Ser (S)	0.7	0.13	U C A G
	AUC	Ile (I)	2.7	0.46	ACC	Thr (T)	2.4	0.43	AAC	Asn (N)	2.6	0.61	AGC	Ser (S)	1.5	0.27	
	AUA	Ile (I)	0.4	0.07	ACA	Thr (T)	0.1	0.30	AAA	Lys (K)	3.8	0.76	AGA	Arg (R)	0.2	0.04	
	AUG	Met (M)	2.6	1.00	ACG	Thr (T)	1.3	0.23	AAG	Lys (K)	1.2	0.24	AGG	Arg (R)	0.2	0.03	
G	GUU	Val (V)	2.0	0.29	GCU	Ala (A)	1.8	0.19	GAU	Asp (D)	3.3	0.59	GGU	Gly (G)	2.8	0.38	U C A G
	GUC	Val (V)	1.4	0.20	GCC	Ala (A)	2.3	0.25	GAC	Asp (D)	2.3	0.41	GGC	Gly (G)	3.0	0.40	
	GUA	Val (V)	1.2	0.17	GCA	Ala (A)	2.1	0.22	GAA	Glu (E)	4.4	0.70	GGA	Gly (G)	0.7	0.09	
	GUG	Val (V)	2.4	0.34	GCG	Ala (A)	3.2	0.34	GAG	Glu (E)	1.9	0.30	GGG	Gly (G)	0.9	0.13	

Figure 1.5: Codon frequency for the *Escherichia coli* genome system:

Represents (%) average codon usage per 100 codons and (Ratio) represents codon abundance relative to all the codons per amino acid. Adapted from Arabidopsis Research companion <http://weed/mgh.harvard.edu> (Maloy *et al.*, 1996).

The Figure 1.6 designates that these amino acids and their codons are underrepresented to about 8% less, as against other codon families. A scenario where 61 tri-nucleotide codons encode 20 amino acids, may have led to a varying degree of preference for some selected codons whose essential function of amino acid translation into proteins may have been limited.

Interestingly, recent *E. coli* expression cells are modified to incorporate these rare codons that support appropriate protein expression, folding, required post-translational modifications to some extent. The presence of rare codons can be addressed by using codon optimization or host modification. Codon optimization increases the expression of recombinant protein by many folds (Rosano and Ceccarelli, 2014; Gupta S. K. *et al.*, 2019; Rosano *et al.*, 2019). A few examples of these rare codons optimized *E. coli* cells are BL21 Codon plus cells which are genetically engineered competent cells derived from BL21, that contain the lambda DE3 prophage which contains the gene for T7 RNA polymerase under the control of a lacUV5 promoter. The BL21 Codon Plus contain rare codons like AGA, AGG, AUA, and CUA but do not have the recA1. The C41 (DE3) and C43 (DE3) Competent cells were genetically engineered from BL21 (DE3) effective *E. coli* cell expression systems for overexpressing toxic genes cloned into any T7 vector and membrane proteins from both prokaryotes, eukaryotes, and viral proteins. Additionally, a redox environment and foldases for example, disulfide isomerases (Dsb proteins) and peptidyl-prolyl isomerases (PPIase) are useful in the formation of the correct disulfide bond in the periplasm (Gupta and Shukla, 2017b), which ultimately reducing the concerns of codon bias. The expression of genes with rare codons often results in low expression and results to premature termination of the protein synthesis (Owczarek *et al.*, 2019). Table 1.2 below, shows some of the newly developed *E. coli* recombinant protein expression strains, alongside their key features. These strains are mostly derivatives of BL21 (DE3), whose genome have been manipulated towards maximizing heterologous protein production.

Table 1.2 *Escherichia coli* Strains Frequently Used for Recombinant Protein Production (Adoped from German *et al.*, 2019)

<i>Escherichia coli</i> Strains	Application/key features
OverExpress™ C41(DE3) and C43(DE3)	Production of membrane/toxic proteins/T7RNAP gene under control of <i>p_{lac}Weak</i> promoter. In addition, C43 (DE3) has a mutated <i>Lacl</i> repressor that binds longer to the <i>lac</i> operator, thus delaying T7RNAP transcription. It is also a Lon protease revertant
ArcticExpress (DE3)	Production of aggregation-prone proteins/constitutive expression of chaperones Cpn10 and Cpn60 from the psychrophilic bacterium <i>Oleispira antarctica</i> , which show high refolding activities at 4–12°C
SixPack	Codon bias correction/insertion of extra copies of genes encoding six rare tRNAs in a ribosomal RNA operon
Mutant56(DE3)	Production of membrane/toxic proteins/mutant T7RNAP with weakened binding to the T7 promoter
TatExpress BL21	Enhanced protein secretion/ <i>ptac</i> promoter upstream of <i>tatABCD operon for increased levels of Tat secretion pathway</i>
Marionette	Protein coexpression/independent control of expression using 12 different inducers
RiboTite	Control of basal expression and tunability of protein production levels/introduction of riboswitches upstream of the T7RNAP gene and the coding sequence of interest

Another essential benefit for high-throughput protein expression and purification is the addition of a fusion tag at either the N- or C-terminus of recombinant proteins. For optimal result, a fusion tag must fulfil these conditions: they must enable (i) easy detection of protein expression, (ii) high yield of protein expression and solubility, and (iii) easy isolation of highly pure proteins from most prokaryotic cells example *E. coli* (Kimple and Sondek, 2004). Presently, a broad range of tags are available (Zhao *et al.*,

2013; Costa *et al.*, 2014), the general characteristics of these frequently used tags are listed in table 1.3.

Table 1.3: An overview of fusion tags, availed for recombinant protein overexpression Adopted from: (Jia and Jeon, 2016).

Tag	Length/size (kDa)	Matrix/elution	Typical uses	Reference
His - tag	2-10, typically 6/ (0.84)	divalent metal ion (Ni, Co, Cu, Zn)/imidazole or low pH	purification and detection	Gaberc-Porekar <i>et al.</i> , 2001
FLAG	8/ (1.0)	FLAG antibody/low pH, EDTA or FLAG peptide	purification and detection	Schmidt <i>et al.</i> , 2012
Strep-II	8/ (1.1)	Strep-Tactin/biotin or desthiobiotin	purification and detection	Schmidt and Skerra, 2007
Fh8	69/ (8.0)	Ca ²⁺ -dépendent hydrophobic interaction/ EDTA	purification increased solubility and expression	Costa <i>et al.</i> , 2014
Trx	109/(11.7)	Phenylarsinine oxide/ thiol-containing reducing agents	purification and increased solubility	Dyson <i>et al.</i> , 2004
SUMO	100/(12.0)	His-affinity must be added to	increased solubility and expression	Zuo <i>et al.</i> , 2005
GST	211/ (26.0)	Glutathione/ reduced glutathione	purification, detection and increased expression and solubility	Kimple and Sondek, 2004
GFP	238/ (26.9)		detection increased solubility and expression	Hammon <i>et al.</i> , 2009
Halo Tag	312/ (34.0)	Chloro-alkane/Halo Tag buffer and TEV protease	purification increased solubility and expression	Kimple and Sondek, 2004
MBP	396/ (42.0)	Cross-linked amylose/ maltose	purification, detection, increased expression and solubility	Zhao <i>et al.</i> , 2013

1.9.0 Regulation of Gene Expression

The regulation of gene expression allows for the modulation of any of the gene expression processes, ranging from DNA-RNA transcription to post-translation modification. From a broader perspective, protein overexpression is regulated through different mechanisms, depending on which expression host systems in use either prokaryotic or eukaryotic (Rosano and Ceccarelli, 2014). Regulation of gene expression is vital for viruses, prokaryotes, and eukaryotes, as it enhances the

adaptability and versatility of an organism to its cellular environment, which allows the cells to express proteins only when needed. With the focus of this thesis, being on overexpression of a range of different proteins, using *E. coli* expression system (prokaryotic), regulation of gene expression is therefore presented to focus on the lactose dependent metabolism in *E. coli*, which is regulated by an inducible system.

1.9.1 The *lac* operon

In bacteria, gene expression is regulated through the robust activity of the *lac* operon (Klug *et al.*, 2016). The *lac* operon, which are made up of three main components; (i) the regulatory genes (activators and repressors), (ii) regulatory genes (promoters and operators) and 3: structural genes (encoding enzymes). Also, the operon is controlled via two main types of gene expression regulation mechanism, which are the negative and positive control (Hardison and Chu, 2019). Incidentally, gene overexpression is found to be repressed in the absence of lactose and the presence of glucose and induced in the presence of lactose and absence of glucose (Hardison and Chu, 2019).

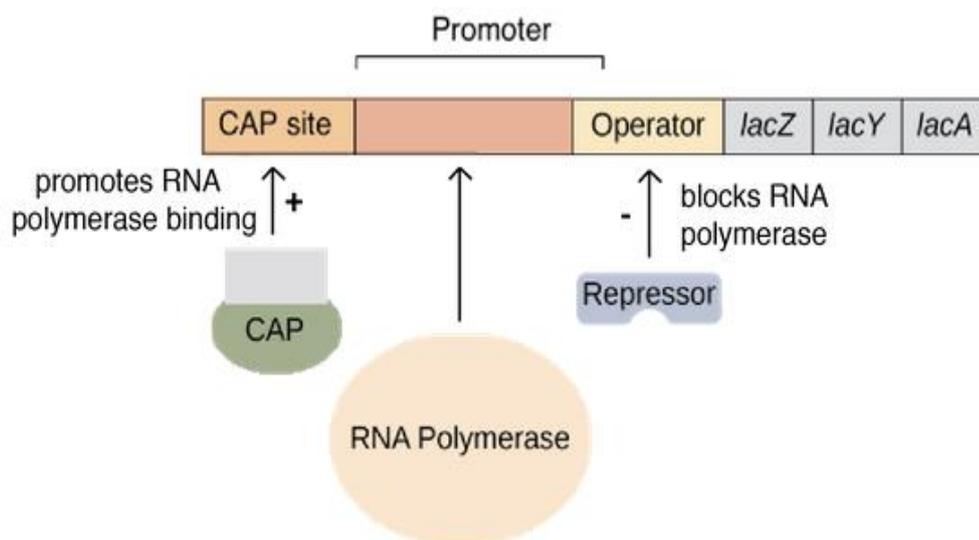


Figure 1.6: Schematics of Prokaryotic gene regulation of the *lac* operon (OpenStax College, Biology)

Furthermore, biological processes like cell differentiation and morphogenesis are driven by the regulation of gene expression (Chen *et al.*, 2013). The lactose serves as both an inducer, and energy source in glycolysis for bacteria gene regulation, constituting inducible enzymes, that are responsible for lactose metabolism. The amino acid sequence of these enzymes *β-galactosidase*, *permease*, and *transacetylase* are encoded by three structural genes *lacZ*, *lacY*, and *lacA*, respectively (Santillán and Mackey, 2008). The functions of these enzymes have been well studied and presented in several kinds of literature. Other prokaryotic gene regulation machinery includes the tryptophan operon and the arabinose operon.

1.9.2 The *lac* repressor

The gene that encodes for *lac* repressor protein is the *I* gene. The *lac* repressor is a protein that represses (inhibits) the transcription activity of the *lac* operon. Presented a mechanism of regulation (of/on). It does this by binding to the operator, which partially overlaps with the promoter. When the operator is bound, the *lac* repressor forms a blockage on the RNA polymerase's way and prevents it from transcribing the operon.

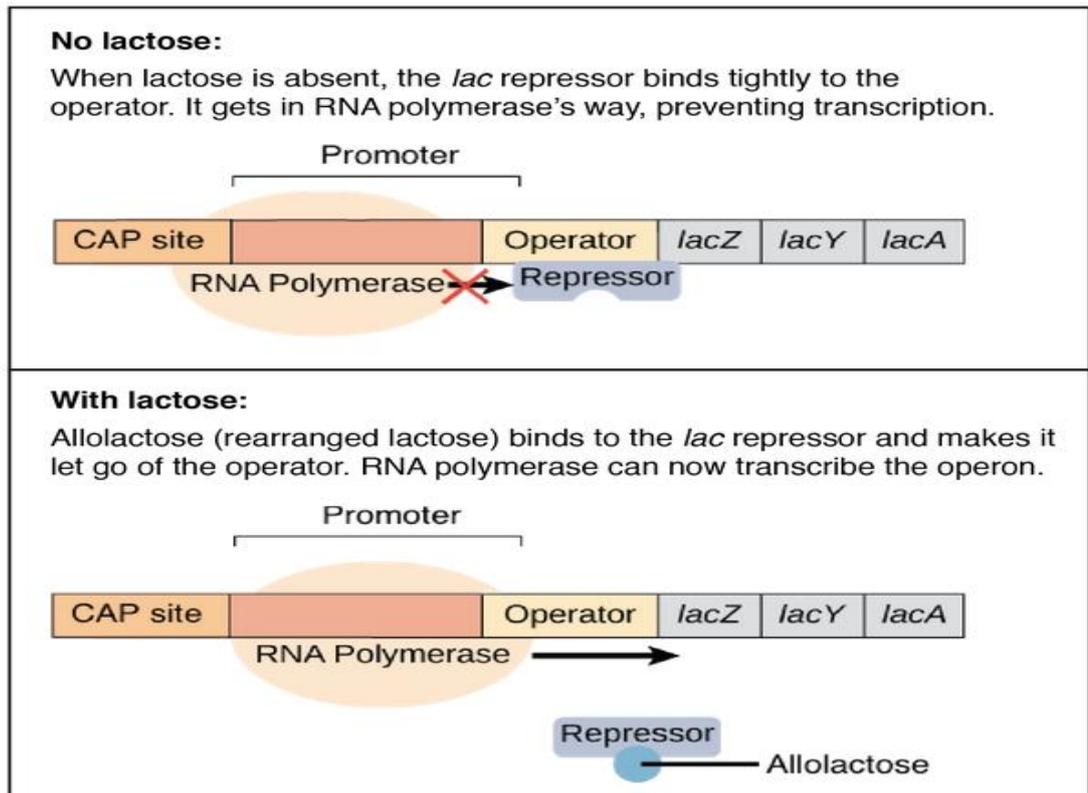


Figure 1.7: Schematics of Prokaryotic gene regulation of the *lac* repressor (OpenStax College, Biology)

When lactose is not available, the *lac* repressor binds tightly to the operator, preventing transcription by RNA polymerase. However, when lactose is present, the *lac* repressor loses its ability to bind DNA. It floats off the operator, clearing the way for RNA polymerase to transcribe the operon as demonstrated in figure.

This change in the *lac* repressor is caused by the small molecule allolactose, an isomer of lactose. When lactose is available, some molecules will be converted to allolactose inside the cell. Allolactose binds to the *lac* repressor and makes it change shape so it can no longer bind DNA. Allolactose is an example of an inducer, a small molecule that triggers expression of a gene or operon. The *lac* operon is considered an inducible operon because it is usually turned off (repressed) but can be turned on in the presence of the inducer allolactose.

1.9.3 Catabolite activator protein (CAP)

With the RNA polymerase only being unable to bind tightly to the lac operon promoter, the catabolite activator protein (CAP), supplements the RNA polymerase activity in the lac operon, by binding to the region of the DNA just before the lac operon promoter, thereby enabling a tighter binding of the RNA polymerase to the promoter. This results to a higher level of transcription. This is demonstrated in figure 1.8.

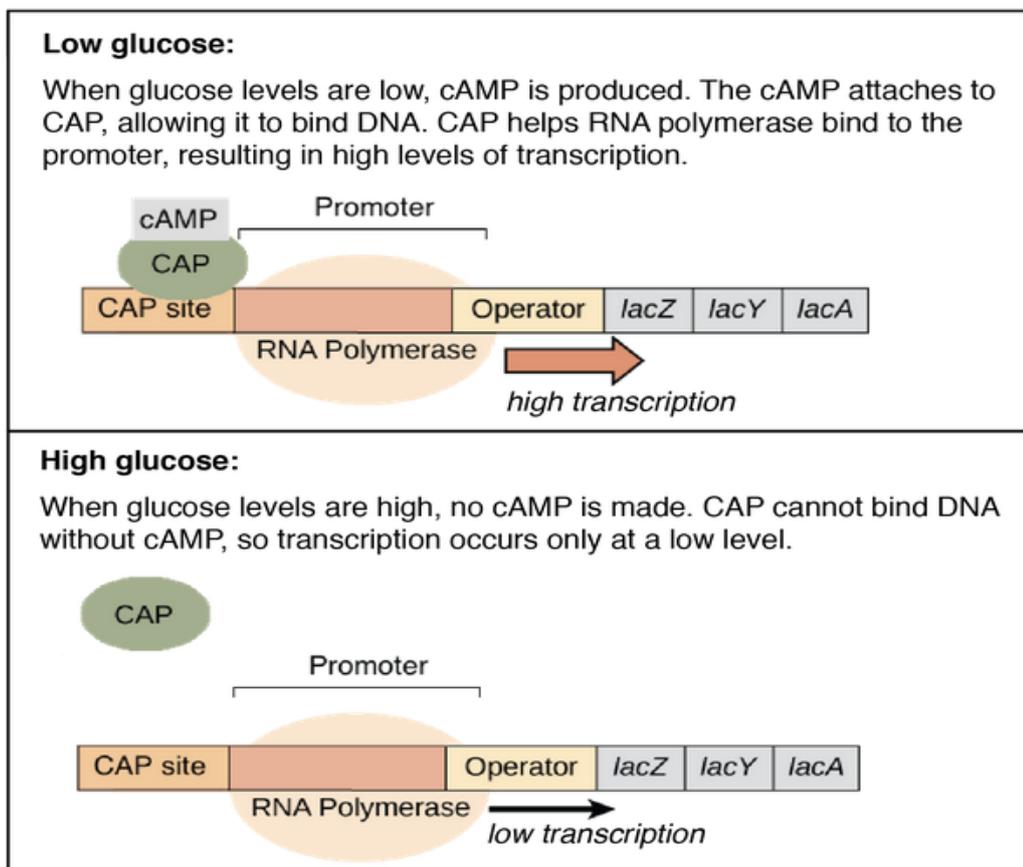


Figure 1.8: Schematics of Prokaryotic gene regulation of the Catabolite activator protein (OpenStax College, Biology)

As the CAP is not always active (able to bind DNA), it is however regulated by a cyclic AMP (cAMP) molecule. The cAMP is a "hunger signal" made by *E. coli* when glucose levels are low. The cAMP binds to CAP, changing its conformational shape and making it able to bind DNA and promote transcription. Without cAMP, the CAP

cannot bind DNA hence will remain inactive. In other for CAP to be active, the glucose levels must be low (cAMP levels are high). Thus, the lac operon can only be transcribed at high levels when glucose is absent. This strategy ensures that bacteria only turn on the lac operon and start using lactose after they have used up all of the preferred energy source (glucose).

The success of protein expression using the *lac* operon in *E. coli* is largely dependent on the regulation of its major component. The table 1.4 below, summarises the entire events of the lac operon and by extension a successful account of recombinant protein synthesis which is vital for the research carried out in this work.

Table 1.4: A Summary of all gene regulation scenarios using the *lac operon* in *E. coli* protein expression

Glucose	Lactose	CAP binds	Repressor binds	Level of transcription
Positive	Negative	Negative	Positive	No transcription
Positive	Positive	Negative	Negative	Low-level transcription
Negative	Negative	Positive	Positive	No transcription
Negative	Positive	Positive	Negative	Strong transcription

Research Questions (Aims of the Study)

To address the current lack of evidence that may allow GPs to choose heart disease treatments based on an individual patient's genetic profile, by understanding the individual binding kinetic affinity of drug classes presently used in the treatment of heart disease and their activity with interacting proteins at the drug site of therapeutic action.

The original aim was to measure the binding affinity of common therapeutics for treating coronary heart disease against their protein targets and to create mutations in these proteins that corresponded to known single nucleotide polymorphisms that had been shown to affect therapeutic efficacy.

Unfortunately, there was a lack of progress in this project due to issues around toxicity of the proteins being expressed and the death of the original Director of studies the late Prof. Roz Anderson who had originally developed the project. This led to the broadening of this research to develop methodologies within the research areas of the new supervisory team. The result of this work can be found in Chapter 3 where the work concentrated on the analysis of two CHD proteins, AGTR1 and HMGCR.

This similarly resulted to a change of focus of the original aim based on the rationale of developing methods within the faculty to address the study of the biophysical and biological properties of recombinant proteins expressed in *E. coli*. This resulted in two distinct projects with a common theme around using

recombinant proteins to study protein: ligand and protein: protein interactions, however with two new aims as follows.

1. To Investigate the biophysical and biological properties of the heparin binding domains in the chemokine CCL5
2. To investigate protein-protein interactions of *Rhodococcus Equi* Virulence-Associated Proteins (VAP).

CHAPTER TWO

Materials and Methods

2. Materials and methods

2.1.0 Chemical and Reagents

All chemicals and reagents used were of analytical grade standard hence were not purified any further. Solutions and buffers were prepared using deionised water and stored at room temperature unless otherwise stated. All PCR primers and out-door synthesized DNA plasmids were supplied by Integrated DNA Technology UK, Source Bioscience UK was used for DNA sequencing.

2.1.1 Media

All liquid media were prepared using deionised water, sterilized by autoclaving, and stored at room temperature except for solid media that were stored at 4°C.

2.1.2 Solid Media

Solid agar media used include LB agar, 2YT agar, and TB agar, and the media were augmented with 2% Agar (Bacteriological No.2) and autoclaved. For 2YT agar media, a mixture of 16g of tryptone powder, 10g of Yeast powder extract and 5g of NaCl (Sigma Aldrich UK), for LB agar media, a mixture of 15g of tryptone powder, 5g of Yeast powder extract and 5g of NaCl (Sigma Aldrich UK), for TB agar media 12g of tryptone powder, 24g of yeast powder extract and 4ml of sterile glycerol were dissolve and adjusted to 1000mL of double-distilled water and autoclaved. A solution of sterile and filtered 100ml of 0.17M KH_2PO_4 and 0.72M K_2HPO_4 were added and accounted for to TB agar media final volume upon cooling. At 50 °C, the agar broth was aseptically poured into designated Petri dishes and allowed to cool at room temperature.

2.1.3 Liquid Media (Broth)

Liquid media (broth) used in the study include Luria-Bertani (LB), Nutrient Broth (NB), and 2 times Yeast Extract (2YT). Preparation is similar to the same as in solid broth except for no agar where added. All liquid broth were stored at 4°C.

2.1.4 Selective Media

Here, liquid (broth) and solid media were prepared as normal. After autoclaving, each medium could cool at about 50°C and the sterile-filtered antibiotic stock solution was added to the concentration of choice and gently shaken. Antibiotics used were ampicillin (100µg/mL).

2.1.5 Bacterial Strains, Plasmids and Cloned Genes

Cloned genes used include, AGTR1, HMGCR, Wild-type-CCL5, ⁴⁴BBXB⁴⁷ mutated CCL5 region (CCL5-Mut-LOW), ⁵⁵BBXB⁵⁹ mutated CCL5 region (CCL5-Mut-HIGH), both regions mutated CCL5 (CCL5-Mut-BOTH), and the R VAP A, D, E, G, H genes. Bacterial strains used include *Escherichia Coli* Subcloning Efficiency DH5α Competent Cells (ThermoFisher Scientific UK), BL21 (DE3) Competent *Escherichia Coli* cells (New England Biolabs UK) and OverExpress™ C43 (DE3) Chemically Competent *Escherichia Coli* cells (Sigma-Aldrich UK). (Hayashi *et al.*, 2007)

Competent cells	<i>E. coli</i> genotypes
OverExpress(tm)C43(DE3)	F ⁻ ompT gal dcm hsdS _B (r _B ⁻ m _B ⁻)(DE3)
BL21(DE3)	<i>E. coli</i> str. B F ⁻ ompT gal dcm lon hsdS _B (r _B ⁻ m _B ⁻) λ(DE3 [<i>lacI lacUV5-T7p07 ind1 sam7 nin5</i>]) [<i>malB</i> ⁺] _{K-12} (λ ^S)
DH5α	F ⁻ endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG purB20 φ80dlacZΔM15 Δ(<i>lacZYA-argF</i>)U169, hsdR17(r _K ⁻ m _K ⁺), λ ⁻

E. coli DH5α competent cell was used as a host for cloning while *E. coli* BL21 (DE3) and C43 (DE3) competent cells were used for the recombinant gene expression.

Bacterial strains were stored and maintained at -80°C while plasmids were stored at -20°C. The vector plasmids used include pGEX-6p-1, pMAL-c5x, pMAL-p5x, and pET32a. The relevant features of these vector plasmids have been briefly described.

2.2.0 DNA Methods

2.2.1 Extraction of Plasmid DNA

Plasmid DNA was extracted using the plasmid miniprep kit (Thermo fishers Scientific USA) gel extraction following the manufacturer's protocol. Plasmid DNA concentrations were quantified using a NanoDrop 1000 spectrophotometer following the manufacturer's protocol after calibration using a blank (control).

2.2.2 Plasmid Maps

Plasmid maps were constructed using PlasMapper, a web-based plasmid map drawing software designed by Dong *et al.*, (2004). CDS templates of genes were infused into the cloning restriction digest site of the plasmid sequence and fed into the program which generates the plasmid-DNA map. The software was available at <http://wishart.biology.ualberta.ca/PlasMapper/>.

2.2.3 Electrophoretic analyses of DNA

Agarose gel casting was prepared by boiling 1g of dry agarose powder in 100mL of TAE-buffer (Tris-Acetic acid-EDTA) to obtain 1% agarose gel. A comb was inserted into the electrophoresis tray, and the molten agarose, which contains 10µL of Gel Red, was poured into the casting chamber and allowed to set. Upon setting, the casted agarose was placed in the electrophoresis chamber and filled with TAE

buffer up to the designated mark in the tank. The comb was carefully withdrawn and DNA sample mixture which contains DNA loading buffer loaded on each well. The agarose gel was run at a constant current of 120mA for 30 min. Agarose images were resolved under UV light using Bio-Rad ChemiDoc imaging machine and software.

2.2.4 Restriction Analysis

The double digest restriction method was applied to all plasmid-DNA, respectively, using *Bam*HI and *Eco*RI restriction endonuclease enzymes (Fermentas Thermo Fisher Scientific USA). This was done to generate cohesive ends in a 10 times fast digest buffer (Fermentas Thermo Fisher Scientific USA) which supports maximal activity for digestion using two enzymes. The concentration of the plasmid-DNA (8.6ng/μL), *Bam*HI (38.6ng/μL), and *Eco*RI (38.5 ng/μL) was recorded as Restriction digest mixtures as shown below were incubated for 15min at 37°C. Restriction digest analysis was resolved by agarose gel electrophoresis at a constant current of 120mA for 30 min.

Table: 2.0 Restriction Digest Mixtures

Restriction enzymes digest mixture	
Items	Volume (μl)
Plasmid –DNA	30
10 x Fast Digest Buffer	12
<i>Bam</i> H1	4
<i>Eco</i> R1	4
PCR-water	50
Total	100

2.2.5 Cloning of amplified DNA in Vector through Ligation

The genes were cloned into the BamH1 and EcoR1 cloning sites of vectors (pGEX-6p-1, pMAL-c5x, pMAL-p5X, and pET32a) MCS as shown in DNA-vector maps. Purified human AGTR1, HMGCR, CCL5 gene inserts were separately digested with *Bam*H1 and *Eco*R1 restriction enzymes to generate cohesive ends in the NE buffer 3, that supports optimal activity for double enzymes digestion see table. The restriction digestion mixture is shown in table 4.0. Each digestion mixture was incubated at 37°C for 20mins (now linearized) and the DNA inserts were gel purified using QIAGEN mini-prep protocol after that. DNA inserts were ligated using a sticky end ligation method to the linearized plasmid using the help of DNA ligase. The ligation reaction mixture is shown in table 4.1. The ligation reaction products were resolved by agarose gel electrophoresis and gel purified. An online tool known as ligation calculator (http://www.insilico.uni-duesseldorf.de/Lig_Input.html) was used to calculate the amount of insert to be used at a 3:1 insert to vector ratio, and this was derived by factoring the size of the insert (bp), size of the vector (bp) and concentration of the vector (ng), which was deduced using a table-top nano-drop spectrometer. The T4 DNA ligase reagents are from (Thermo Fishers Scientific).

Table 2.1 Sticky End Ligation Mixture

Ligation Mixture	
Items	Volume (μL)
Vector	1
5x Ligase buffer	1.4
Ligase buffer	4
PEG 4000	1
Distilled- Water	1.5
Total	20

2.2.6 VAP Protein Plasmid-DNA Constructs

All *Rhodococcus equi* virulence-associated protein (Vap) DNA-plasmid clones, were received as a kind gift from Dr. Lynn Dover of Newcastle University, under adequate refrigerated conditions. The *R. equi* vap clones were transformed firstly into *Escherichia Coli* DH5α and later re-transformed into *Escherichia Coli* C43 (DE3) in the methods described below in section (4.2.7).

2.2.7 Transformation of chemically competent *Escherichia Coli* DH5α

An *E. coli* DH5α strain (New England Biolabs UK) was used as a first-line cloning strain. Five microliters of suspected ligated DNA constructs were added to a 50μL aliquot of freshly thawed stock of DH5α, chemically competent *E. coli* and incubated on ice for 30 min. It was heat-shocked at 42°C for 45 seconds and transferred immediately to the ice for 2 min to recover. A 2YT medium (500μL) was added to the solution and incubated at 35°C for 1 hour. A 200 μL and 50μL sample of suspension (containing transformed cells) were respectively spread on a 100μg/mL ampicillin-resistant LB agar plate and incubated at 37°C overnight. Single

colonies (transformants) were selected and inoculated overnight in a 5 mL of ampicillin-resistant 2YT medium and incubate overnight at 37°C with shaking at 160 rpm, acquired broth was used for plasmid extraction and sent for DNA sequencing.

2.2.8 Transformation into BL21 (DE3) chemically competent *Escherichia Coli* Strain

An *E. coli* BL21 (DE3) strain (Sigma Aldrich UK) was used to sub-clone sequence-verified plasmid DNA products from the original DH5 α cloning. 5 μ L of sequence-verified plasmid was added to a 50 μ L aliquot of freshly thawed stock of BL21 (D3) chemically competent *E. coli* and incubated on ice for 30 min. Sample suspension was heat-shocked at 42°C for 45 seconds and transferred immediately to the ice for 2 min to recover. A 1% glucose supplemented 2YT medium (500 μ L) was added to the solution and incubated at 35°C for 1 hour. A 200 μ L and 50 μ L sample of suspension (containing transformed cells) was spread on a 100 μ g/mL ampicillin-resistant LB agar plate and incubated at 37°C overnight. Single colonies (transformants) were selected and inoculated overnight in a 5 mL of ampicillin-resistant 2YT medium and incubate overnight at 37°C with shaking at 160 rpm, lysate recovered was used for plasmid extraction and sent for DNA sequencing.

2.2.9 Transformation into C43 (DE3) chemically competent *Escherichia Coli* Strain

Ten microliters of suspected ligated DNA constructs were added to a 40 μ L aliquot of freshly thawed stock of C43 (DE3) chemically competent *E. coli* and incubated on ice for 30 min. It was heat-shocked at 42°C in a water bath for 45 seconds and transferred immediately to the ice for 2 min to recover. A recovery growth medium

(250 μ L) was added to the solution and incubated at room temperature for 1 hour. After that, a 50 μ L and 200 μ L sample of suspension (containing transformed cells) was spread on LB agar plates containing 100 μ g/mL ampicillin and incubated at 37°C overnight. Selective single colonies (transformants) were selected after the overnight incubation and incubated overnight in a 5mL of 2YT broth (with ampicillin resistance) at 37°C with shaking at 160 rpm. The transformant broth was pooled together and extracted according to Gene Jet plasmid miniprep protocol (Thermo scientific UK) and sent for DNA sequencing after plasmid concentration had been resolved using nano-drop spectrometer.

2.3.0 Post-Overexpress Plasmid Stability Assay

About 100 μ L of lysate samples were withdrawn at 2hours, 3 hours and overnight, post-IPTG-induction (0.5mM IPTG) of AGTR1-Fusion protein overexpression. The lysates were treated using plasmid preparation methods as described in 2.3.0, to extract lysate DNA plasmid. The plasmid concentrations were obtained and recorded using a tabletop nano-spectrometer and further clarified using agarose gel electrophoresis methods as described in 2.2.3. The stability and quality of AGTR1-plasmid DNA are thereafter analysed.

2.4.0 Protein Method

2.4.1 Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE Analysis)

A discontinuous electrophoresis gel system (Laemmli, 1970) was used with a 12% 15 well-precast TGX stain-free gels (Criterion[®] Bio-Rad UK). The buffers required

include a loading buffer 4:1 leammli buffer: β -mercaptoethanol(Biorad UK)), SDS running buffer of 1x TGX buffer and precision plus ladder. The precast TGX strain-free gel was clamped and placed inside the chamber filled with 1xs TGX running buffer. Each overexpression protein samples, and other lysates were mixed with a 5:1, loading buffer. SDS-PAGE suspensions were sonicated using a water bath sonicate for 5min and heated for 99°C for 5min. The fitted comb attached to the precast gel was removed before loading 10 μ L of protein samples unto the stain-free gel wells. Gels were electrophoresed for 45 min at a constant voltage of 180V. The gels were activated before images were taken. For Coomassie blue staining, gels were soaked overnight in coomassie blue and de-stained after 24 hours of Coomassie blue staining. SDS-PAGE gels were captured and visualized by the ChemiDoc system and software (Bio-Rad UK).

2.4.2 Protein Overexpression

Gene fusion protein overexpression was performed after a recombinant plasmid had been inserted into a host. This is achieved by transforming a gene into a chemically competent cell (BL21 (DE3) and C43 (DE3) with each of the plasmids as detailed in 3.2.5. Antibiotics-resistant colonies were selected and grown in 5 mL 2YT broth overnight at 37°C left to shake at 250 rpm. 1mL portions of the overnight grown colonies were sub-cultured into 50mL LB or 2YT broth (ampicillin-resistant). The 50mL culture was grown at a selected temperature at 180 rpm. At an optical density of 0.6 (OD₆₀₀), the overexpression culture was induced by a lactose analog 1mM Isopropyl β -D-thiogalactopyranoside (IPTG) at different concentrations and incubated at different times before harvesting by centrifugation.

2.4.3 Large-scale gene expression

Large-scale expression is achieved by selecting a single ampicillin-resistant colony (transformant) and grown in 10mL LB or 2YT broth (supplemented with ampicillin and 0.5M sorbitol and 1mM Betaine) overnight at 37°C left to shake at 250 rpm. 1 – 3mL portions of the overnight pre-culture were used to inoculate a 100 mL LB or 2YT broth (ampicillin-resistant). The 100mL culture was grown at a selected temperature and 180 rpm to an optical density of 0.6 – 1 at 200 rpm, 37°C. Protein expression was induced upon the addition of optimal IPTG concentration and incubated at temperature and time duration that gave maximum protein overexpression following optimization.

2.4.4 Production of Lysates

After a 300 mL protein overexpression, the cell products were harvested by centrifugation at 3000 x g for 10 min at 4°C. The pelleted cells were re-suspended with 30mL phosphate buffer saline (PBS), corresponding to the equivalent volume of overexpression medium. Thereafter the PBS re-suspended lysates were lysed mechanically by sonicating the cell lysates, under these conditions 10 cycles of 30sec “ON” and 30sec “OFF” sonication, using ultrasonic disintegrator (MSE Soniprep 150 plus UK). The cell lysates were sonicated in a sterile 50 mL universal tube place on the ice, during the cell lysis procedure. The protein lysates were clarified by ultra-centrifugation at 20,000 g for 20 min at 4°C. The supernatant recovered from the lysate production was stored as a soluble fraction; 20µL of the sample were verified using SDS-PAGE analysis and pellet stored for further solubilization analysis, all storages were at -20°C.

2.4.5 Protein Refolding and Solubilisation

The lysate production pellets were weighed and resuspended in 1/10th of expression volume firstly in 6M Urea and later in an optimized Urea detergent (8M Urea, 0.5% Tween 20, 10mM DTT in 20mM Tris pH 9). The suspension was incubated overnight at 4 °C on a roller. Thereafter, the overnight incubated suspension was centrifuged 3000 x g for 10 min at 4°C and supernatant recovered and refolded. The refolded fraction was obtained by diluting the supernatant with about 20-folds of refolding buffer (20mM Tris buffer at PH 9). The mixture was stirred slowly while adding the refolding buffer. A fraction of these products, including soluble fraction, insoluble fraction, and refolded fraction, were resolved by SDS-PAGE analysis. The rest of the samples stored at -20°C using absolute (100%) glycerol for long-term preservation.

2.4.6 Extraction of Soluble Protein by Further Mechanical Approach

Initial steps towards harvesting soluble recombinant fusion protein samples are described in (2.5.4). Further methods were applied, when soluble protein was unable to be harvested sufficiently via the steps described in (2.5.4). Protein lysates were re-suspended with centrifugation supernatants obtained at (2.5.4) and stored at -20°C for 48 hours. Protein lysate sample, thereafter, withdrawn and freeze-thawed, after which they were sonicated under these conditions, 10 cycles of 30sec “ON” and 30sec “OFF” sonication, using Soniprep 150. Thereafter protein lysate samples were clarified by ultra-centrifugation at 20,000 g for 20 min at 4°C. The supernatant was recovered as the soluble recombinant protein lysate and stored at -20°C, while 20µL of the sample, were verified using SDS-PAGE analysis.

2.5.0 Protein Purification Analysis

The recombinant fusion proteins produced in this research were purified from their clarified lysate using two main protein purification methods, which are the batch purification and Ni-NTA Spin purification method.

2.5.1 Batch Purification

Maltose Binding Protein tagged recombinant fusion protein lysates (soluble fractions) were purified through batch purification. This was by applying an Amintra MBP resin (Expendeon UK) for MBP tagged protein. A 50% slurry working stock was made-up by re-suspending 0.5mL of neat MBP slurry in 0.8mL phosphate buffer saline (PBS). For low scale batch purifications, 5mL of lysate suspension were added onto 100 μ L of the 50% slurry in a tightly sealed 15 mL sterile universals, this sample suspension, was placed on tabletop roller machine 60mins in a 4 °C environment. After that, the suspension was centrifuged at 6000 rpm for 10mins, and supernatant marked and stored as bound protein. 10mL of PBS was applied to washing the pellet recovery post-centrifugation for three times in 15 min. To wash, the PBS embedded pellet is collected in 15 mL sterile universal tube and placed those on a tabletop roller, following the same conditions and time described above. After the third wash, pure recombinant fusion proteins were eluted by applying 100 μ L of 10mM Maltose to the slurry bound lysates. To analyse the success of this procedure, a small fraction of the eluted sample was analysed using SDS-PAGE analysis and the rest of the samples stored at -20°C.

2.5.2 IMAC Column Chromatography Protein Purification

N-terminal and C-terminal Histidine-tagged *R. equi Vap*-fusion protein were purified using Ni-NTA IMAC column (Bio-Rad UK) on the Biologic column chromatographic system. The application of 1 x NP, 20mM imidazole, positioned as running buffer A, was applied throughout the manual mode protein purification, at 1 mL/min flow rate. Buffer A was run through the system until the system was flushed and a straight baseline, consistently observed on the LP data view. The system flow was stopped, and the filling loop isolated and the IMAC-column was connected unto the system. Buffer A was running through again, through the system, until the baseline was observed to be stable. About 2 mL of recombinant *Vap*-fusion protein lysate sample was collected on a sterile syringe and loaded into the fill-loop. Also, the outflow was connected to the fraction collector and started, at a collection volume of 1mL. The filling loop was then connected in-line and purification initiated by pressing “start” and “record.” The sample could continue to run until unbound *Vap*-fusion proteins flow through completely and the chromatogram is returned to baseline. This was repeated until an optimal pure sample set of 8 – 10 fractions were collected. The system was thereafter stopped and buffer A, replaced with 1 x NP imidazole 500 mM buffer B, and the buffer B could run until final pure proteins were eluted, and baseline stabilized. The sample collection was stopped, and LP data view recording saved. The IMAC-column was dismantled and gently flushed by applying 0.1 M NaOH using a syringe.

2.5.3 Bradford Assay

Bradford assay is a rapid and sensitive protein quantitation method for testing for the presence of microgram proteins, using protein binding dyes. This assay measures the concentration of absolute proteins in a sample. About 30 μ L each of some selected eluted Ni-NTA purified proteins, was added unto a sterile 96 well plate. Assay blank was established by adding 30 μ L of water on a separate well. Thereafter, 1.5 mL of Bradford reagents were added unto each well-containing the sample and mixed appropriately by pipetting up and down. The reaction was observed at room temperature for 5mins and measured thereafter at 595nm using a plate reader machine.

2.5.4 Cleavage Restriction

The purified MBP fusion protein lysate sample was measured for its concentration using a tabletop nano-drop spectrometry. The sample concentration per mg/mL was used to calculate the total volume of pure MBP fusion protein lysate required to perform cleavage restriction. A variable amount of PBS was used to mark up the volume of pure lysate to 20 μ L in cases where the concentration of the pure MBP fused protein lysate required a low amount due to its original concentration. Thereafter, 1 μ L of 200 μ g/mL of factor-Xa protease (NEB, UK #P80105) was added to the purified MBP fusion protein lysates respectively and mixed by vortex for 30 seconds. The only 5 μ L of the factor-Xa protease mixture was applied to the analysis, alongside MBP5 paramyosin (NEB UK, #E80525), which was used as a positive control. Cleavage removal mixture was incubated for a 24hours period and samples were taken at the 4th hour, 6th hour and 24th hour period respectively.

Thereafter, 2.5µL of 4x loading buffer was added after cleavage was completed at every stage of sample withdrawal (4H, 6H, and 24H). MBP restricted sample was heated for 5min and analyzed using SDS-PAGE analysis. Most of the samples are thereafter stored at -20°C.

2.6.0 Protein Isolation

After the recombinant CCL5-MBP fusion proteins were successfully cleaved, using factor Xa protease (Bio Lab UK). A few other techniques were employed to harvest and isolate and independent recombinant CCL5 protein, from its original fusion protein construct. Below are the methods applied with the aim of isolated non-tagged recombinant CCL5 wild type and mutant proteins.

2.6.1 Protein Isolation by MBP Reverse-Batch Purification

1 mL of Factor-Xa digested, pure CCL5 - fusion protein lysates, was withdrawn from -20°C and freeze-thawed and transferred into a 1.5mL microcentrifuge tube. 400µL of MBP Amintra Resin 20% Ethanol (Expendon UK), was added to the pure cleaved fusion protein sample and mixed gently. The sample mixture was thereafter incubated at 4°C for a period of 1 hour and 12 hours respectively on a tabletop rolling machine. After the incubation, the sample mixture was centrifuged at 14,000 rpm for 15 mins on a table-top microcentrifuge. The sample mixture supernatants were recovered and a 20µL analyzed using SDS-PAGE for the presence of isolated pure CCL5 protein. The pellets containing the MBP-resin were all analysed by SDS-PAGE. The rest of the samples were stored at -20°C thereafter.

2.6.2 Protein Isolation by Heparin Reverse-Batch Purification

5mL of soluble Wildtype-CCL5 – MBP fusion protein lysates, was withdrawn from -20°C and freeze-thawed and transferred into a 10 mL sterile universal tube. 100µL of Capto™ Heparin Resin (GE Healthcare Science UK), was added to the soluble protein sample and mixed gently. The sample mixture was thereafter incubated at 4°C for a period of 2 hours on a table-top rolling machine. After the incubation, the sample mixture was centrifuged at 14,000 rpm for 15 mins. The sample mixture supernatant decanted, while the resin pellet was recovered and washed three times intermittently with PBS, by incubating on a table-top roller for 30min at 4°C. After the final wash, the heparin resin was recovered and analysed using SDS-PAGE for the presence of isolated pure CCL5 protein.

2.6.3 Isolation by Centrifugal Filtration

The Amicon Ultra-15 Centrifugal Filter Device is a single spin protein recovery approach of the retained protein (CCL5) via ultra-filtration. The device was calibrated by adding 5 mL of PBS onto the filter chamber unit and centrifuging at 3000 rpm for 5 mins, to wet the filter. 1.5 mL of Factor-Xa digested, freeze-thawed pure CCL5 - fusion protein lysates were loaded onto the upper chamber of a 30,000 Da Amicon Ultra-15 Centrifugal Filter Device (Millipore UK). The filter chamber was inserted back to the Amicon device and tightly sealed to avoid leakage. Thereafter, Amicon tube was centrifuged at 6,000 rpm for 15 Min at 25°C. The CCL5-fusion protein samples, retained at the filter chamber of the Amicon tube, were obtained by pipetting and designated as a supernatant sample. While the samples that had flown through the 30,000 Da pores, were retained at the lower part of the

chamber, and collected by pipetting and designated as flow-through samples. Both the supernatant and flow-through samples of the 30,000 Da Amicon Ultra-15 Centrifugal filtration, were analysed for the presence of native recombinant CCL5 protein from its fusion construct. About 20 μ L the samples were further analysed using SDS-PAGE analysis. The remaining samples are thereafter stored at -20°C.

2.6.4 Buffer Exchange Optimised Centrifugal Filter Isolation

A buffer exchange reaction was initiated on purified factor-Xa cleaved, CCL5-MBP fusion protein samples, by applying 50 mM Tris pH 8.5 150 mM NaCl buffer, to purge the purified protein lysate of Phosphate buffer saline (PBS). A 1 mL sample of factor-Xa digested, purified CCL5 - fusion protein lysate of 1.28mg/mL was withdrawn from -20°C and freeze-thawed. Thereafter, the sample was loaded onto the filter chamber unit of a 30,000 Da, Amicon Ultra-15 Centrifuge device and made up to 3 mL volume by adding 2 mL 50mM Tris pH 8, 150 mM NaCl buffer. The filter chamber was inserted back to the Amicon device and tightly sealed to avoid leakage. Thereafter, Amicon tube was centrifuged at 5,000 rpm for 20 Min at 25°C. This buffer exchange and centrifugation process were repeated three times, by applying a total volume of 6 mL of 50mM Tris pH 8, 150 mM NaCl buffer on the PBS buffered, purified CCL5-fusion protein sample. After the last centrifugation, the purified CCL5 protein samples, retained at the filter chamber of the 30,000 Da Amicon tube, were obtained by pipetting at stored at -20°C. About 20 μ L of the samples were further analysed using SDS-PAGE analysis to verify the presence of 8kDa pure CCL5 protein.

2.6.5 2nd Buffer Exchange and Size Differential Optimised Centrifugal Filter

Isolation

A second buffer exchange reaction was initiated on purified factor-Xa cleaved, CCL5-MutBoth-MBP fusion protein samples by applying 20mM Tris, 1mM EDTA, 1mM DTT, 5mM MgCl₂ pH 8.0 buffer, to purge the purified protein lysate of Phosphate buffer saline (PBS). A 1 mL sample of factor-Xa digested, purified CCL5-MutBoth-MBP protein lysate of 3.731 mg/mL was withdrawn from -20°C and freeze-thawed. Thereafter, the sample was loaded onto the filter chamber unit of a 30,000 Da and 50,000 Da Amicon Ultra-15 Centrifuge device respectively, and made up to 500 µL of 20mM Tris, 1mM EDTA, 1mM DTT, 5mM MgCl₂ pH 8.0 buffer. The filter chamber was inserted back to the Amicon device and tightly sealed to avoid leakage. Thereafter, Amicon tube was centrifuged at 6,000 rpm for 35 Min at 25°C. This buffer exchange and centrifugation process were repeated three times, by applying a total volume of 1.5 mL of 20mM Tris, 1mM EDTA, 1mM DTT, 5mM MgCl₂ pH 8.0 buffer on the PBS buffered, purified CCL5-MutBoth-MBP protein sample. After the last centrifugation, the supernatant and flow-through CCL5-MutBoth-MBP samples, retained at the two chambers of both 30,000 Da and 50,000 Da Amicon tubes, were obtained by pipetting at stored at -20°C. About 20µL the samples were further analyzed using SDS-PAGE analysis to verify the presence of 8kDa pure isolated CCL5-MutBoth protein.

2.6.6 Size Differential Exclusion Based Centrifugal Filter Devices Isolation

A 4 mL purified CCL5 -fusion protein lysate product (1.28 mg/mL) which had been buffer exchanged, and protease cleaved was withdrawn from -20°C and freeze-thawed. Thereafter, the sample was loaded unto the filter chamber unit of a 30,000 Da, Amicon Ultra-15 Centrifuge device. The filter chamber was inserted back to the Amicon device and tightly sealed to avoid leakage and was centrifuged at 5,000 rpm for 20 Min at 25°C. About 400µL of protein samples, retained at the filter chamber of the 30,000 Da Amicon tube, were obtained by pipetting and stored at -20°C, while about 2.5 mL 30KDa and 50kDa Amicon separation of CCL5MutBoth-MBP flow-through sample was recovered and analysed further, by concentrating protein sample, using a 3,000 Da Amicon Ultra-15 Centrifuge device. The samples were loaded unto the filter chamber unit of the 3,000 Da, Amicon Ultra-15 centrifuge device and centrifuged at 5,000 rpm for 20 Min at 25°C. 200µL supernatant samples retained at the filter chamber were obtained by pipetting at stored at -20°C. Thereafter, 20µL of each the samples were further analysed using SDS-PAGE analysis to verify the presence of 8kDa pure CCL5 protein. A second size differential protein isolation was assayed, by applying a 4 mL buffer exchanged, protease cleaved purified CCL5 - fusion protein lysate to a 30,000 Da and 50,000 Da Amicon Ultra-15 Centrifuge device, respectively. The rest of the methods followed through as described in 2.7.2 except with centrifugation time been increased to 1 hour. Both supernatant samples (held at the filter chamber) and flow-through samples of the 30,000 Da and 50,000 Da Amicon Ultra-15 Centrifuge devices were recovered and 20µL of each sample verified using SDS-PAGE analysis.

2.7.0 Isolation of Disrupted Cellular Debris by Dialysis

Before SPR investigation, all purified protein samples used in this work were subjected to nucleic acid removal treatment by adding 0.05mg DNase (Sigma Aldrich USA), and 0.005mg RNase (Sigma Aldrich USA) into 4mL purified protein samples, to remove unwanted nucleic acid /disrupted cellular materials from the recombinant protein in order to reduce viscosity. Thereafter, the protein sample was loaded upon a molecular porous dialysis membrane tubing MWCO-35Kd (Spectrum Laboratories Inc. USA) after water had been added to the dialysis tubing to moisturize and allow for easy loading protein sample. Care was taken to ensure that samples were loaded carefully before and after a double airtight knot was formed. About 500 mL of dialysis buffer 10mM HEPES, 2M NaCl, 50mM EDTA, 0.005% Tween 20 (for CCL5-proteins) and 10mM HEPES, pH 7.4, 150 mM NaCl, 3mM EDTA 3mM, Tween 20 0.005% respectively were added unto 1000 mL beaker, with a magnetic stirrer aid added. Thereafter the purified protein knotted dialysis tubing was dropped into the buffer-loaded beaker and place on a stirrer to stir at 30 rpm for 24 hours at 4°C. The protein concentration after dialyzed chemokine-CCL5 and Vap-proteins are presented in chapters 4 and 5 respectively.

2.7.1 Dialysis of Biotin-Heparin Sodium Salt

Heparin-biotin sodium salt $\geq 97\%$ of 15kDa lot (B9806) acquired from Sigma Aldrich was found to have excess unbound biotin contamination. To clean up the heparin-biotin sodium salt, dialysis was applied using a Slide-A-Lyzer 2K (2,000 MWC) dialysis cassette (Thermo Scientific. USA). Heparin-biotin sodium salt was bought in a 10mg/mL stock concentration. A 1mg/mL concentration to a 3mL volume was

obtained by diluting 300 μ L of 10mg/mL heparin-biotin sodium salt, in a 2.7mL of 300mM NaCl salt. The 3mL of heparin-biotin sodium salt 1mg/mL was collected using a 500 mL sterile syringe and injected into the dialysis cassette, by channelling the needle through the tiny hole of the dialysis cassette. About 400 mL of 10mM HEPES, 2M NaCl, 50mM EDTA, 0.005% Tween 20 (dialysis buffer) was added into a 1000 mL beaker, with a magnetic stirrer aid added. Thereafter the heparin-biotin loaded dialysis cassette was dropped into the buffer-loaded beaker and placed on a stirrer to stir at 30 rpm for 5 hours at 4°C. The dialysis buffer was changed intermittently hourly. About 1.5 mL, 63 μ g/mL (0.063 mg/mL) of dialyzed heparin-biotin sodium salt sample was recovered at the end of the dialysis.

2.8.0 Surface Plasmon Resonance

Surface Plasmon resonance (SPR) is a specialized and reproducible microfluidic technique used to investigate the binding interaction, kinetics, and affinity between non-labelled molecules and substrates as previously described ((Sarrazin *et al.*, 2005, Martínez-Burgo *et al.*, 2019)). The SPR used for this study is a Biacore system, this instrumentation works by immobilizing a molecule often referred to as a “ligand” onto the surface of a sensor chip, after which the “analyte” molecule is flown over the immobilized ligand. The binding interaction assayed in this thesis using SPR is (i) WT-CCL5 and its mutants against a biotinylated heparin sugar molecule and the (ii) binding interactions between of *R. equi* Vap proteins. All SPR work was performed in the Newcastle University United Kingdom using a Biacore X100 under the guidance of Dr. Helen Waller (ICaMB, Newcastle University). The running buffer used was HBS-P (10mM HEPES pH7.4, 150mM NaCl, 0.005% P20)

and all reagents used are from GE Healthcare (Uppsala, Sweden), except otherwise stated.

2.8.1 SPR Capturing Methods

2.8.1.1 VapA-pET22a Immobilisation

The outer layer of CM5 Biacore sensor chips are coated with a matrix of covalently linked carboxymethylated dextran; these unbranched polymers of carbohydrate molecules present to be flexible therefore permitting the movement of ligands when attached on them. The binding interaction of *R. equi* VapA-pET23a was done via Surface Plasmon Resonance at 25°C. The immobilization method for CM5 sensor chip was initiated via the amine coupling process (direct or covalent coupling method) as introduced in section 1.2.1. A new CM5 chip was docked unto the BIAcore X100 and was activated using a 1:1 mixture 200µL of 1-ethyl-3-(3-dimethylamino propyl)-carbodiimide (EDC) and *N*-hydroxysuccinimide (NHS) respectively, were injected unto flow path 2-1 for 600 seconds at 5µL/min flowrate. Thereafter injection flow path was changed to flow cell 2 (fc2), leaving the flow cell 1 (fc1) as the reference flow cell. An RNase and DNase purified VapA-pET23a dialyzed in 10mM HEPES, pH 7.4, containing 150 mM NaCl, EDTA 3mM, Tween 20 0.005% overnight at 4°C, and was diluted to 650 µg/mL and injected on the activated sensor chip at 5 µL/min for 120 seconds on flow cell 2. After enough VapA-pET23a was immobilized, the flow path was changed to 1 and 2, where the reactive groups remaining on the activated CM5 surface were blocked with 1M ethanolamine-HCl pH 8.5 at flow rate 5 µL/min on both flow cell 1 and 2 for 60

seconds. VapA-pET23a CM5 biosensor chip immobilization level was found to be 1,911 response units (RU).

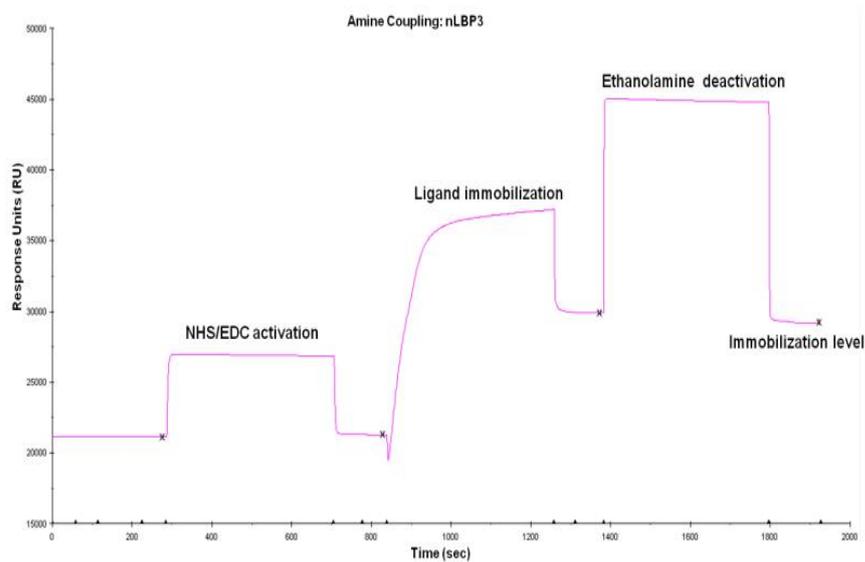


Figure 2.1: A typical sensorgram schematics of ligand immobilization using amine coupling.

Trace showing the surface activation of CM5 chip using EDC/NHS followed by the ligand immobilization and deactivation by ethanolamine. (Jahanshahi *et al.*, 2014)

2.8.1.2 Biotinylated-Heparin Immobilisation

The outer layer of streptavidin (SA) Biacore sensor chips are coated with a matrix of covalently linked streptavidin dextran, these unbranched polymers of streptavidin molecules present to form an irreversible binding complex with biotinylated ligand molecules when attached on them. The binding interaction of biotinylated heparin was done via Surface Plasmon Resonance at 25°C. The immobilization method for the SA sensor chip was initiated via a non-covalent coupling process (indirect coupling method) as introduced in section 1.2.1. A new SA chip was docked unto the Biacore X100 and was normalized by injecting a pre-wash buffer (1.5 NaCl, 50mM NaOH) for 60 seconds at 10µL/min flowrate on fc 2-1. Thereafter injection flow path was changed to flow cell 2 (fc2), leaving the flow cell 1 (fc1) as the reference flow cell. 20µg/ml of biotinylated heparin in 300 mM

NaCl was at 10 μ L/min into the flow cell 2 of the SA Biacore sensor chips for 60 seconds. Thereafter dialyzed biotinylated heparin immobilization reached 77 RU; hence flowrate was changed to 5 μ g/mL and injected for 120 seconds to obtain an immobilization level of 204 RU. This injection was repeated for another 120 seconds until the maximum RU of 517 was reached. The flow path remained unchanged, and the immobilized SA-biosensor chip was regenerated by injecting 2M NaCl, at 5 μ L/min for 120 seconds to remove unbound heparin.

2.9.0. SPR Binding Affinity and Kinetics Interaction Methods

2.9.1 Binding Interaction of *R. equi* Virulence Associated Proteins

The following Vap proteins VapG-pET23a, VapG-pET28, VapH-pET2 and VapE-pET23a, at 50 μ M and 100 μ M (diluted in 10mM HEPES, pH 7.4, 150mM NaCl, EDTA 3mM, Tween 20 0.005%), were passed over the VapA-pET23a activated and immobilized biosensor on a customised automatic run at 25°C. Custom assay was set up at 60 seconds association time and 500 seconds dissociation time, at flow rate of 10 μ L/min and regeneration at 60 seconds at 10 μ L/min using buffer. The binding affinity and kinetics results of VapA-pET23a protein against VapG-pET23a, VapG-pET28, VapH-pET2, and VapE-pET23a, were shown on a sensorgram and analyzed using a Biacore SPR X100, BIAevaluation software.

2.9.2 Binding Interaction Chemokine - Biotinylated Heparin

Following the surface preparation of the SA Biacore sensor chips, SPR binding affinity was performed to compare the binding interaction properties of the immobilized biotinylated heparin, against different concentrations of CCL5

chemokines (wild-type and mutants). A concentration range of CCL5 protein (positive control) acquired from 1 µg/mL – 0.01 µg/mL (control-CCL5), Wt-CCL5-MBP, CCL5-MutBoth-MBP, CCL5-MutBoth-MBP and CCL5-MutBoth-MBP of consistent concentrations (1µg/mL – 0.01µg/mL), were flown over the surface of the immobilized SA chip at 30µL/min for 120 seconds (association), with a dissociation phase of 500 seconds as similarly described (Martínez-Burgo *et al.*, 2019). A flow cell of 2 – 1 was used for the streptavidin immobilized SPR experiment, where the reference flow cell (flow cell 1) was subtracted from the RU and the binding kinetics sensorgram results analysed using a BIAevaluation software.

2.10. Other Binding Interaction Assay Techniques

2.10.1 Trans Endothelium Flow Assay

In order to investigate the neutrophil adhesion in response to endogenous and exogenous CCL5 peptides under physiological *in vitro* condition, the Venaflux platform (Cellix Ltd., Dublin, Ireland) was applied.

2.10.1.1 Cell culture

A 175 cm² tissue flask was grown to confluency in MCDB-131 media (Carter *et al.*, 2003) with HMEC-1 endothelial cells and then stimulated for 24 hours with TNF-α (10 ng ml⁻¹) and IFN-γ (10 ng ml⁻¹). THP-1 monocytes were grown in suspension to confluency and stimulated with TNF-α (10 ng ml⁻¹) and IFN-γ (10 ng ml⁻¹) for 24 hours prior to running the assay.

2.10.1.2 Seeding the flow chamber

A vena8 endothelial+ chip (Cellix, Ireland) is used for the assays and image of the biochip platform is shown below. Each channel of the chip was loaded with 10 μl of Bovine Fibronectin (5mg ml^{-1} in 0.5M TBS pH 7.5, Sigma Aldrich, UK) and left overnight. The Following day the endothelial cells are trypsinised, washed, and then resuspended in 400 μl of complete MCDB-131 media. The Fibrinogen solution was flushed out of the chambers with air and then 5 μl of HMEC-1 cells were pipetted into each of the 8 channels, the number of endothelial cells was in excess.



Figure 2.2: A typical sample image of a vena8 endothelial chip

The flow chamber is then placed in a tissue culture incubator for three hours to allow the endothelial cells to form a monolayer on the lower surface of the chip. After three hours the wells are flushed with warm Gey's balanced salts solution (Sigma-Aldrich, UK) to remove any unbound endothelial cells.

2.10.1.3 Performing the flow assays under flow condition.

The flow assays were performed using an Exigo pump (Cellix, Ireland) controlled using the Smartflow software provided by the manufacturer. The pump was run using a 1ml syringe to hold the fluid and the viscosity of the fluid was set in the software as 0.01 dyne. The chamber was placed on an inverted microscope

connected to a GXCAM hichrome-lite digital camera (GT Vision, UK). The assays for each chamber were run as follows. Each chamber was run with Gey's balanced salt containing either no chemokine or chemokine added at concentration of 100 ng ml⁻¹ for the peprotech CCL5 or at 500 ng ml⁻¹ of the recombinant CCL5-MBP fusion protein, the solution was run through the chamber at 10 µL/min, for three minutes. The recombinant chemokine had been digested to release the MBP tag, but the MBP was still present in solution, so this is accounted for by adding a higher weight: volume. Then the stimulated THP-1 cells were run into the chamber at a flow rate of 10 µl min⁻¹ for 30 secs followed by a flow rate of 1µl min⁻¹ for 4.5 mins and this was recorded using camera device attached to the microscope. The number of cells sticking to the surface of the endothelium was counted for analysis.

2.10.2 Thermal Shift Assay

Thermal shift assay was conducted using manufacturers control and recombinant *R. equi* Vap proteins in accordance with manufacturer's guidelines. The protocol for Thermal Shift™ Dye Kit manufactured by Thermo Fisher Scientific USA was applied.

Chapter Three

The pharmacogenomics implication of CHD
SNP Protein interactions against CHD drugs

3.0 Case Study: Coronary Heart Disease

Coronary heart disease (CHD) or coronary vascular disease (CVD) is a group of diseases caused by the build-up or blockage of coronary arteries by lipid plaques. These arteries supply oxygenated blood to the heart (Nhlbi.nih.gov, 2020). CHD is clinically characterized by angina (chest pain), periodic heartburn and discomfort, often traveling into the arm, jaw, neck, shoulder or back. These clinical indications are often followed by heart attacks and, commonly, heart failure (Benjamin *et al.*, 2019, CDCP, 2020). Many risk factors have been implicated by many risk factors, including high blood cholesterol, genetics, obesity, smoking, diabetes, excessive alcohol, diet, and lack of exercise, among others (Mehta *et al.*, 2014; Mendis *et al.*, 2011). According to the 2018 Global Burden of Disease study and a 2019 review on the epidemiology of CHD in the United Kingdom, CHD has maintained its lead as the number 1 cause of global mortality (Mensah *et al.*, 2019), with an estimated 17.5 million global mortality rates in 2015 (Nowbar *et al.*, 2019) and 17.9 million deaths in 2016 (Who.int, 2020). About 26.6 million people were diagnosed with CHD in the United States in 2014, of which 611,105 eventually died (CDC, 2015; Mozaffarian and Benjamin, 2015).

In the United Kingdom alone, an average of 1.7million instances of CHD-related episodes was recorded in NHS hospitals between 2013 and 2016. Consequently, CHD is the second main cause of death at 28%, with over 155,000 deaths annually (British Heart Foundation, 2015). According to recent reports, CHD accounted for about 404,000 inpatient cases in NHS hospitals, with a total health cost expenditure of about £6.8 billion in 2013 alone (Townsend *et al.*, 2015). A review based on the

Cardiovascular Disease Statistics 2014 shows the northeast of England has the highest prevalence of CHD, with the maximum mortality rate of 4.5% in the United Kingdom (Prachi *et al.*, 2015).

While the coronary heart disease pandemic remains alarmingly on the rise, some epidemiologists have suggested that cardiovascular disease-related mortality rate by the year 2020 might be a leading cause of 25 million deaths, representing about 36% of the global mortality rate. (Katz 2011, Prachi *et al.*, 2015)

This research focuses on the eminent need to achieve clinically proven and acceptably safe, effective, and more personalized medication for millions of patients suffering from coronary heart disease through the application of pharmacogenomics. Earlier research by Miller and Miller (1975); David and John (2002) and Sekar *et al.* (2008) have uniformly reported high intracellular LDL (serum cholesterol) as a significant risk factor associated with increased occurrence of CHD. Cooper-DeHoff and Johnson in 2015 similarly reported the present clinical application of blood pressure control drugs to have only achieved <50% result among patients undergoing hypertension treatment. The angiotensin-renin system via aldosterone in the kidney plays a long-term blood pressure-lowering function in the management of hypertension (Doolittle, 1983). Rose and Mark (2002), however, reported that this mechanism often progresses to heart failure when it becomes maladaptive.

3.1.0 Pharmaceutical CHD Treatment Options

Presently, according to the National Institute for Health and Care Excellence (NICE) prescription clinical guideline, as adopted by the National Health Services (NHS) in the United Kingdom, the available clinical chemotherapeutic options for the management and treating coronary heart disease revolve around six major classes of drugs (NHS 2014).

Table 1.1 presents these six classes of drugs, their precise corresponding binding protein or site of action where the individual classes of drugs have been observed to exert pharmacological activity in the human intracellular or extracellular membrane. More so, the table presents information on the domain characterization of these individual proteins, whether they are enzyme-bound proteins, transmembrane proteins, or receptor-bound proteins.

Table 3.1: NICE Current CHD Treatment Options

	Names of Drug	Drug Site of Action/ Protein Targets	Types of Proteins	Target Location	Protein Classification
1	Statins	3-Hydroxy 3-methylglutanyl coenzyme (HMG-CoA) reductase Inhibitor	Enzyme	Endoplasmic reticulum membrane	Glycoprotein
2	Angiotensin-Converting Enzyme Inhibitors	Angiotensin-Converting Enzyme (ACE)	Enzyme	Cell membrane	
3	Angiotensin II Receptor Antagonists	Angiotensin II Receptor Type 1 (AGTR1)	Transmembrane and Receptor	Cell membrane	G-pro
4	Calcium Channel Blockers	Calcium channel voltage-dependent gamma subunit 3 (CACNA1A)	Receptor	Multi-pass membrane	
5	β -Blocker	Adrenergic Receptor Beta 1 (ADRB1)	Receptor	Plasma membrane	
6	Anti-platelet Agents	Platelet's adenosine diphosphate receptor (P2Y ₁₂)	Receptor	Cell membrane	

While there has been significant therapeutic intervention in the treatment and management of CVD in the last two decades using these classes of drugs (BHF, 2015), studies have shown alarming numbers of CHD patients who experienced no therapeutic benefit, and others with life-threatening adverse drug reaction concerns (Gholami *et al.*, 2008).

These concerns are largely due to the present limitations in understanding the few cases of therapeutic inefficacy and adverse drug reaction concerns associated with the NICE clinical prescription guidelines for the six classes of drugs used in treating CHD. The drugs that are presently available to hospitals and GPs have continued to be prescribed to patients as a first, second, and third line of CHD prescription medication without recourse to individual patient's genomic profile. The challenges posed by the present prescription guidelines have ultimately resulted in a non-patient specific one-dose-fits-all medication approach. These worrisome realities may lead to more complications of patient disease and health states, toxicity-related adverse drug reaction concerns, and perhaps death in some cases.

3.2 Mechanism of Action of CHD Drugs

3.2.1 Statins

Statins or HMG-CoA reductase inhibitors are classified under the anti-cholesterol family of drugs. Their major effectiveness is evident in both the primary and secondary inhibition of ischemic cardiac attack (David and John, 2002). Statins, also known as 3-hydroxy-3-methylglutanyl coenzyme (HMG-CoA) reductase Inhibitors, acts by competitively inhibiting the enzyme 3-hydroxy-3-methylglutanyl coenzyme (HMG- CoA) reductase or HMGCR, and act as therapeutic agents in the reduction

of CHD risk. HMGCR is classified as a rate-regulating enzyme in the mevalonate pathway of cholesterol biosynthesis using four electrons in the reduction of HMGCR enzyme to mevalonic acid and coenzyme A (CoA) and two NADPH as electron donors. Mevalonate (mevalonic acid), an isoprenoid, is a precursor of cholesterol in the cholesterol biosynthesis pathway; other products of mevalonic acid biosynthesis are heme and farnesyl-pyrophosphate. The activity of the HMGCR enzyme is regulated by the availability of intracellular cholesterol plasma levels, through a series of controlled transcription and translation mechanisms. These mechanisms involve the degradation of present enzymes and a protein kinase-catalyzed phosphorylation of Ser⁸⁷², which results in a reduced affinity for NADPH and a consequent decline in HMGCR activity (Istvan and Deisenhofer, 2000, Istvan *et al.*, 2001). The affinity of statins for the enzyme HMGCR is nearly 3-fold greater than the HMG-CoA natural substrate (Moghadasian, 1999). The overall pharmacokinetics of statins because of an up-regulation of low-density-lipoprotein receptors in the liver, leads to an intracellular decrease of low-density lipoprotein cholesterol (LDL-C) of between 20-60% patient users, depending on the dosage and the exact statin administered (David and John, 2010). Statins do to exhibit pleiotropic characteristics hence the inhibition of the mevalonate pathway is not only associated with the treatment of coronary heart disease but also with inhibiting proliferation signals for cancerous cells, as review by Swanson and Hohl (2006).

Table 3.2: Generational classification of major statin drugs

First Generation	Reg: Year	Second Generation	Reg: Year	Third Generation	Reg: Year
Simvastatin	1981	Fluvastatin	1993	Pitavastatin	2009
Lovastatin	1987	Atorvastatin	1996	Rosuvastatin	2010
		Pravastatin	1996		

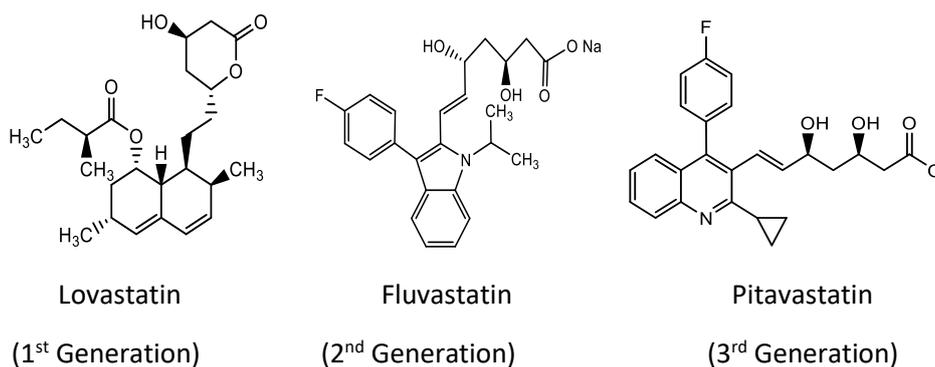


Figure 3.0: Structure of selected Statin based on their generation.

3.2.2 Statin Binding Receptor

3.2. 2.1 Hydroxy-3-methylglutaryl coenzyme A reductase (HMG-CoA) Reductase Enzyme

The enzyme 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) reductase is a member of the transmembrane glycoprotein encoded by the human HMGCR gene, which has a length of about 2500bp in its open reading frame (ORF) and positioned on chromosome 5 at the 5q13.3-q14 position (Friesen and Rodwell, 2004). Human HMG-CoA reductase, also designated as HMGCR, has a molecular weight of 97,476Da with 888 amino acids (Navdar *et al.*, 2003). HMGCR is a four-electron oxidoreductase which performs a rate-limiting enzymatic activity in the conversion of mevalonate to cholesterol in the biosynthesis of cholesterol and other isoprenoids (Istvan and Deisenhofer, 2001, Friesen and Rodwell, 2004). The

activities of HMGCR are implicated in the regulation of cholesterol biosynthesis by activation of sterol regulatory element-binding proteins (SREBP) signaling (Osborne and Espenshade, 2009) metabolism. Kersten (2008) demonstrated the role of HMGCR in the metabolism of intracellular fatty acid, triacylglycerol, and ketone bodies; furthermore, HMGCR plays a role in lipid metabolism regulation through the activation of peroxisome proliferator-activated receptor alpha (PPAR α).

HMGCR is routinely regulated by a negative feedback mechanism with the aid of sterols and non-sterol metabolites derived from mevalonate. The catalytic domain of the human HMGCR gene approximately 200kDa in size is composed of four monomeric polypeptide units consisting of two α -subunits and two β -subunits which are joined to form a $\alpha\beta$ -dimers (Holm and Rosenstrom, 2010).

3.2.2.2 Statin Interaction with HMGCR

The statin binding interactions with HMGCR remains a crucial aspect of understanding their intracellular pharmaco-efficacy in human species. All statins share a similar HMG-like moiety, often present in an inactive lactone form. During statin metabolism, these pro-drugs become enzymatically hydrolyzed to form their active hydroxy acids (Corsini and Maggi, 1995). Istvan (2003) similarly reported the formation of ionic and polar interactions between HMG-like statins and HMGCR enzymes interaction, offers flexibility and a degree of conformational rearrangement on the carboxyl-terminal residue that permits for the statin hydrophobic group to fit between the L-domain helices.

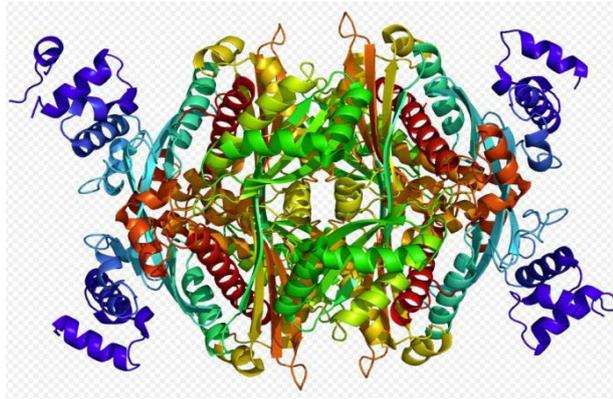


Figure (3.1): Quaternary structure of HMG-CoA reductase source: <http://www.genecards.org/cgi-bin/carddisp.pl?gene=HMGCR> Date accessed: 13/10/2015

Statins have similar rigid, hydrophobic groups that are covalently linked to the HMG-like moiety. Lovastatin, pravastatin, and simvastatin resemble the substituted decalin-ring structure of mevastatin, a naturally occurring HMGR inhibitor sourced from a microorganism in the '70s called *Penicillium citrinum* (Endo *et al.*, 1976). Fluvastatin, cerivastatin, atorvastatin, and rosuvastatin are known as type 1 statins and bind through hydrogen bonds. Furthermore, the decalin-ring interactions are entirely synthetic HMGR inhibitors, which consist of larger groups linked to the HMG-like moiety. They allow for a range in hydrophobicity from very hydrophobic (e.g., cerivastatin) to partly hydrophobic (e.g., rosuvastatin) and bind through their fluorophenyl groups (figure 3.2).

All statins are competitive inhibitors of HMGR and effectively binds the HMGR active site substrate; this, however, does not apply to the binding of NADPH in HMGR active site (Endo *et al.*, 1976). The inhibition constant (K_i) values for the statin-enzyme complexes range between 0.1 to 2.3nM for all statins (Corsini and Maggi, 1995), whereas the Michaelis constant (K_m) for HMG-CoA is 4nM (Bischoff and Rodwell, 1992).

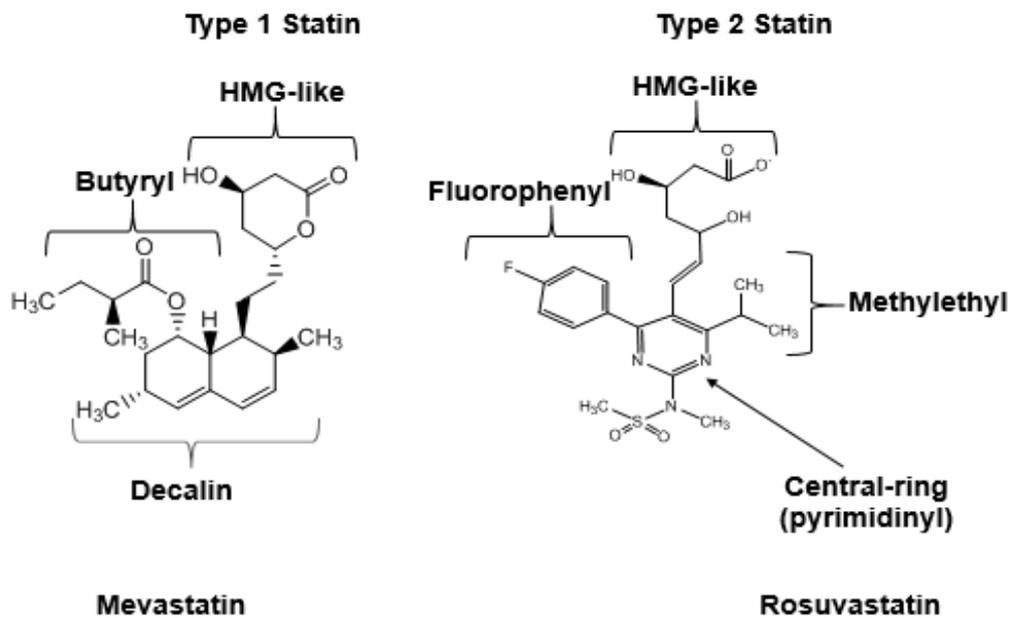


Figure 3.2: Comparison of basic type 1 statin structure (mevastatin) and type 2 statin structure (Rosuvastatin)

Furthermore, statins maintain the same structural fold around the active site residue. The 3-dimensional X-ray crystallography structure of human HMGCR showed that the monomeric $\alpha\beta$ -dimers form firm tightly associated tetramers (Istvan *et al.*, 2000). However, HMGCR only becomes active for catalytic function in its dimeric form.

Individual monomers of HMGCR consist of three distinct domains that form its tertiary structure; a helical amino-terminal domain (N-domain) constituting of residues 460 – 527, a large domain (L-domain) constituting of residues 528 – 590 and 694-872, made up of a 27-residue of α -helix as the central element, and a small domain (S-domain) constituting of residues 592 to 682, containing a central antiparallel four-stranded β -sheet (Roche, 2004).

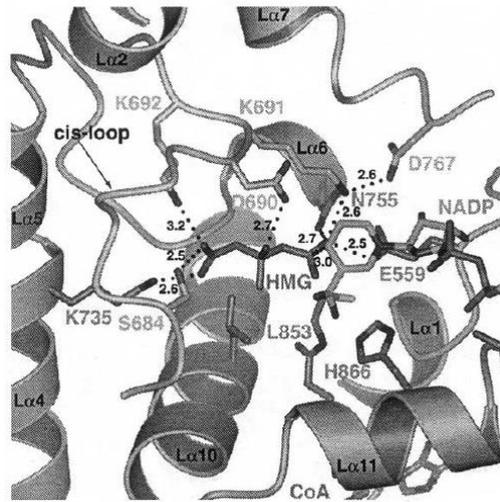


Figure 3.3: Enclosed view of HMGR enzyme active site, showing inhibitor (statin) binding. Binding networks designated with dotted lines, with numbers that indicate angstrom distances. Reproduced from (Istvan 2003).

The structure of the catalytic portion of human HMGR is characterized by a loop (residues 682– 694, called the “cis loop”), making it ideal for the binding and catalysis of statins. Because statins are competitive to HMGR, their HMG-like moieties bind to the HMG-binding portion of the HMGR active site. This binding occurs at the L-domain of a single monomer, with Tyr⁴⁷⁹ of the nearest monomer forming a hydrophobic shield above the adenine base that encloses the catalytic binding pocket (Istvan, 2003). However, the HMG-binding bulky hydrophobic group clashes with residues that constitute the narrow pocket during binding, which is meant to provide accommodations for the pantothenic acid moiety of CoA. The binding affinity of most statin molecules (particularly the HMG-like portion) to the HMGR enzyme is characterized by a conformational change in the less-rigid C-terminal of the HMGR enzyme α -helices, a movement that allows a mild hydrophobic binding site to accommodate the bulky hydrophobic part of the statin molecules. Istvan and Diefenhofer (2001) argued that the exhaustive mechanism of statin inhibition of HMGR has remained unresolved.

The ability of all statins to function similarly by binding firmly to the HMGCR active site thereby inhibits the activities of HMGCR in cholesterol synthesis. Statins are found to possess different structural variations that partly account for differences in the degree of statin therapeutic potency when administered and their behavior on HMGCR enzyme inhibitory activity (Eva, 2003).

Interestingly, the enclosed view of the human HMGCR active site allows for a better understanding of why the HMGCR enzyme active site is considered structurally ideal for the binding of statin molecules and ultimate functional in HMGCR inhibition (Istvan and Deisenhofer 2009). The 'cis-loop' feature permits a conformational positioning of different amino acid residues, which binds to the HMG-moiety of statin. Several amino acids, such as Lys⁶⁹¹, Ser⁶⁸⁴, Lys⁶⁹², and Arg⁶⁹⁰, are reported to binding effectively with HMG moiety, while Glu⁵⁵⁹ and His⁸⁶⁶ have been suggested as two important amino acids residue that binds the CoA moiety of statin (Istvan, 2003). These amino acids are found on the right-hand side of the 'cis-loop'.

A recent study credited *eubacterium (Pseudomonas mevalonii)* as the host in which the first class II-HMGCR was discovered (PmHMGCR), possessing desirable characterization for structural and mechanistic investigation. An HMGCR sequence identity-based alignment study of HMGCR (class I) and PmHMGCR (class II) showed that the active catalytic domains of both classes are conserved. However, the two proteins differ when they form quaternary structures, in that the PmHMGCR monomers form a hexamer, in contrast to the formation of a tetramer by human HMGCR (Istvan, 2001). Also, one of the major differences between the PmHMGCR

and the human HMGCR is the absence of the cis loop on the PmHMGCR, hence, the Human HMGCR is the focus of this study.

While many advanced studies have been ongoing to understand the pharmacotherapeutics and pharmacogenetics of statin interaction with HMGCR, most researchers have only focused on the effect of one or two individual statins on patients of different ages and racial backgrounds. Krauss *et al.*, (2015) reviewed variation in the HMGCR gene against ethnic differences in LDL cholesterol response to simvastatin treatment; it was concluded that the overall simvastatin-mediated LDL cholesterol reduction per ethnicity was not significant, accounting just for only 2% of total variance in the study response of sample population of 57 people (24 blacks and 23 whites), in which a total of 79 SNPs was distinctly studied.

Other vital research on human HMGCR relating to its pharmacodynamics and pharmacokinetic have also implicated individual single nucleotide polymorphism (SNP) variation and cytochrome P2D6 (CYP2D6) polymorphism as influential factors that affect the cholesterol-lowering ability of selected statins (Nordin *et al.*, 1997, Salonen *et al.*, 1999, David and John, 2002).

These studies have further highlighted the need for a new understanding of statin and HMGCR gene interaction from a more genomic and pharmacogenomics perspective. Hence information was acquired from a pharmacogenomics database and used to identify possible HMGCR gene SNPs and mutations that have strong clinical annotation among statin users. The focus was to understand the pharmacogenomics dynamics associated with statin dosage and usage, and in to identify its mutants/SNPs that show clinical annotation.

3.3.0 HMGCR Gene Mutants and Polymorphisms

The initial objective of this research with the HMGCR gene was to clone, over-express, purify and characterize a recombinant human HMGCR gene (Wild Type) alongside available variants (mutants) that show convincing clinical annotation and prevalence across wider patient statin users. The first task was to acquire information on existing human HMGCR mutant genes and identify available mutations (mutants)/SNPs that may affect how the drug acts across a range of different SNP user categories. A bioinformatics and pharmacogenomics database known as PharmGKB used to investigate and retrieve this data. The PharmGKB was used to explore the available genomic information on human HMGCR and its SNPs. The computational data obtained from PharmGKB indicated that human HMGCR has five recognized mutants (Table 2.0).

Table 3.3: Human HMGCR gene mutants

HMGCR Variant	HMGCR Variant	Allele Variation	Variant Characterization	Variant Amino Acid
rs12654264	Intron 12 of IVS 12	A>T	Non-coding region (Intronic)	Not Applicable
rs17238540	SNP 29	T>G	Intronic	Not Applicable
rs17671591	SNP 12	A>T	Not Available	Not Applicable
rs3846662	N/A	A>G	Intronic	Not Applicable
rs10474433	N/A	T>C	Not Available	Not Applicable
H7	N/A	N/A	Not Available	Not Applicable

Source: PharmGKB <https://www.pharmgkb.org/gene/PA189#tabview=tab1&subtab=33>. Date accessed 7/03/2016

However, rs17238540 and rs3846662 human HMGCR gene mutants generated from PharmGKB data were observed to be intronic mutant genes, meaning that they have only non-coding sequences within the human HMGCR gene open reading

frame, hence upon protein translation, the mutations do not translate to any amino acid substitution. Currently, there are not sufficient literature information that exists of human HMGCR rs12654264, rs17671591 and rs10474433 gene variants. Available literature shows that the above mentioned human HMGCR gene variants cannot be transcribed among statin patient's users (Table 2.0). The original aim of studying how the inhibition of human HMGCR gene by statins could be affected by human HMGCR mutations appears not to be possible because the non-coding sequences of the human HMGCR gene mutants would not reflect mutations as amino acid substitution even though it could undergo protein overexpression in effect.

3.3.1 HMGCR Gene Exon 13 Splicing

Interestingly, Marisa and Ronald (2009), identified a degree of functional variability in a non-functional H7 haplotype mutant of recombinant human HMGCR protein in their investigation on the role of human HMGCR gene alternative splicing on statin efficacy. HMGCR-H7 has been discovered to be linked with baseline lipid levels The Multi-Ethnic Study of Atherosclerosis (MESA) and Cholesterol and Pharmacogenomics Society study populations. The CAP study showed that African American H7 carriers had lower reference line total cholesterol and LDL-cholesterol level compared to non-HMGCR-H7 carriers (Krauss *et al.*, 2008). On the other hand, MESA investigation with a study population of 612 European Americans, 597 African Americans, 108 Hispanic Americans, and 627 Chinese Americans, a HMGCR H7-haplotype with SNPs 12, 20144 and 29 were reportedly linked with significantly

reducing reference line triglycerides within the African American and Hispanic American ethnic participants of the study population (Chen *et al.*, 2009).

Many researchers had earlier reported that an allele of HMGCR v1 haplotypes could be associated with reduced LDL-C response to statin (Chasman *et al.* 2004, Krauss *et al.* 2008, Marisa & Ronald 2009). HMGCR variant H7 haplotype consists of three intronic SNPs (rs17244841, rs3846662, and rs17238540) with no obvious functionality (splice site) or disequilibrium linkage with exonic SNPs.

Furthermore, recombinant human HMGCR gene, with an alternative splice of SNP rs3846662 located at exon 13 are predicted, to be associated with most of statin-induced expression when HMGCR mutant H7 exon 13 splice Hmgcr13 (-), which means the alternative deficiency of a single HMGCR-H7 exon 13 monomer from its theoretical tetramer structure (Medina *et al.*, 2008, Marisa and Ronald 2009). Alternative splicing has been implicated as an integral mechanism, regulating the risk of most cardiovascular diseases (Komamura *et al.*, 2004; Mango *et al.*, 2005).

In another study by the Cholesterol and Pharmacogenetic Society (CAPs) on full-length human HMGCR transcript HMGCR 13(+) and an alternatively spliced HMGCR 13(-) transcript without exon 13 in an immobilized lymphocytes cell line of 944 genotyped individuals (335 black and 609 white men and women), 40mg/day simvastatin for six weeks showed that alternatively spliced HMGCR mRNA lacking exon 13 (SNP rs3846662), along with other CHD risk factors helped to inform nearly 24% variation of LDL- cholesterol response in the study population (Marisa *et al.*, 2008). Marisa and Ronald (2009) further hypothesized that successful regulation of the human HMGCR gene alternative splicing mechanism might be crucial in

understanding how HMGCR exon 13 splicing could inform more effective statin response to patients with CHD (Figure 2.3).

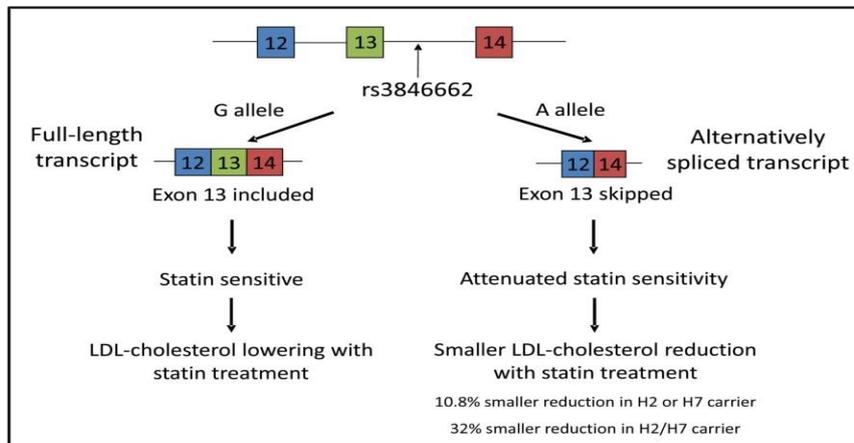


Figure (3.4): Schematics of HMGCR alternative splicing Adopted from (Marisa & Ronald 2009).

Although the SNP rs3846662 allele of HMGCR H7 haplotype is known for its role in statin-induced expression of HMGCRV_1, the CAPs studies have suggested it may not be the sole determinant of exon 13 alternative splicing and hence may not be directly or significantly associated with variation in statin response (Marisa *et al.*, 2008 and Krauss *et al.*, 2008). Interestingly, full-length and HMGCRV_1 transcript was identified in all 944 genotypes subject tested in the study. Considering the limitation of the CAPs study, in which immobilized lymphocyte cell lines were used as a model system for investigating full-length HMGCR expression in vitro.

3.4.0 AGTR1 Genes and Angiotensin II Receptor Blockers

Angiotensin II type 1-receptor protein, also designated as AT1-receptor protein, is a 41kDa protein, which forms an integral part of the renin-angiotensin system, which mediates major cardiovascular functions of the angiotensin II. The human

AT1-receptor protein is located on chromosome 3 at the q24 location 148405658 – 148463790 positions of the human genome (Koike *et al.*, 1994; Lazard *et al.*, 1995). This transmembrane protein is vital in the regulation of blood pressure and modulation of intracellular salt and body fluid level in patients with hypertension, myocardial infarction, and coronary heart disease (Wilko *et al.*, 2005, Renato *et al.*, 2014 and Cooper-DeHoff and Johnson 2011). The AT1-receptor proteins are encoded by human angiotensin receptors 1, also known as the AGTR1 gene (Thu *et al.*, 2003), and classified under the transmembrane G-protein coupled receptor family (Balakumar and Jagadeesh 2015).

The drug classified as angiotensin receptor II blockers (ARBs), and known as angiotensin receptor II antagonist binds specifically, to the angiotensin II receptor subtype 1 (AGTR1) or angiotensin II subtype 1 receptor gene. The ARBs act by selectively and effectively blocking the angiotensin II activation at the angiotensin II AT1 –receptor, which prevents the formation of angiotensin 1. This blockage results in a vasodilatory effect and consequent de-activation of the renin-angiotensin system. On the contrary, when the renin-angiotensin system becomes active, vasopressin is produced, thereby reducing the intracellular concentration of secreted aldosterol. The aldostreol is found in the kidney as an effective osmoregulator, which helps to achieve an increased renal tubule by speeding up the re-absorption of potassium, sodium, and water in the blood and urine. This phenomenon results necessitates the maintenance of osmolarity in the kidney and heart and effectively reduces the risk of CHD and high blood pressure.

The renin-angiotensin system is primarily involved in a range of activities including, the regulation of intracellular plasma volume, moderation of electrolyte, and

regulation of arterial blood pressure (Richard *et al.*, 1998). The renin-angiotensin system consists of two distinct enzymes called the rennin, which is an aspartyl protease enzyme and angiotensin-converting enzyme (ACE), which are characterized as a kinase II zinc protease enzyme. The eventual product of the renin-angiotensin system on endogenous angiotensinogen is to produce angiotensin II that acts as an effective vasoconstrictor, which affects the renal and cardiovascular function (Yel *et al.*, 2004). Interestingly, AGTR1 genes have been associated with cancer (Suganuma *et al.*, 2005; Rhodes *et al.*, 2009). A DNA meta-COPA analysis, using 3200 microarrays from 32 breast cancer bioinformatics databases, identified AGTR1 and ERBB2 as the two most prevailing breast cancer genes (Tomlins *et al.*, 2005, Burshra *et al.*, 2009)

Concerning CHD, the overall role of ARBs is to inhibit the renin-angiotensin pathway, one of the earlier developed ARBS (1995) called losartan, and an analog of imidazole-5-acetic acid was classified, and an effective antidepressant drug known as a selective but less competitive agonist of angiotensin II receptor 1. Structural alignment of losartan's benzyl group of S-8308 directly towards the N-terminus of angiotensin II, alongside other molecular modifications, resulted in a high receptor affinity of losartan ($IC_{50}=0.019\mu\text{mol/L}$) and administrable through oral route (Thomas *et al.*, 2013). Most types of ARBs possess a similar structure with losartan except valsartan, which differs not only in structure as a non-imidazole ARBs but also exhibits stronger potency of $IC_{50} = 0.0089 \mu\text{mol/L}$ (Bauer and Reams, 1995). Table 1.0 below shows different generations of ARBs and the year of release into the market. The dosage concentration and generation of ARBs are fundamental factors to be considered in studying the binding affinity of ARBs –

AGTR protein interaction. Allowing for a clearer understanding, whether the new generation ARBs (Olmesartan and Irbesartan) have become more pharmacogenetically compliant.

Table 3.4: Generational classification of ARB Drugs

First Generation		Second Generation	
Drug	Reg: Year	Drug	Reg: Year
Losartan	1995	Olmesartan	2002
Eprosartan	1997	Irbesartan	2012
Valsartan	1997		
Candesartan	1998		
Telmisartan	1998		

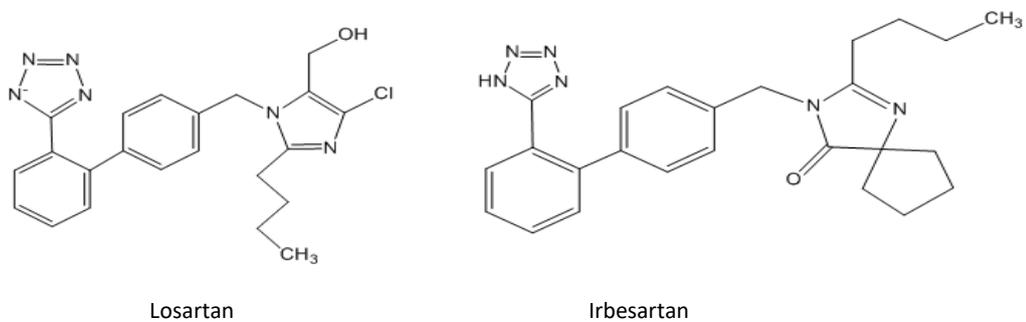


Figure 3.5: Structure of selected Angiotensin receptor II blockers based on their generations

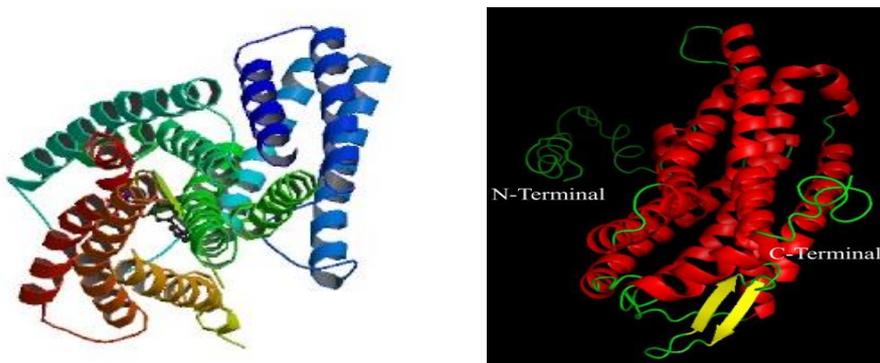


Figure (3.6): Tertiary structure of Angiotensin II Receptor Subtype 1 (AGTR1) Homology models (3D structures) Ang II-AT₁ receptor. Colours are according to the secondary structure. The red color in the figure indicates α -helix, yellow colour indicates β -sheet, and turns are in green color structures with C- and N-terminal ends.

3.4.1 AGTR1 Gene Mutants and Polymorphisms

Single nucleotide polymorphism (SNPs) are important markers in understanding and resolving the behaviors of DNA sequences with alterations that lead to phenotypic changes (Syvanen, 2001). The application of SNP information has helped to identify SNP interactions linked to common diseases (Li *et al.*, 2014). Thereby providing new genomic treatment options that accommodate genetic susceptibility due to existing polymorphic phenotypes. SNPs are the most occurring variation in the human genome, with about 9million SNPs reported in SNPs databases (Kim and Misra, 2007).

An online-based pharmacogenomics database, called pharmGKB was used, to search for SNPs information associated with the AGTR1 gene. AGTR1 gene pharmGKB search showed many SNPs, among which only two were selected and studied further, being the rs12721226 and the rs5182. At the DNA level, both rs12721226 and rs5182 of the AGTR1 gene showed an SNP alteration of G > A and C > T, respectively. The rs5182 SNP, however, showed an amino acid translation of Leucine to Leucine, which had no phenotypic variation and, consequently no clinical relevance. On the other hand, rs12721226 SNP of the AGTR1 gene codes for the amino acid translation of residue A163T showing alteration from alanine to threonine at position 163 (Arsenault *et al.*, 2010). The use of specialist high thorough put binding interaction techniques will be used to acquire binding affinity data of ABR drugs of the wild-type AGTR1 and its rs12721226 phenotype. Experimental drug binding interaction will be measured in a concentration-dependent manner until an optimal drug concentration-dependent binding interaction is attained.

Table 3.5: AGTR1 SNPs

Available AGTR1 SNPs on pharmGKB			
AGTR1 Mutants	Alleles	Function	Amino Acid Translation
rs12721226	G > A	SNP	A163T
rs2640543	A > G	SNP	N/A
rs275651	T > A	SNP	N/A
rs5182	C > T	SNP	L191L
rs5186	A > C	SNP	N/A

The pharmacodynamics of ARB drugs differ among patients who carry wild-type AGTR1 genes, in comparison to others who carry rs12721226 mutant phenotype (Arsenault *et al.*, 2010). The extent of drug binding discrepancy was verified by investigating the hypothesis that AT1-receptor proteins and ARB drug binding interaction differ among patients with different AGTR1 phenotypes. An earlier independent study by Arsenault *et al.*, (2010), had investigated the binding interaction of wild-type AGTR1 and rs12721226 against antihypertensive antagonists, where mammalian cells were used to express wild-type AGTR1 and rs12721226 proteins. The study had generated a relative binding affinity data, by assessing the quantity of 3H-inositol phosphate secreted when antihypertensive antagonists interacted with wild-type AGTR1 and rs12721226 respectively.

That previous study differs from this present work in a few ways. Firstly, the previous work was an independent study of the AGTR1 gene alone, and its mutant and drugs. The study used a mammalian cell line for the overexpression of wild-type AGTR1 and rs12721226 proteins and applied biological assays rather than biochemical assays in generating a relative non-absolute binding interaction data.

Chapter Aim

This research hoped to fill a gap in knowledge by generating a list of kinetic affinity binding data of wild-type HMGCR and AGTR1 proteins and their mutants that show clinical annotation in CHD patients. And to demonstrate that the binding behaviors of these proteins and their SNPs against factors like the generation of drugs and dosage of administration of the drugs may provide insight into the optimal dose for patients carrying different SNPs.

Results and Discussion

3.5.0 Amplification of Full-length HMGCR Gene

Reverse transcriptase-polymerase chain reaction (RT-PCR) was applied to obtain amplified full-length HMGCR gene, using the reverse and forward full HMGCR transcript primers, and the methods described in section 2

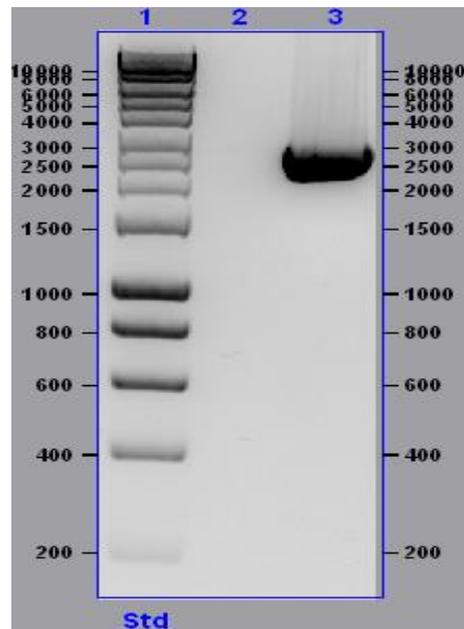


Figure 3.5 RT-PCR DNA Amplification Analysis of Full-length HMGCR Gene

Lane: Showing hyper ladder 1kb and lane 3 showing RT-PCR amplified HMGCR gene.
Resolved by 1% agarose gel electrophoresis.

3.5.1 DNA Extraction of Full-length WT- HMGCR Gene

A larger volume of RT-PCR amplified sequences of the wild-type full-length HMGCR gene were resolved using agarose gel electrophoresis, using the same method as applied in section (2.2.3). The amplicon patterns generated in the agarose gel screening of figure 3.6 were compatible with expectations. The hyper ladder 1kb band that migrates between 2000 and 3000bp markers showed to be consistent with full-length WT-HMGCR gene, which is a plasmid of 2500bp in length. The gels were cut and pooled out using a sterile blade. Qiagen DNA extraction protocol was used according to the manufacturer's instruction for the extraction of the HMGCR gene (2.2.1). The extracted HMGCR gene concentration was measured using the benchtop nano-drop reader according to manufacturer's instruction.

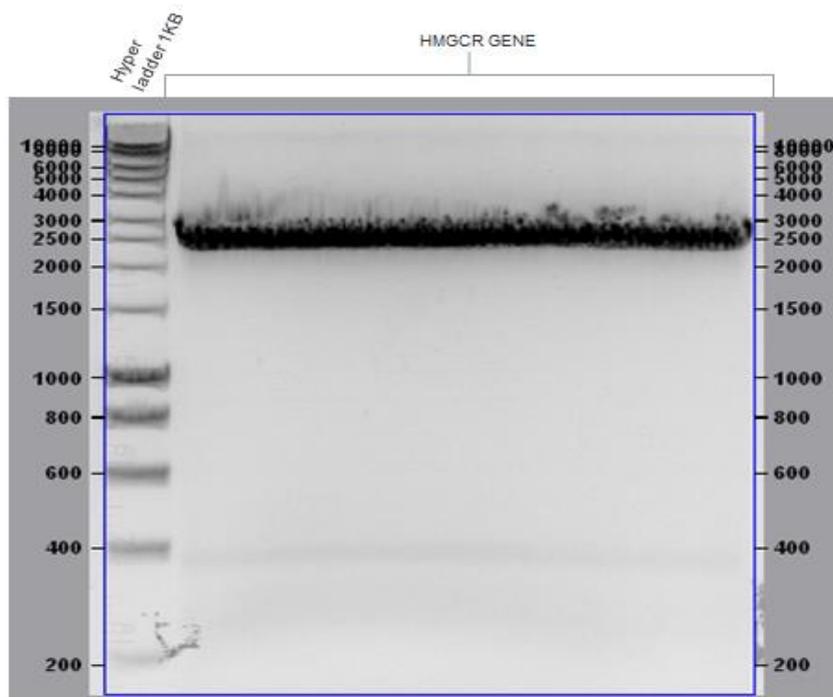


Figure 3.6: DNA Extraction Analysis of Full Length HMGCR Gene

Lane 1: showing hyper ladder 1kb, other lanes, showing: High Fidelity reverse transcriptase PCR amplified full-length HMGCR gene. Resolved by 1% Agarose Gel Electrophoresis.

3.5.2 Cloning of Amplified WT-HMGCR into pGEX-6p-1 Vector and Analysis

The purified WT-HMGCR amplicon was prepared to be cloned by restriction digest into a pre-digested treated pGEX-6p-1 vector *E. coli* via a sticky end ligation method (2.2.5). The ligation suspensions were used to transform competent *E. coli* DH5 α strain, and transformants were selected in solid culture with ampicillin. The full-length HMGCR-pGEX-6p-1 plasmid was extracted from selected transformants after small scale LB or 2YT broth culture and analysed through restriction digest profiling with the same enzymes used for plasmid DNA cloning (2.2.4) and resolved by agarose gel electrophoresis (figure 3.7).

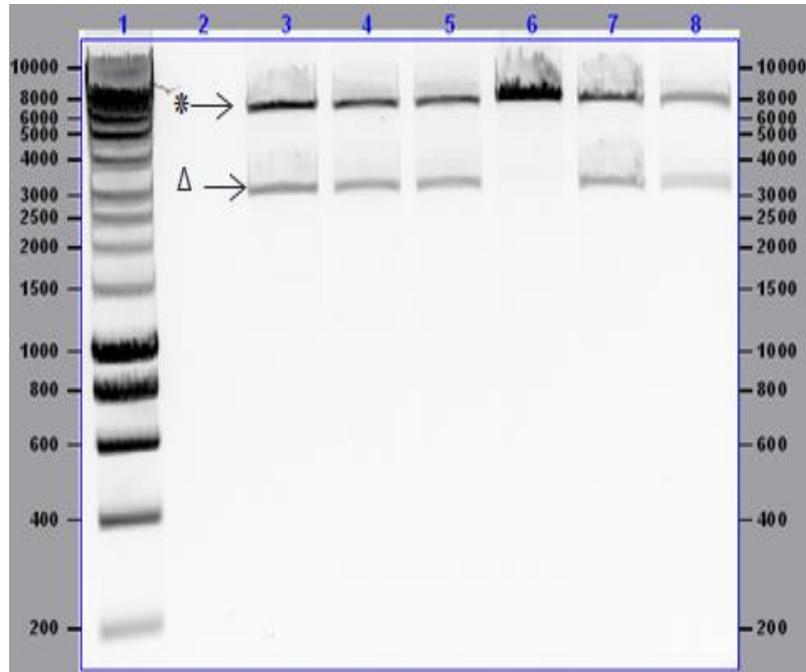


Figure 3.7: Cloning of full-length Wild-Type-HMGCR Gene into pGEX-6p-1 plasmid vector using DH5 α *E. coli* strain and evaluation by restriction digest.

Lane 1: showing hyper ladder 1kb and lane 3-8: various full-length Wild-type-HMGCR-pGEX-6p-1 constructs. Digested with *Bam*H1 and *Eco*R1 restriction endonuclease enzyme and resolved by 1% Agarose Gel Electrophoresis. * Linear, pGEX-6p-1 Δ full-length Wild-type HMGCR fragment.

The amplicon patterns generated in the agarose gel screening of figure 2.7 (Lane 3, 4, 5, 7, and 8) were compatible with expectations. The hype ladder 1kb band that migrates between 5000 and 6000bp markers showed to be consistent with pGEX-6p-1, which is a plasmid of 4984bp in length. The WT-HMGCR gene transformants all generated bands that migrated consistently with expectations, as defined in section 2.2.1., which implied that the WT-HMGCR gene, effectively cloned into the pGEX-6p-1 expression vector. However, nucleotide sequencing could not validate the successful cloning of the HMGCR gene.

3.5.3 HMGCR gene cloning and over-expression challenges and troubleshooting attempts

The cloning and overexpression of the recombinant full-length human HMGCR gene are vital to this study. However, earlier attempts made to clone and overexpress the gene have been unsuccessful due to a cascade of unexplainable events. A previous study though using a different expression vector called pGEX-cs plasmid had used similar methods as applied in this work to successfully clone and overexpress the catalytic domain of wild-type HMGCR gene (Eva *et al.*, 2000). Furthermore, there is no literature-based evidence to suggest that full-length HMGCR had been successfully cloned and expressed using a bacterial host expression system anywhere. This section, therefore, conveys the rationale behind the troubleshooting approaches applied to accomplish bacterial production of recombinant WT full-length HMGCR gene overexpression. The size of the full-length coding sequence of the HMGCR gene is about 2670 base pairs (Dong *et al.*, 2004). The production of recombinant HMGCR

protein via bacterial host system needed for this investigation; hence, the HMGCR gene was to be cloned into other bacterial expression (eg: pMAL-c5x expression vectors) and expressed using the DH5 α and C43 (DE3) *E. coli* cell.

3.5.4 Recombinant Overexpression of Full-length WT-HMGCR-GST Fusion Protein.

The HMGCR-pGEX-6p-1 plasmid clone could not be validated through DNA sequencing, the overwhelming evidence observed on the plasmid-DNA cloning and restriction digest agarose gel (figure 3.7) informed the decision to attempt protein overexpression of a selected clone further. At first, the recombinant HMGCR-pGEX-6p-1 plasmids were transformed into *E. coli* DH5 α strain, and later sub-cloned into an *E. coli* C43 (DE3) strain (figure 3.10). The heterologous protein overexpression of the HMGCR catalytic domain (Eva *et al.*, 2000) had applied the use of low temperature. Also, the overexpression media (2YT) was supplemented with 0.5M Sorbitol and 1mM Betaine consistent with other researchers' experiment protocol (Eva *et al.*, 2000). Therefore WT-HMGCR-GST overexpression culture was grown to a mid-log phase of 24°C and induced with 0.5mM IPTG after increasing to an induction temperature of 30°C for 2 hours duration. The SDS-PAGE analysis of clarified HMGCR-GST fusion protein overexpression lysates showed a non-definitive synthesis of HMGCR-GST recombinant protein was induced under these conditions when protein expression was induced using 0.5mM IPTG at 30°C (figure 3.9).

A review of the HMGCR-GST fusion overexpression SDS-PAGE gel at time (T2) figure 3.9, showed the precision plus band that migrates between 100 and 110kDa markers, which is not consistent with the expected recombinant HMGCR-GST fusion protein size of about 123kDa.

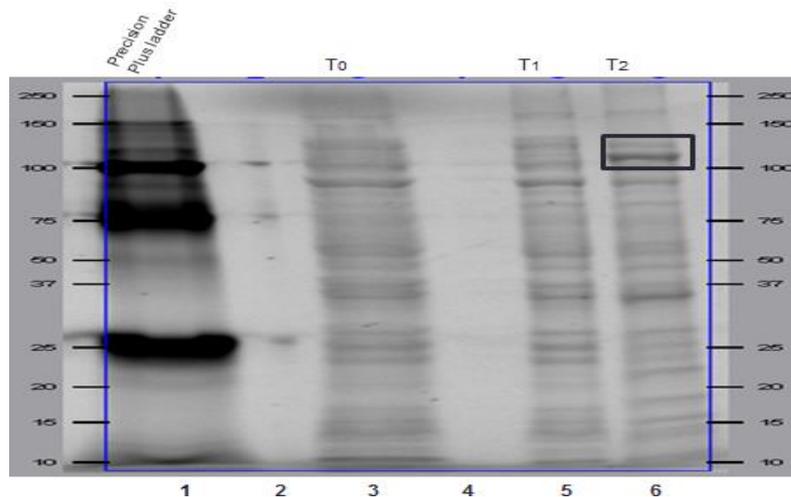


Figure 3.9 Protein Overexpression Analysis of Wild-Type HMGCR-GST fusion overnight protein expression in C43 (DE3) *E. coli* strain at 30°C in 0.5M Sorbitol and 1mM Betaine 2YT Media.
 Lane1: precision plus marker, Lane 3 – 8 Wildtype HMGCR-GST expression profiles showing post-IPTG (1mM) induction. T0 (L3), T1 (L5), T2 (L6). Resolved by SDS-PAGE methods at 150 Volts.

3.5.5 Comparative HMGCR-GST fusion protein Overexpression in DH5α *E. coli* strain and HMGCR-GST fusion protein expression in C43 (DE3) *E. coli* strain

In other to investigate the effect of different *E. coli* strains, or the overexpression ability of the non-verified HMGCR-pGEX-6p-1 plasmids. Two lots of HMGCR-pGEX-6p-1 plasmids were used for comparative HMGCR-GST overexpression analysis. The first was transformed by been transformed using a competent DH5α *E. coli* strain, and the second was transformed by being transformed using a competent C43 (DE3) *E. coli* strain (figure 2.9) respectively and both preserved in absolute glycerol. The synthesis of HMGCR-GST fusion protein utilised a yeast extract media (2YT) supplemented with 0.5M Sorbitol and 1mM Betaine. The two-glycerol preserved HMGCR-pGEX-6p-1

transformed plasmids (BL21 (DE3) and C43 (DE3)) where cultures were grown to a mid-log phase of 37°C and induced with 1mM IPTG respectively. Overexpression profiling was analysed by withdrawing a sample from both expressions, at 1 hour, 2hours, and overnight after IPTGS induction, at a temperature of 30°C. The Coomassie blue-stained and stain-free SDS-PAGE analysis, of clarified 100% glycerol preserved HMGCR-GST fusion protein overexpression lysates, in BL21 (DE3) and C43 (DE3) respectively. A review of the HMGCR-GST fusion overexpression SDS-PAGE gel (Coomassie blue-stained) for DH5α *E. coli* strain (figure 2.11a lane 3-6), showed no sign of successful protein overexpression (figure 2.11b lane 6). Although there was a migration of the precision plus bands between 37 and 50kDa position, it was not consistent with the expected recombinant HMGCR-GST fusion protein size of about 123kDa. Similarly, the HMGCR-GST fusion overexpression for C43 (DE3) *E. coli* strain showed that the putative HMGCR-GST fusion protein band migrated between 20 and 25kDa markers, which is not consistent with the expected recombinant HMGCR-GST fusion protein size of about 123kDa but relatively consistent with GST protein size of about 26kDa.

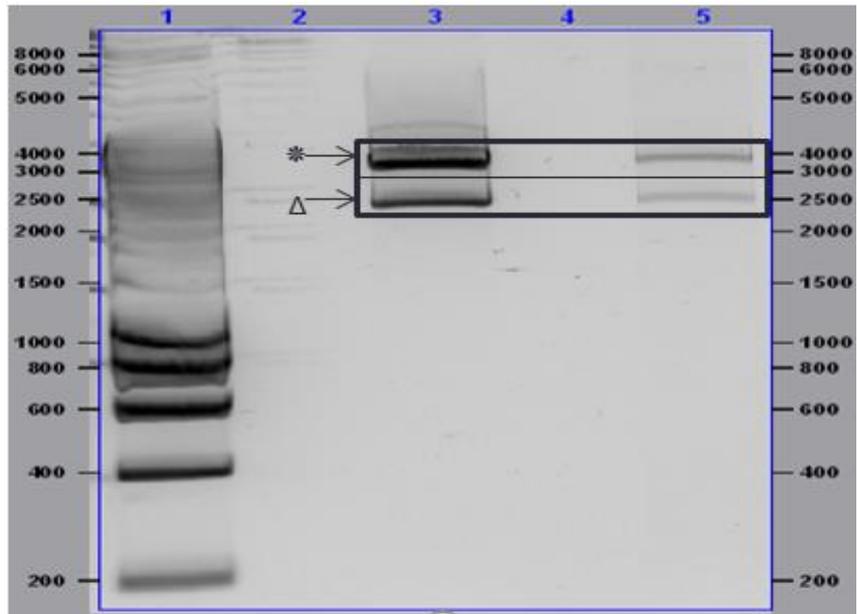


Figure 3.10: Comparative Wild-Type Recombinant HMGCR-pGEX -6p-1 plasmid in DH5 α Strain and HMGCR-pGEX -6p-1 plasmid Sub-cloned C43 (DE3) Strain.

Lane 1: showing hyper ladder 1kb and lane 3-5: various. Wild type-HMGCR-pGEX-6p-1 constructs. Digested with BamH1 and EcoR1 restriction endonuclease enzyme and resolved by 1% Agarose Gel Electrophoresis.

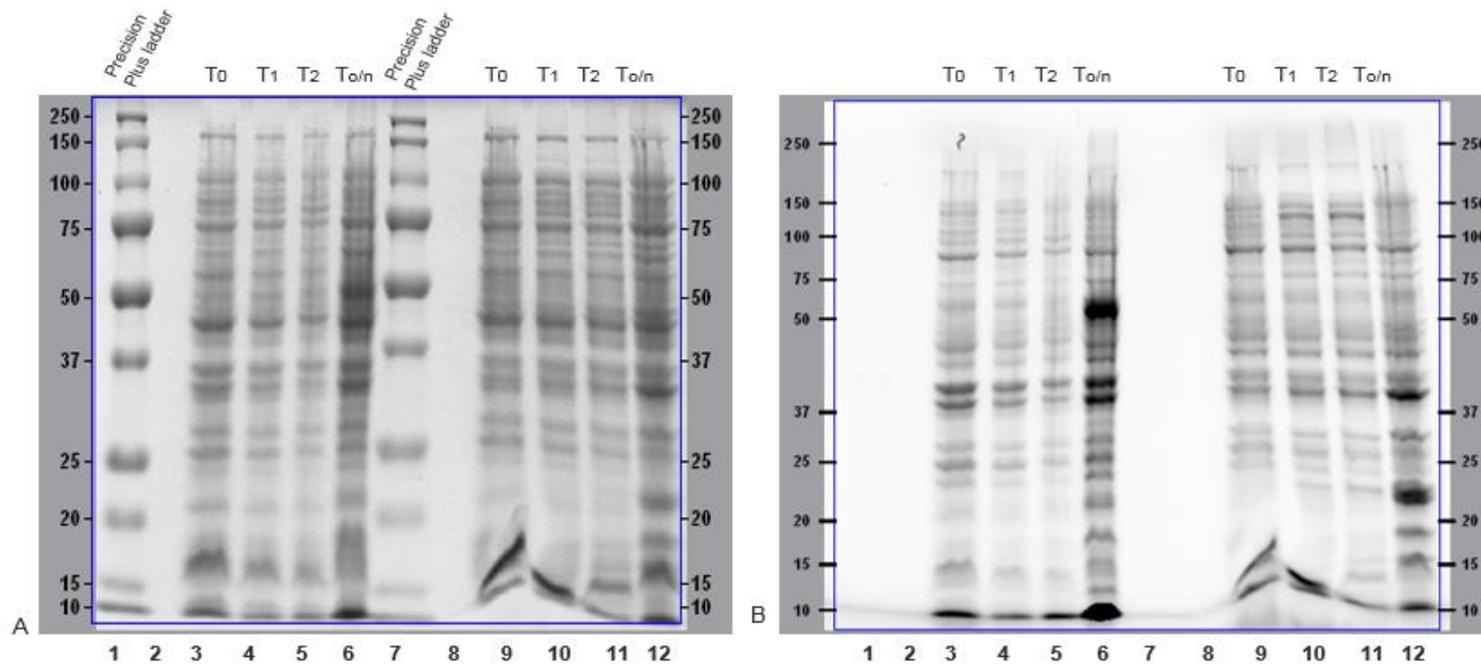


Figure 3.11 Comparative Protein Overexpression Analysis of an HMGCR-GST fusion protein in DH5α *E. coli* strain and HMGCR-GST fusion protein expression in C43 (DE3) *E. coli* strain with glycerol preservation in 0.5M Sorbitol and 1mM Betaine 30°C in 2YT media.
 (A) Lane 1: Precision plus marker, Lane 4 – 6, showing Glycerol preserved HMGCR-GST fusion protein expression profile post IPTG(1mM) induction in DH5α *E. coli* strain. T0 (L3), T1 (L4), T2 (L5), To/n (L6). Lane 7: Precision plus marker, Lane 10 – 12, showing Glycerol preserved HMGCR-GST fusion protein expression profile post IPTG (1mM) induction in C43 (DE3) *E. coli* strain. T0 (L9), T1 (L10), T2 (L11), To/n (L12). **(B)** Lane 3 – 6, showing coomassie blue stained gel of glycerol preserved HMGCR-GST fusion protein expression profile post-IPTG (1mM) induction in DH5α *E. coli* strain. T0 (L3), T1 (L4), T2 (L5), To/n (L6). Lane 9 – 12, showing Glycerol preserved HMGCR-GST fusion protein construct-5 expression profile post IPTG (1mM) induction in C43 (DE3) *E. coli* strain. T0 (L9), T1 (L10), T2 (L11), To/n (L12). Resolved by Coomassie blue-stained and Stain-free SDS-PAGE methods, respectively, in 150 Volts.

Based on the results of this comparative DH5 α and C43 (DE3) *E. coli* transformed HMGCR-GST tag fusion protein expression, it could be concluded that it may be, that this overexpression machinery is unable to synthesize a full-length HMGCR-GST fusion protein. The result (figure 3.10) shows the outcome of different overexpression profiles in the two different *E. coli* strains, although figure 3.9 previously showed the recombinant HMGCR-pGEX-6p-1 plasmid-DNA as being consistent with the expected amplicon sizes after gene cloning. Additionally, the inability to obtain an accurate recombinant HMGCR-pGEX-6p-1 plasmid (figure 2.7), through gene sequencing, confirms the hypothesis that wild-type recombinant HMGCR-pGEX-6p-1 plasmid, was unable to be cloned and protein synthesis through bacteria overexpression could not produce the expected HMGCR-GST recombinant protein.

3.5.6 Solubilisation and Refolding of HMGCR-GST Fusion Protein

Protein overexpression lysates from HMGCR-GST fusion protein overexpression in C43 (DE3) *E. coli* strain were solubilized separately, using 6M urea and solubilization detergent containing 8M urea, 0.5% Tween 20, 10mM DTT in 20mM Tris pH nine according to the protocol described in section 2.3.5., as shown in figure 2.11 and 2.12. A review of the HMGCR-GST fusion overexpression SDS-PAGE gel (Coomassie blue stained and stain-free gels) figure 2.11 lane 8 and 9, showed a relative expression that is not consistent with the expected recombinant HMGCR-GST fusion protein and even clearer on the stain-free SDS-PAGE gel (figure 2.10b lane 6). The fusion protein band migrated

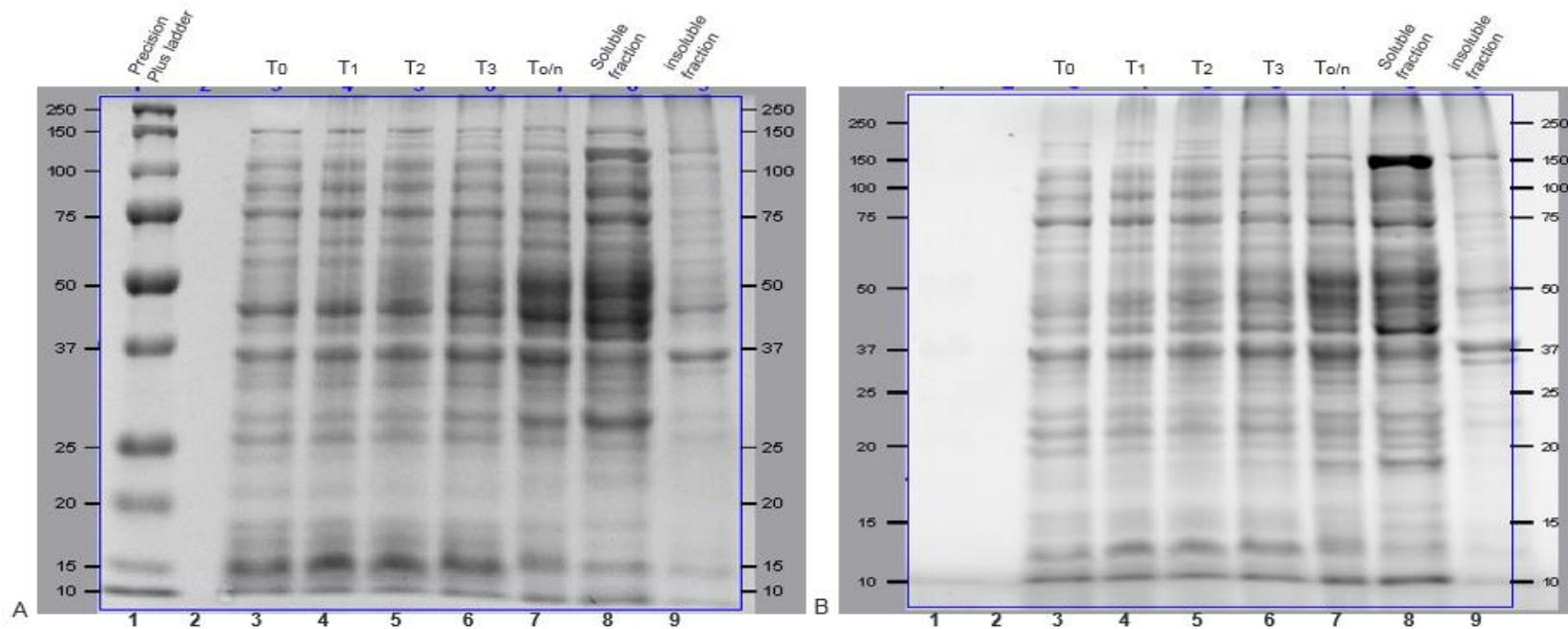


Figure 3.12 Protein Solubilisation Analysis of HMGCGR-GST fusion protein in C43 (DE3) *E. coli* strain 37°C in LB media.
 (A) Lane 1: Precision plus marker, Lane 4 – 7, HMGCGR-GST fusion protein expression profile post IPTG (1mM) induction in C43 (DE3) *E. coli* strain. T0 (L3), T1 (L4), T2 (L5), T3 (L6), To/n (L7), soluble fraction (L8) insoluble fraction (L9) (B) coomassie blue stain of 2.11a: Lane 4 – 6, showing HMGCGR-GST fusion protein expression profile post IPTG (1mM) induction in C43 (DE3) *E. coli* strain. T0 (L3), T1 (L4), T2 (L5), T3 (L6), To/n (L7), soluble fraction (L8), insoluble fraction (L9). Resolved by Coomassie blue-stained and stain-free SDS-PAGE methods, respectively, in 150 Volts.

between 37 and 50kDa markers, which is not consistent with the expected recombinant HMGCRC-GST fusion protein size of about 123kDa. Evaluating the C43 (DE3), *E. coli* transformed wild-type HMGCRC-GST fusion protein expression profile in figure 2.11. The presences of dense (but not so clear) layouts on lane 6: T3, lane 7: To/n, and lane 8: soluble fraction indicates some extent of migration between the 37kDa – 50kD, similarly, the presence of an amplified layout on lane 8: soluble fraction indicated a migration between the 100kDa – 150kDa of the precision plus ladder. The observed amplification on lane 6 and 7 is not consistent with the expected size, based on the expression profile of a recombinant HMGCRC-GST fusion protein. However, the amplification on lane 8 (soluble fraction) migrating between 100kDa – 150kDa, showed a protein size of about 120kDa. This size closely relates to the expected size of wild-type recombinant HMGCRC- GST protein of 123kDa. However, it may not be accurate to claim that figure 2.11, lane 8 amplification was a synthesized recombinant wild type HMGCRC-GST protein, since the HMGCRC-pGEX-6p-1 plasmid clones and were unable to be validated through protein sequencing.

3.6.0 AGTRI Protein Synthesis

This section presents most of the laboratory work that had been carried out in order to address the limitations of AGTR1-GST protein expression. Previously, the wild-type angiotensin receptor 1 (WT-AGTR1) gene, had been studied and successfully cloned into *E. coli* pGEX-6p-1 (Figure 1.2.1), pMAL-p5x and pMAL-c5x *E. coli* vector systems. The pGEX-6p-1 vector and pMAL-c5x vectors encode for GST gene fusion protein of 26kDa size and MBP fusion protein of 42.5kDa size respectively (Guan *et al.*, 1987). These two vectors are both used for expression, detection, and purification of their respective tagged proteins produced in *Escherichia coli* cells (Maina *et al.*, 1988). In the context of this work, AGTR1 was thoroughly attempted to be synthesized using GST, MBP and His tag fusion protein, with each having a higher degree of genetic modification and promoter, whose features were exploited to synthesise AGTR1 in *E. coli* expression system.

The pGEX-6p-1 bacterial vector constitutes of a *tac* promoter for chemically inducible, high-level expression of GST-tagged recombinant proteins. Towards its application for bacteria protein overexpression, pGEX-6p-1 plasmid vector, contains internal *lacI^q* gene for use in any *E. coli* host (Dilworth *et al.*, 2018). The pGEX being a host to a hybrid *tac/lac* promoter exhibit both an IPTG inducible relatively weak promoter characteristics and for its *tac* promoter features, a hybrid promoter between Trp and Lac promoter, semi-strong in gene expression regulation (Dilworth *et al.*, 2018).

The use of DH5 α and C43 (DE3) *E. coli* strains respectively, for the overexpression of AGTR1 protein synthesis are reported in the chapter as well. Recombinant AGTR1 plasmids were verified using DNA sequencing methods. The results obtained, validated that recombinant AGTR1- pGEX-6p-1, AGTR1- pET32a and AGTR1- pMAL-c5x constructs were within frame the ORF of the vectors.

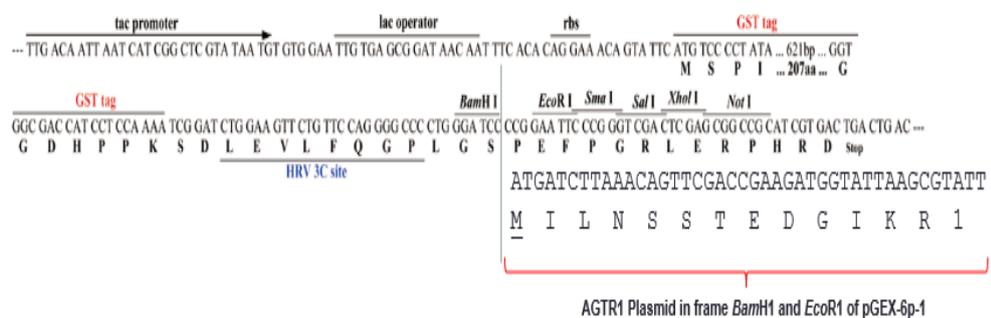


Figure 3.14 Sequence Verified AGTR1-pGEX-6p-1 plasmid-DNA Polylinker
Showing AGTR1 plasmid fitted between *Bam*H1 and *Eco*R1 within MCS frame of the pGEX-6p-1 overexpression vector.

3.6.1 WT-AGTR1 Cloning into *E. coli* Vector pGEX-6p-1

A synthetic WT-AGTR1 plasmid was commercially acquired (IDT UK). The WT-AGTR1 gene was retrieved from the commercial vector using a preliminary restriction digestion reaction. After that, the WT-AGTR1 plasmid amplicon was extracted and cloned into a pGEX-6p-1 vector via sticky end ligation (2.2.5). The ligation suspensions were used to transform competent *E. coli* BL21 (DE3) strain. The AGTR1-pGEX-6p-1 was extracted from selected transformants after small 2YT broth culture and analysed through restriction digest profiling *Eco*R1 and *Bam*H1

restriction endonuclease enzymes (2.2.4). The AGTR1-pGEX-6p-1 cloning was further analysed using agarose gel electrophoresis methods (figure 3.15).

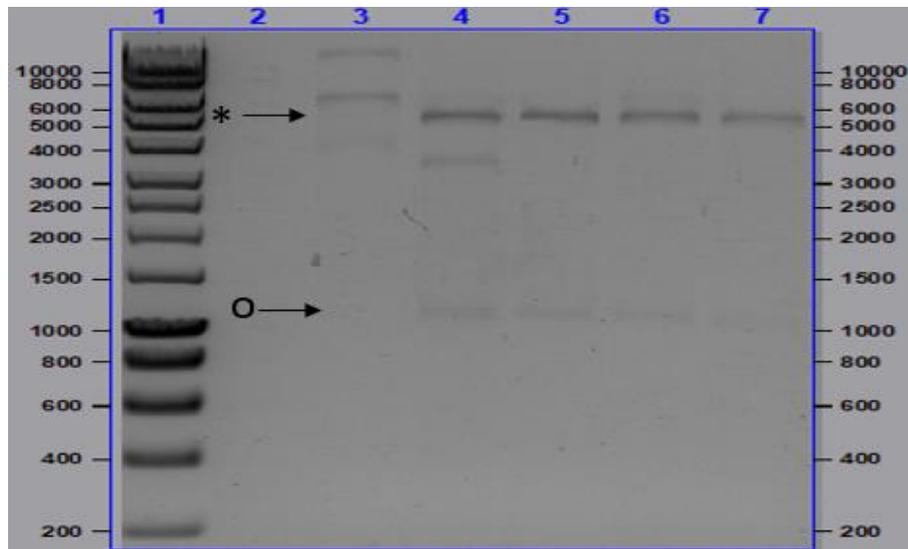


Figure 3.15: Cloning of full-length Wild-Type-AGTR1 Gene into a pGEX-6p-1 plasmid vector using DH5 α *E. coli* strain and evaluation by restriction digest.

Lane 1: showing hyper ladder 1kb and lane 3-7: various full-length Wild-type-HMGCR-pGEX-6p-1 constructs. Digested with *Bam*H1 and *Eco*R1 restriction endonuclease enzyme and resolved by 1% Agarose Gel Electrophoresis. * Linear, pGEX-6p-1, and O full-length Wild-type AGTR1 fragment.

The amplicon patterns generated in the agarose gel screening of figure 3.15 (Lane 3, 4, 5, 7, and 8) were compatible with expectations. The marker ladder 1kb band that migrates between 5000 and 6000bp markers showed to be consistent with pGEX-6p-1, which is a plasmid of 4984bp in length. The WT-AGTR1 gene transformants on lane 4-7 of figure generated bands that migrated consistently with expectations as defined in section 3.2.1. The result shows that the WT-AGTR1 gene effectively cloned into the pGEX-6p-1 expression vector. Further recombinant plasmid sequencing validated that the WT-AGTR1-pGEX-6p-1 construct successfully cloned.

3.6.2 Overexpression of recombinant WT-AGTR1-GST fusion Protein

Further attempts to overexpress recombinant AGTR1-pGEX-6p-1 construct, by transformed the recombinant plasmid into an improved overexpression *E. coli* strain BL21 (DE3) at 37°C was unsuccessful figure 3.16. What remained unclear was the specific reason behind AGTR1-GST recombinant events, which occurred against time (post-induction). This unresolved expression (figure 3.16) and AGTR1 gene termination (figure 3.17 lane 3 and 4) were observed in 6 hours post IPTG induction plasmid stability assay using BL21 (DE3) *E. coli* strain at 37°C in an AGTR1-GST fusion protein overexpression.

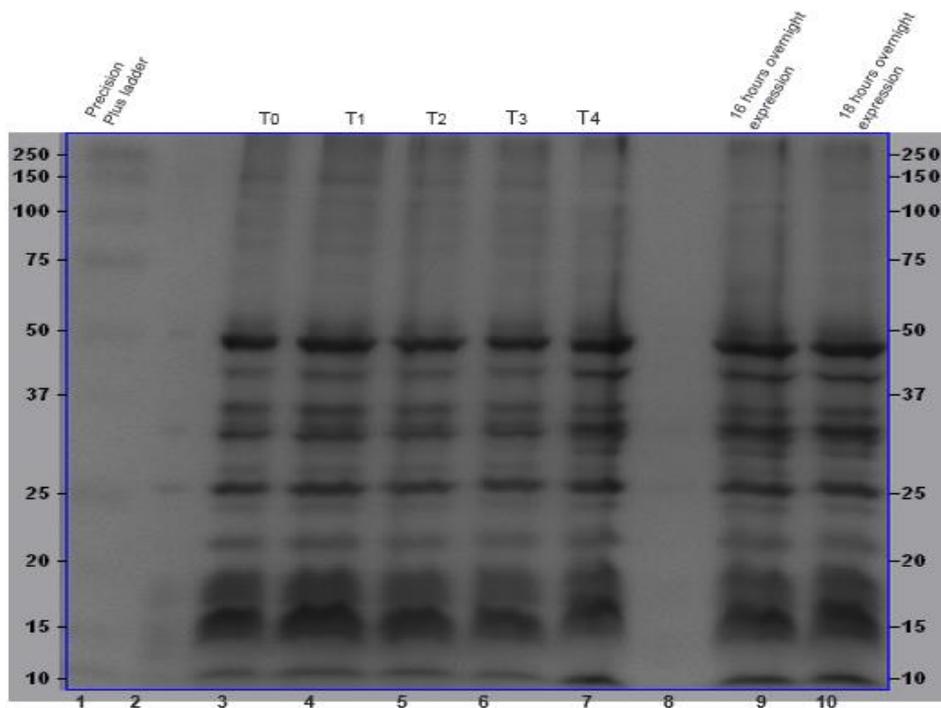


Figure 3.16: Protein Overexpression Analysis of AGTR1-GST Tag fusion protein Expression in BL21 (DE3) *E. coli* strain at 28°C in LB media:

Lane 1 – 10 showing expression profile post-IPTG (1mM) induction T0 (L1), T1 (L2), T2 (L3), T3 (L4), T4 (L5), 16 hours overexpression (L9), 18 hours overexpression (L10). Resolved by stain-free SDS-PAGE method.

Below are the troubleshooting approaches that were applied to improve on the apparent lack of AGTR1-fusion protein overexpression, including the Introduction of pET32a vector and other protein expression parameter interventions. It is crucial to highlight that there was a lack of success in synthesizing recombinant AGTR1-fusion protein, impacted on the continuity of this project.

3.7.0 AGTR1 over-expression challenges and troubleshooting attempts (problem-solving attempts)

3.7.1 Recombinant protein AGTR1 Overexpression using GST-fusion protein

The estimated AGTR1-GST fusion protein size should be about 67KDA, which constitutes the 41,061 Daltons being the size of AGTR1 protein and 26,000 Daltons being the size of GST protein. The recombinant AGTR1-pGEX-6p-1 plasmid was initially used to transform DH5 α *E. coli* strain and subsequently transformed into BL21 (DE3) *E. coli* strain according to methods described in section 4.2.7. In earlier AGTR1-GST fusion protein overexpression assays, 37°C was used as induction temperature without successful protein production. The overexpression condition of the AGTR1-GST fusion protein synthesis was optimized by growing culture to mid-log phase and inducing with 0.5mM IPTG at an induction temperature of 28 °C for 18 hours post-induction. SDS-PAGE Electrophoretic analysis of clarified lysates prepared from recombinant AGTR1-GST overexpression showed that there was no overexpression of AGTR1-GST fusion recombinant protein (shown in figure 3.17).

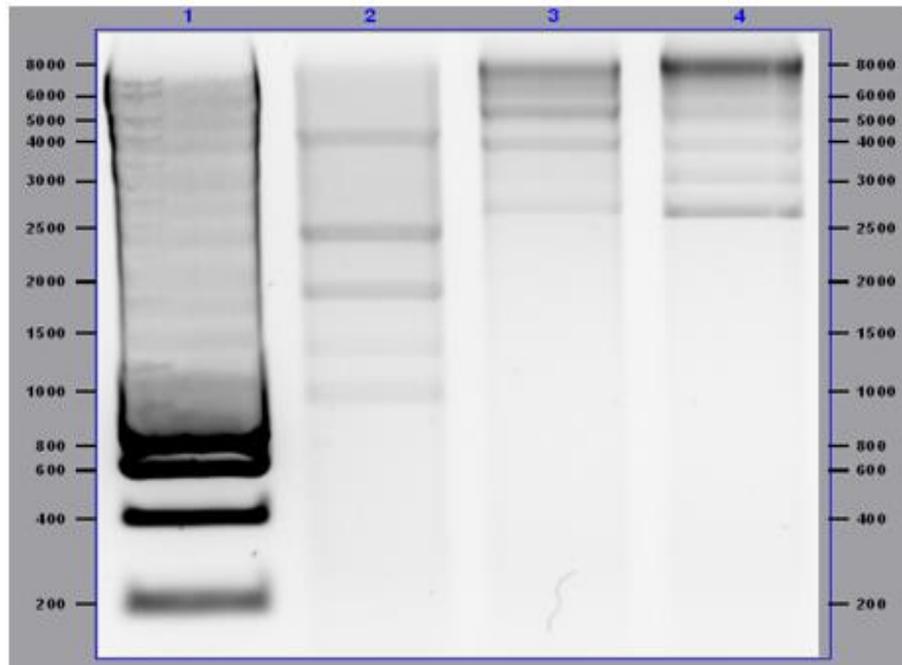


Figure 3.17: Plasmid Stability Analysis of the Wild-Type AGTR1-GST Fusion Overexpression in BL21 (DE3) *E. coli* Strain at 37°C in 2YT Media.

Lane 1: showing hyper ladder 1kb and lane 2-4: various AGTR1-GST fusion protein of 2 - 6 hours post 1mM IPTG induction lysates using 2YT media. Lane 2: AGTR1-GST fusion protein at 2hours post-induction, Lane 3: AGTR1-GST fusion protein at 4hours post-induction, and Lane 4: AGTR1-GST fusion protein at 6hours post-induction. Digested with *Bam*H1 and *Eco*R1 restriction endonuclease enzyme and resolved by 1% Agarose Gel Electrophoresis at 150Volts.

Further plasmid stability analysis, using *Bam*H1 and *Eco*R1 digested post-induction lysates, of recombinant WT AGTR1-GST fusion protein at 37°C, which were taken at a different time interval. Revealed that AGTR1-GST fusion protein expression at 2 hours post-induction (lane 2 figure 3.17), may have been truncated as the bands observed were not consistent with the expected recombinant AGTR1-pGEX-6p-1 plasmid size. The complete disappearance of the AGTR1-protein band of the recombinant AGTR1-GST fusion protein lysate (lane 3 and 4 of figure 3.17) shows that AGTR1-protein within the pGEX-6p-1 (GST-fusion) may have been eliminated because of series of unexplainable transcription and translation events in pGEX-6p-1 plasmid in the BL21 (DE3) *E. coli* strain. The GST-tag fusion protein (26kDA) earlier used in AGTR1 fusion protein expression had shown no success. The major factors

considered that could be limiting AGTR1 fusion protein overexpression may be due to the AGTR1 gene's characteristic features as a transmembrane GPCR protein. Hence the AGTR1-GST fusion protein may be undergoing leaky expression and other unknown recombinant events. These events may have led to truncation (Figure 3.17) of the fusion protein in the tac-promoter of the pGEX-6p-1 vector during bacterial overexpression (Frangioni *et al.*, 1993).

Additionally, GPCRs are also widely noted for their difficulty to overexpress in bacterial host systems (Saida *et al.*, 2006). The pGEX-6P-1 expression system is known to have a less tightly regulated overexpression of chemically inducible toxic genes during bacterial overexpression (Rathore and Gautam, 2014). Protein synthesis results for AGTR1 gene overexpression, suggests, that bacterial system probably may not be suitable for recombinant protein synthesis of AT1-receptor proteins. Shukla and co-workers reported, other overexpression systems, like the mammalian system and baculovirus/insect system, have previously overexpressed AGTR1 protein for functional studies (Shukla *et al.*, 2009). Unfortunately, those kinds of overexpression systems produced very low quantities of recombinant AGTR1 proteins at a high cost and taking a lot of time. Therefore, to optimize wild-type AGTR1 fusion protein overexpression, using conditions that may enhance gene expression of sequenced verified wild-type AGTR1-plasmid genes, formed the primary focus of AGTR1 further studies after the last review. Systematically, the effects of following overexpression elements were studied closely ranging from the application of different overexpression vectors, host cell line, induction temperatures, and AGTR1-host cell toxicity assay.

3.7.2 Optimization of AGTR1 Overexpression Parameters

3.7.2.1 Vectors

The choice of a cloning vector in the recombinant *E. coli* overexpression system is fundamental to successful overexpression. Toxic or hard to express genes, require tight regulation during protein overexpression. Consequently, the use of tightly regulated hybrid promoters (T7lac, ARA_pro) is known to ease the difficulty in expressing very toxic proteins (Saida *et al.*, 2006). Different bacterial overexpression vectors (fusion tags) have been produced, to enhance protein overexpression, improve protein folding, solubility, and ease the purification processes of some toxic genes. These vectors are often genetically modified in their promoters and selective marker regions, to enhance tighter regulation of protein expression as obtained with the pET32a vector systems and to enhance protein folding as obtained with pMAL vector systems. The pET32a vector system (His-tag) has a molecular size 1kDa and contains a T7lac promoter, while the pMAL-p5x vector system (MBP-tag) has a molecular size of 42.5kDa and consists of a P_{tac} promoter.

Protein expression vector-based optimization of AGTR1 fusion protein includes the expression of recombinant AGTR1-pET32a and AGTR1-pMAL-p5x plasmids. The estimated size of AGTR1-HIS and AGTR1-pMBP fusion proteins should be about 42kDa and 83.5kDa respectively. The AGTR1 protein constitutes 41 kDa in both fusion protein complexes, respectively.

3.7.2.2 Cloning of WT-AGTR1 into *E. coli* Vector pET32a

The open reading frame encoding the codon-optimized AGTR1 gene was amplified with the signal peptide region for cloning pET32a vector for production with a C-terminal hexahistidine tag. The amplified was purified and digested using restriction endonucleases *Bam*H1 and *Eco*R1 and extracted from agarose. The purified and digested AGTR1 gene was ligated to pET32a (also digested with *Bam*H1 and *Eco*R1) using sticky end ligation, as described in the methods section. The ligation mixture was used to transform competent *E. coli* BL21 (DE3). The plasmid DNA was recovered from ampicillin resistant transformant and verified for the release of the AGTR1 gene (migrating consistent with the expected 1100bp of AGTR1 fragment) as shown in (Figure 1.2.2) after digestion with *Bam*H1 and *Eco*R1 and resolved using electrophoresis. The product after that was gel extracted and purified from the agarose gel and subjected to restriction digestion for cloning into the pET32a vector.

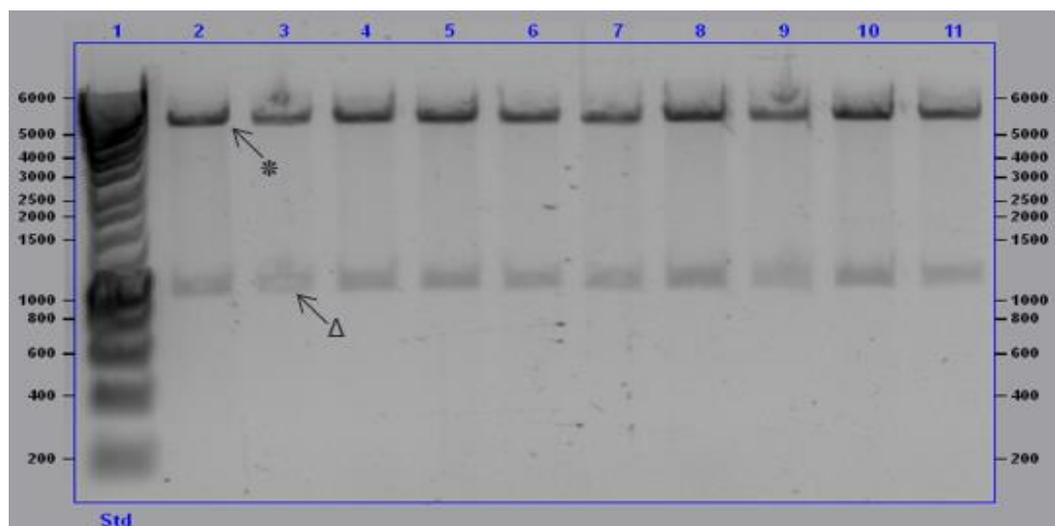


Figure 3.18: Restriction Digestion Analysis of Recombinant pET32a ligated with AGTR1

Lane 1: showing hyper ladder 1kb and lane 2-11: various AGTR1-pET32a constructs. Digested with *Bam*H1 and *Eco*R1 restriction endonuclease enzyme and resolved by 1% Agarose Gel Electrophoresis. * Linear, pET32a fragments, O Linear AGTR1 gene.

Nucleotide sequence verification of selected AGRT1-Pet32a clones is present below (Figure 1.2.3).

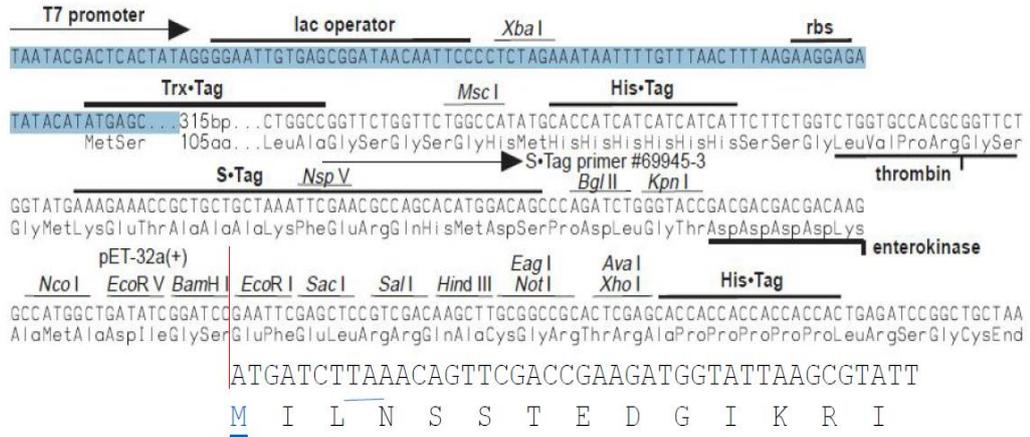


Figure 3.19: Sequence Verified AGRT1-pET32a plasmid-DNA Polylinker. Showing AGRT1 plasmid fitted between *Bam*H1 and *Eco*R1 within MCS frame of the PET32a overexpression vector in BL21 (DE3) *E. coli* strain.

3.7.2.3 Cloning of WT- AGTR1 into *E. coli* Vector pMAL-p5x

The amplified codon-optimized AGTR1 gene was cloned into the pMAL-p5x vector down-stream of the *malE* gene, which encodes the maltose-binding protein.

3.7.3 AGTR1-Fusion Protein Overexpression Temperature Viability Assay

Heterologous recombinant protein production in *E. coli* depends essentially on temperature, among many other factors (Weickert *et al.*, 1996). Induction temperature influences the process of protein folding and stability. In *E. coli* recombinant protein expression, induction temperature has a major impact on the production of toxic or difficult to express recombinant proteins. A few studies have shown that high overexpression yield of toxic or difficult to express proteins are achievable when overexpressing at low induction temperature ranging from 4°C - 20°C (Teresa *et al.*, 2013, Rosano and Ceccarelli 2014).

Further analysis of recombinant AGTR1-fusion protein overexpression in this work, considered the application of low temperature. The sequence-verified recombinant plasmid AGTR1-pET32a was used to transform BL21 (DE3) *E. coli* strain and previously overexpressed of glycerol-preserved sequence-verified AGTR1-pET32a plasmid at 37°C using 1mM IPTG induction and 3 hours post-induction, showed no synthesis of AGTR1-His-Tag fusion protein. Comparatively, the same plasmid was overexpressed for 23 hours using 1mM IPTG induction after cooling from 37°C to 17°C induction temperature. Upon resolving the clarified lysate by SDS-PAGE electrophoretic analysis (figure 3.21), the result showed AGTR1His-tagged protein synthesis was unsuccessful. Hence the reduction of

induction temperature may not be a sufficiently independent factor for the optimization of AGTR1-His-Tag protein synthesis through bacteria overexpression. However, induction temperature for bacterial protein expression has been found to have a pronounced effect on protein stability and folding (Weickert *et al.*, 1996). Consistent with other research opinions (Natalia and Salvador, 2006), Perez-Bermudez *et al.* (2010) studied the effect of various temperatures various during recombinant protein production in *E. coli* and concluded that the highest soluble recombinant protein yield was acquired within 4°C and 15°C over 20 hours. Low induction temperatures were found to aid overexpression yield and solubility in the production of hard to express proteins (Chi-Yen *et al.*, 2015). Also, the reduction of overexpression temperature, have become a widespread strategy for minimizing aggregate formation during recombinant protein production of polypeptides and membrane proteins (Vasina and Baneyx, 1996). For toxic membrane protein MBP-P5βR2, optimal overexpression temperature was obtained at 15°C and 4°C overexpression at 18 hours and 72 hours, respectively (Teresa *et al.*, 2013).

3.7.4 Comparative Expression of AGTR1-HIS tagged fusion protein at 37°C and 17°C

Two respective colonies of AGTR1-pET32a plasmid-DNA sequencing results showed to be positive, containing properly cloned recombinant AGTR1-pET32a plasmid-DNA, which indicates, that AGTR1 genes were in frame with no mutation. The recombinant AGTR1-pET32a plasmids had been previously cloned using a DH-5α *E. coli* strain, and now, we're used to retransforming *E. coli* BL21 (DE3) cells.

Preliminary cell culture, of the selected AGTR1-pET32a colonies, were grown under the following conditions, 10mL yeast extract broth containing 10 μ L of ampicillin for antibiotic resistance, incubated overnight at 37 $^{\circ}$ C, shaking at 200rpm. The protein overexpression of two respective colonies AGTR1-HIS tagged fusion protein at 37 $^{\circ}$ C and 17 $^{\circ}$ C was initiated by inoculating a 150mL of yeast extract broth containing 150 μ L of ampicillin with 2mL of overnight grown samples and grown until cultured attained 0.7 optimal densities relative to OD600. After that, protein overexpression was induced by adding 1mM IPTG (isopropyl β -D-thiogalactopyranoside) after pre-induction expression lysates were withdrawn at times zero (T0). Protein expression was observed for a period of 3hours with sample withdrawn every hour intermittently, after which incubation could continue 23 hours post-overnight (figure 1.3). In other to further understand the effect of different overexpression temperatures, in the recombinant production of AGTR1-HIS fusion proteins. Two lots of AGTR1-pET32a colonies were expressed at different temperatures, after OD600 at 37 $^{\circ}$ C (figure 1.3) and 17 $^{\circ}$ C (figure 1.4) respectively and used for comparative AGTR1-HIS overexpression analysis.

A further review of the 17 $^{\circ}$ C AGTR1-HIS, tagged fusion overexpression SDS-PAGE gel (Coomassie blue-stained) for BL21 (DE3) *E. coli* strain (figure 3.21), showed that at overnight expression (lane 7-10), there appeared to be a degree of non-confirmed expression. However, the fusion protein band migrate between 50 and 75kDa markers showed the expression not to be consistent with the expected recombinant AGTR1-HIS fusion protein size of about 50kDa.

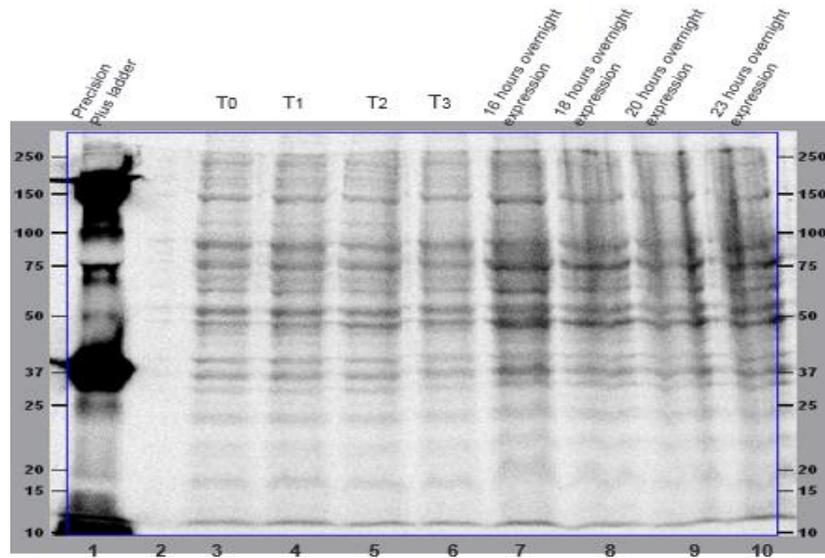


Figure 3.21: Protein Overexpression Analysis of AGTR1-HIS Tag fusion construct protein expression in BL21 (DE3) *E. coli* Strain at 17°C in LB media.

Lane 1: Precision plus marker, Lane 3 – 10, showing AGTR1-HIS Tag fusion protein expression profile post IPTG (1mM) induction. T0 (L3), T1 (L4), T2 (L5), T3 (L6), 16 hours overexpression (L7), 18 hours overexpression (L8), 20 hours overexpression (L9), and 23 hours overexpression (L10). It is resolved by Coomassie blue-stained SDS-PAGE methods at 150 Volts.

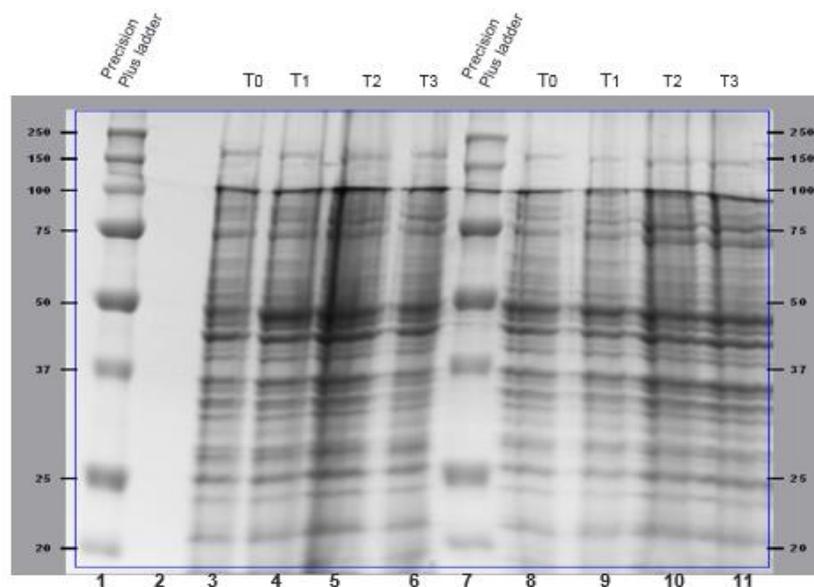


Figure 3.22: Protein Overexpression Analysis of AGTR1-HIS Tag fusion protein construct-1 expression in BL21 (DE3) *E. coli* strain with and without glycerol preservation at 37°C in LB media.

Lane 1: Precision plus marker, Lane 3 – 6, showing Glycerol preserved AGTR1-HIS Tag fusion protein construct-1 expression profile post IPTG (1mM) induction. T0 (L3), T1 (L4), T2 (L5), T3 (L6). Lane 7: Precision plus marker, Lane 8 – 11, showing non-glycerol preserved AGTR1-HIS Tag fusion protein construct-1 expression profile post IPTG (1mM) induction. T0 (L8), T1 (L9), T2 (L10), T3 (L11), Lane 12: Precision plus marker. It is resolved by Coomassie blue-stained SDS-PAGE methods at 150 Volts.

The AGTR1-pET32a plasmid stability was also investigated simultaneously by analysing the effect of glycerol preserved recombinant AGTR1-pET32a plasmid against the non-glycerol preserved AGTR1- pET32a plasmid DNA.

The glycerol preserved and non-glycerol preserved AGTR1-pET32a, BL21 (DE3) transformed plasmid-DNA were cultured and grown to a mid-log phase of 37°C and induced with 1mM IPTG at 37°C respectively. The protein expression profile was analysed by withdrawing a sample from expressions, 1 hour, 2hours, and overnight post-induction (figure 3.22). The Coomassie blue-stained and stain-free SDS-PAGE analysis, of clarified glycerol, preserved and non-glycerol preserved AGTR1-HIS tagged fusion protein overexpression lysate in *E. coli* BL21 (DE3), showed a no synthesis of a protein not consistent with the expectation of AGTR1-HIS recombinant protein under these conditions when protein expression was induced using 1mM IPTG at 37°C (figure 1.4).

3.7.5 Optimizing Choice of Bacteria Host Cell Strain

Unlike most soluble proteins, membrane proteins (GPCR like proteins) are generally introduced into bacteria host membranes to adopt their appropriate conformation (Ren *et al.*, 2009). This effectively induces toxicity to the host cell during protein overexpression. Hence of empirical importance is the choice of the most suitable host strain appropriate for the overexpression of the toxic gene. Routinely, the first line *E. coli* competent cell used in this research was a DH5 α competent *E. coli* cell. This *E. coli* strain is primarily known for its ability to ensure

insert stability, due to a controlled resistance for heterologous recombination in the presence of its *recA* mutation.

Additionally, DH5 α is commonly deficient in plasmid degrading endonucleases, whose activities digest the plasmids through an isolation process. Gene constructs were sub-cloned into BL21 (DE3) competent *E. coli* cell. BL21 (DE3) is known for its lack of Lon protease and ompT outer membrane protease; hence, degrading recombinant proteins during purification (Grodberg and Dunn, 1988). Also, Studier F.W. and Moffat B.A (1991), reported that T7 *gene 1* of BL21 (DE3) expression strain was found to be regulated by a promoter *lacUV5* in a poorly modulated manner under repressive conditions. Ideally, protein expression in the *E. coli* BL21 (DE3) work by the incorporating lambda DE3 with the gene which encodes the T7 RNAP, under the control of the *lacUV5* promoter. Upon induction of IPTG (isopropyl β -D-thiogalactopyranoside), a non-metabolized analog of lactose, the T7 RNAP precisely transcribes the gene interest inserted in the T7 expression plasmid downstream of the T7 promoter. The mRNA of the gene of interest is highly expressed, thereby producing the desired protein (Augius *et al.*, 2016). In the case of toxic genes or hard to express proteins, mRNA of the gene of interest, overloads the translation machinery, which often results to ribosome destruction (Dong *et al.*, 1995).

However, consistent with AGTR1-fusion protein's high toxicity in the bacterial cells, the use of BL21 (DE3) *E. coli* strain was unable to demonstrate successful recombinant AGTR1 fusion protein overexpression. This observation was consistent with an earlier similar work by Laurence *et al.*, (2004), where a comparison of toxic recombinant proteins overexpression was assayed among

BL21 (DE3), C41 (DE3) and C43 (DE3) *E. coli* cells respectively. The study findings showed that the transformation efficiency of 62% was recorded for BL21 (DE3), while 100% transformation efficiency was recorded in C41 (DE3) and C43 (DE3), respectively. Interestingly, the percentage of toxic plasmid showed BL21 (DE3) transformed reaction, to have demonstrated 96% toxicity, which was evidenced by the absence of single colonies on the LB ampicillin plates contrary to C41 (DE3) and C43 (DE3). These transformed cell lines had 50% and 4% respectively, which demonstrates plasmid instability in BL21 (DE3) during the expression of toxic proteins in *E. coli* (Laurence et al., 2004). The findings of this work are consistent with other research that have compared a series of *E. coli* strains (Miroux and Walker, 1996; Dumon-Seignovert *et al.*, 2004). Hence the rationale for choosing C43 (DE3) competent *E. coli* cell as a viable host cell for sub-cloning sequence-verified AGTR1-plasmids (AGTR1-pGEX-6p-1, AGTR1-pET32a, AGTR1-pMAL-c5x, and AGTR1-p5x) DNA.

3.7.5.1 AGTR1-HIS-Tagged Fusion Protein Expression in C43 (DE3) Strain

A comparison of an SDS-PAGE resolved recombinant AGTR1-HIS fusion tag overexpression for construct 1 and 2 using *E. coli* strain BL21 (DE3) at 17°C (figure 1.3), at 37°C (figure 1.4) and *E. coli* strain C43 (DE3) strain, transformed overexpression at 28°C see (Figure 1.5), showed a degree of expression in the C43 (DE3) expressed fusion protein when 2YT media was applied. Although the migration on the AGTR1-HIS tag BL21 (DE3) transformed overexpression at 17°C and 37°C showed no valid successful recombinant protein overexpression (figure 1.3 and figure 1.4), the C43 (DE3) transformed AGTR1-HIS-tagged construct 1 and

2 fusion protein, overexpression at 28°C, in the presence and absence of glycerol preservation, showed a rather faint expression, especially with the non-glycerol preserved AGTR1-HIS tag contrast 1 and 2, which appears to be consistent with the expected molecular size of a recombinant AGTR1-HIS tag fusion protein of about 50kDa.

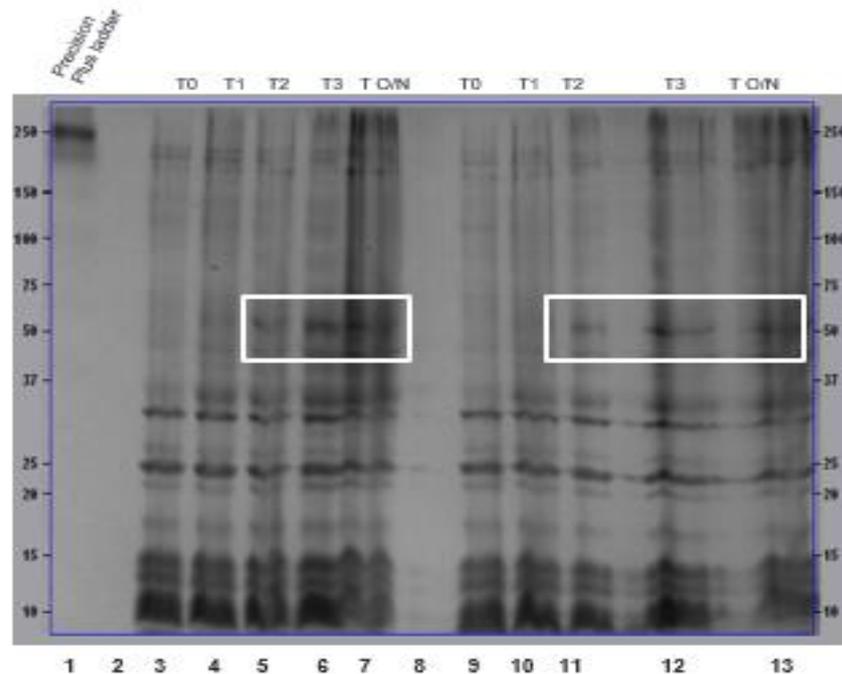


Figure 3.23: Protein Overexpression Analysis of AGTR1-HIS Tag fusion protein construct-1 and construct-2 expression in C43 (DE3) *E. coli* strain without glycerol preservation at 28°C in 2YT media.

Lane 1: Precision plus marker, Lane 3 – 7, showing non-glycerol preserved AGTR1-HIS Tag fusion protein construct-1 expression profile post IPTG (1mM) induction. T0 (L3), T1 (L4), T2 (L5), T3 (L6), To/n (L7). Lane 9 – 13, showing non-glycerol preserved AGTR1-HIS Tag fusion protein construct-2 expression profile post IPTG (1mM) induction. T0 (L9), T1 (L10), T2 (L11), T3 (L12), To/n (L13). It is resolved by Coomassie blue-stained SDS-PAGE methods at 150 Volts.

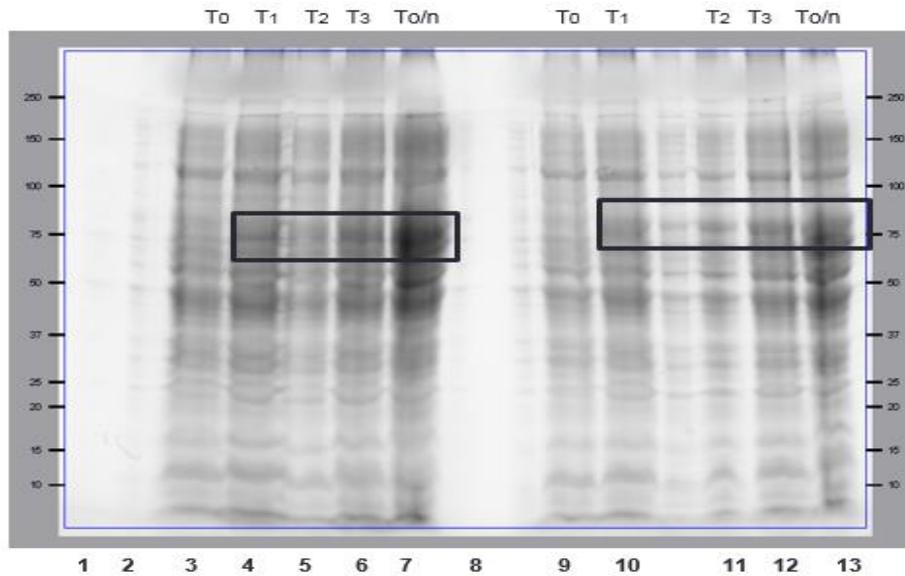


Figure 3.24: Protein Overexpression Analysis of AGTR1-HIS Tag fusion protein construct-1 and construct-2 expression in C43(DE3) *E. coli* strain with glycerol preservation at 28°C in 2YT media.

Lane 1: Precision plus marker, Lane 3 – 7, showing Glycerol preserved AGTR1-HIS Tag fusion protein construct-1 expression profile post IPTG (1mM) induction. T0 (L3), T1 (L4), T2 (L5), T3 (L6), To/n (L7). Lane 8: Precision plus marker, Lane 9 – 13, showing Glycerol preserved AGTR1-HIS Tag fusion protein construct-2 expression profile post IPTG (1mM) induction. T0 (L9), T1 (L10), T2 (L11), T3 (L12), To/n (L13). Resolved by Coomassie blue-stained SDS-PAGE methods at 150 Volts.

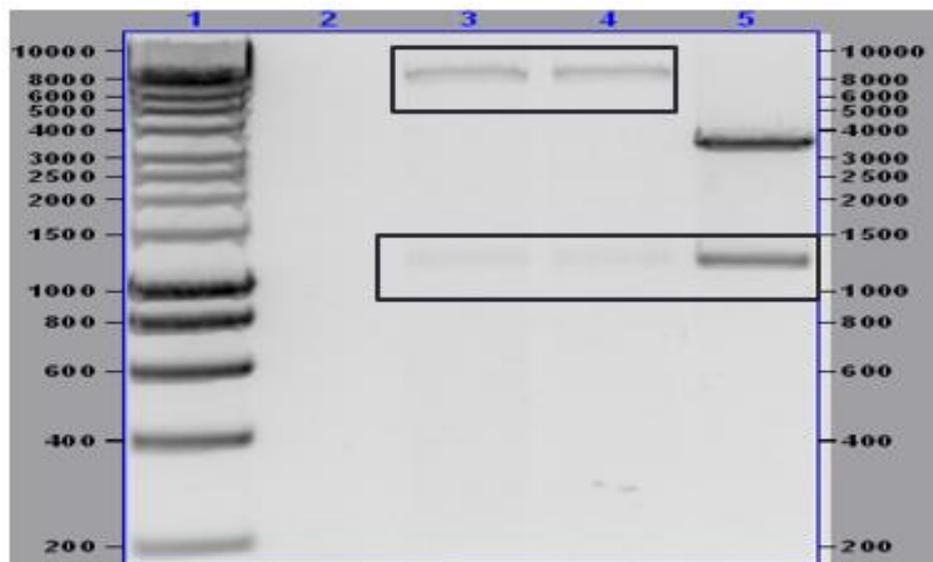


Figure 3.25: Plasmid Stability Analysis of Wild-Type AGTR1-HIS Tag fusion protein overexpression in C43 (DE3) *E. coli* strain at 37°C in 2YT Media.

Lane 1: showing hyper ladder 1kb and lane 3-5: Showing different AGTR1-HIS fusion protein post 1mM IPTG induction lysates using 2YT media. Lane 3: 3-hours post-induction AGTR1-HIS overexpression lysate. Lane 4: 6-hours post-induction AGTR1-HIS overexpression lysate. Lane 5: Pre-induction AGTR1-HIS lysate at time zero. Digested with BamH1 and EcoR1 restriction endonuclease enzyme and resolved by 1% Agarose Gel Electrophoresis.

The evaluation of AGTR1 plasmid stability analysis, during a wild-type AGTR1-HIS fusion tag protein expression, both before and after IPTG induction (figure 3.25), showed a clear degradation of AGTR1 gene, in a C43 (DE3) *E. coli* transformed expression. The obvious loss of AGTR1 plasmid integrity was observed after, a plasmid preparation, and restriction digest analysis of samples, isolated at 3 hours and 6 hours post IPTG induction (figure 3.25, lane 3 and 4), which were compared with the AGTR1-HIS tag fusion overexpression samples, isolated pre-IPTG-induction (figure 2.25, lane 5). This apparent AGTR1 gene degradation during protein expression maybe because of the AGTR1 gene toxicity effect and also connected to the difficulty experienced during AGTR1 gene fusion protein overexpression.

3.8.0 Other Troubleshooting Interventions and chapter conclusion

Cellular osmolytes concentration was among the troubleshooting interventions employed to enhance overall recombinant protein expression and solubility. Folding integrity of native proteins has been maintained during bacterial cell protein synthesis in the presence of betaine, high salt, and sorbitol (Oganseyan *et al.*, 2007). Bacterial chaperon co-expression and increase of intracellular osmolyte concentration has shown to aid formation of soluble native proteins, thereby preventing protein aggregation and formation of properly formed disulphide bonds and increasing the expression of difficult to express proteins (Goenca *et al.*, 2001, Zhang *et al.*, 2002, Chen *et al.*, 2003 and Liu *et al.*, 2005). These findings formed the rationale for supplementing 1mM of betaine and 0.5M of sorbitol into AGTR1-HIS-Tag and AGTR1-MBP-Tag protein overexpression media.

Chapter 4

CCL5 – Heparin Binding Interaction

4.0. Chapter Four

4.1 Introduction

Finding a novel approach of tightly regulating the chemokine system, has been a major focus of many researches to find cure to diseases, such as infectious diseases like HIV (Suresh and Wanchu, 2006, Paolo, 2015), inflammatory diseases like rheumatoid arthritis (Iwamoto *et al.*, 2008, Zoltan *et al.*, 2010 and Rory and Christian, 2010), auto-immune diseases like multiple sclerosis (Adam and Jacek, 2006; Mazzi, 2015 and Tahereh *et al.*, 2018) kidney dysfunction, reproductive endocrinology and cancer diseases (Stephen *et al.*, 2007, Raymond *et al.*, 2016). Inflammatory disease is used in this chapter as a case study. Here also we looked at the binding affinity interactions and implications of WT-CCL5 chemokines against glycosaminoglycans (heparan sulphate). A thorough understanding of this binding interaction may reveal how to disrupt the CCL5 – GAG interaction, hence attenuating the mobilization of white blood cells towards sites of inflammatory insults *in vivo*. We believe resolving this hypothesis may portends a viable anti-inflammatory strategy.

4.1.1 Chemokine and Chemokine Receptors

Chemokines are small peptides and a subfamily of the cytokine family; the name is derived from the contraction of chemotactic cytokines, because of its principal chemo attractant characteristics (Fernandez and Lolis, 2002). Chemokines are small soluble proteins that act as molecular signals to influence cellular mobility during inflammation (Dandan *et al.*, 2013). This means that chemokines can provide directional signals for the migration of cells both *in vivo* and *in vitro* (Kunkel and Butcher, 2003), and aid in cell positioning throughout development and

homeostasis. Chemokines play important functions in leukocyte recruitment and activation (Kunkel and Butcher, 2003), other functions of chemokines include their ability to induce cellular activation (Sokol and Luster, 2015 and Alex Klaren Beek *et al.*, 2012), to feature as a co-receptor for HIV infectivity and to play important role in cellular development (Sokol and Luster, 2015). Most chemokine proteins are about 8 – 10KDa in size, of which about 50 different chemokines have been discovered to date. Most chemokines and their corresponding gradients are detected by their ability to bind specific chemokine receptors (Sokol and Luster, 2015). However, there has been considerable redundancy between chemokines, largely due to a limited number of chemokine receptors, when compared with available chemokine ligands, ultimately leading to promiscuous chemokine ligand-receptor binding. A simple taxonomical system was invented for chemokines and its corresponding receptors, which had previously been classified based on their original identified functions (Bacon *et al.*, 2002). There are two types of chemokine, the inflammatory chemokines, and the homeostatic chemokines. The inflammatory chemokines are critical for chemo-attractant features of leukocyte mobilization and are expressed in inflamed tissues, whereas homeostatic chemokines sustain immune surveillance by leukocytes and maintain physiological traffic (Moser and Loetscher, 2001).

4.1.2 Chemokine Structures

Generally, the chemokine family, consist of about 50 low molecular weight polypeptides, which are composed of a tertiary conserved structure of four (4) cysteine motifs (White *et al.*, 2013), which are distinctly divided, into four

subfamilies of CC, CXC, CX3C and C sub-family (Murphy, 2002 and Allen *et al.*, 2007) as illustrated in figure 3. The classifications of the CC subfamily are essentially defined by the positioning of the two inherent cysteine residues at the N-terminus. These cysteine residues participate in disulphide bonding and are adjacent, oriented, or juxtaposed (CC) or separated by 1 or 3 non-conserved amino acid (CXC and CX3C, respectively). The CXC subfamily, has its amino acid between them, while the CX3C family of one member, has three (3) amino acids between the CC motif, while the C-subfamily have one amino acid within its structural orientation (White *et al.*, 2013).

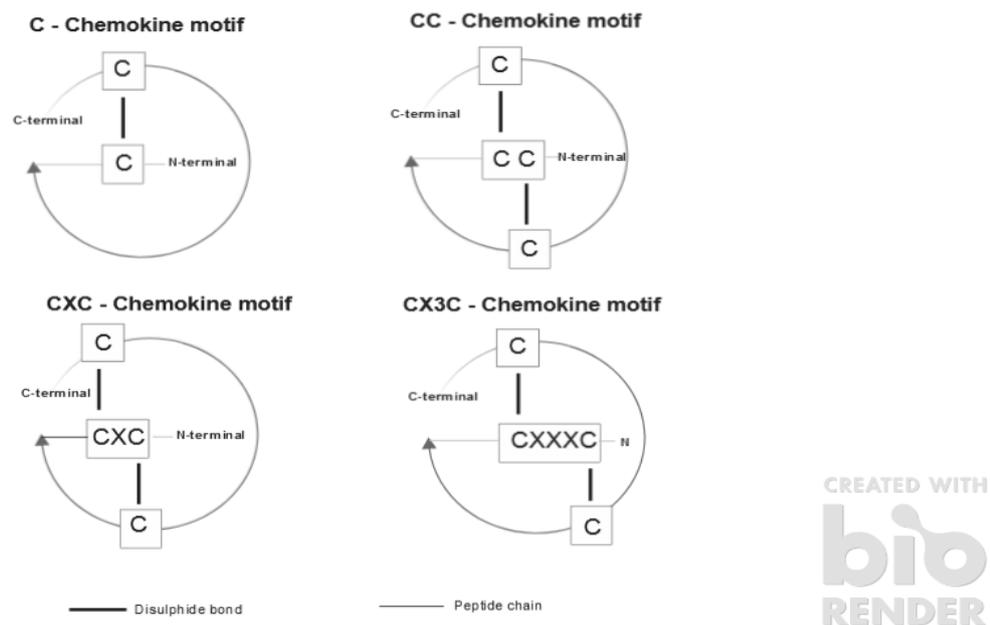


Figure 4.1: Structure of the four chemokine families

Schematics showing the disulphide connecting bonds, the peptide chains, and C and N terminals of the four chemokine families – C, CC, CXC, and CX3C. Further indicating the position of their cysteine placement. The image was created using bio-render software.

All the CC subfamilies are thought to act through the transmembrane spanning G-protein-coupled receptors (section 4.1.1), a feature that has made the studying of chemokines an interesting target for the pharmaceutical industry. The chemokines are interestingly the only kind of cytokines that act upon the 7-transmembrane receptors (Arimont *et al.*, 2017). Chemokines are made up of a common tertiary fold, a Greek key motif (Richardson, 1977), containing three antiparallel β -sheets, which is stabilized by multiple hydrogen bonds and hydrophobic interaction (Handel and Domaille, 1996). At the C-terminal α -helix of the chemokine structure, the structure is stabilised aided by disulphide bonds between the conserved cysteine residues (Schwiebert, 2005). The N-terminus of the chemokine structure plays a crucial role in receptor binding and activation. It's largely disorderly and the well-oriented regions create a scaffold to ensure appropriate conformation of protein binding interaction, even though they are often not directly involved in binding (Clark-Lewis *et al.*, 1995).

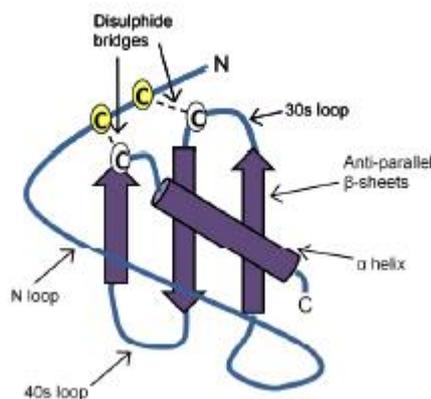


Figure 4.2 Chemokine Structure

Illustration of general chemokine structure. The Greek key-like motif containing three antiparallel β -sheets and an α -helix.

4.1.3 Cellular Recruitment for Chemokine Activities

Most biologic effects of chemokines are facilitated by their interaction with their receptors on the surface of the cell. Leukocyte mobilization into the tissues during inflammation, for example, is regulated by chemo-attractive cytokines and chemokines (Hans and David, 2012), which is potent upon interaction with GAGs (Farrugia *et al.*, 2017). Upon intracellular damage of cells, pro-inflammatory cytokines like TNF α and IL-1B are secreted, thereby activating the endothelium (Proudfoot *et al.*, 2017). The activated endothelium induces the production of chemokines, which are immobilized, to the endothelial cell surface-expressed GAG to form a chemotactic gradient with a directional signal (Proudfoot and Ugucioni, 2016). Some of the main events at the point of leukocyte recruitment and during the inflammation process are tightly regulated by GAGs, mostly embedding the surface of endothelial cells and leukocytes (Moore *et al.*, 1995; Parish, 2006; Webb *et al.*, 1993; Gotte, 2012; Oschatz *et al.*, 2011).

To arrive at sites of injury, leukocytes must exit the blood by traversing the endothelium, by attaching to the surface of the endothelium of post capillary venules (HEV). At this point, the cell starts to roll along the endothelium cells surface; this phenomenon creates a transient interaction of selectin-mediated leukocyte interactions (Lawrence and Springer, 1993). Selectins are lectin-like linkage glycoproteins that mediate leukocyte rolling, and they also facilitate the reduction of the velocity of leukocyte movement along endothelial cell, which results to firm adhesion on the surface of the endothelial matrix (Grange and Kube's, 1994, Bauer *et al.*, 2000, Hickey *et al.*, 1999, Gavins *et al.*, 2007). The

chemokine in two major steps, further activates the integrins via 1) integrin-mediated adhesion (tethering) and rolling of the flowing leukocytes on the endothelial cells matrix and 2) via stimulating the movement of adherent leukocyte across the endothelium and through the extracellular matrix as shown in Figure 4.6 (Springer, 1994; Butcher and Picker, 1996). The slow pace of rolling leukocytes on selectins presents an advantaged encounter with the chemokines which are overlaid on the apical surface of the endothelium by glycosaminoglycans (Tanaka *et al.*, 1993; Morla, 2019). These interactions provide strong leukocyte attachment to the endothelium, which expedites leukocyte haptotactic trans-endothelial migration (Springer, 1994; Rot, 1993 and Schimmel *et al.*, 2017).

Heparin appears to be one of the principal GAGS involved in the binding and release of bio-immune mediators and chemokines, that modulate almost all stages of leukocyte recruitment, adhesion, rolling and transmigration of leukocytes from circulation (figure 4.6) (Parish, 2006, Archana *et al.*, 2015).

The binding interaction between chemokines and its receptors expressed on the cell surface of leukocyte leads to a series of cellular events like the changing of $\beta 2$ integrin viability, mostly the CD11b/CD18, on the leukocyte cell surface, and other activities which ultimately, aims at eliminating the inflammatory stimuli. Other events, such as enhanced locomotion due to changes in cell shape, production of superoxide anions, and secretion of lysosomal enzymes, are also vital to the recruitment of leukocytes. Therefore, leukocytes mobilization toward the point of inflammation activates a cytokine-rich milieu which is sustained until the invading stimuli are eliminated. The behaviour of monomeric and oligomeric chemokines

during leukocyte recruitment, including systemic details of leukocyte mobilization for the inflammatory response, during trans endothelial migration, were all recently reviewed by Schimmel *et al.*, (2017).

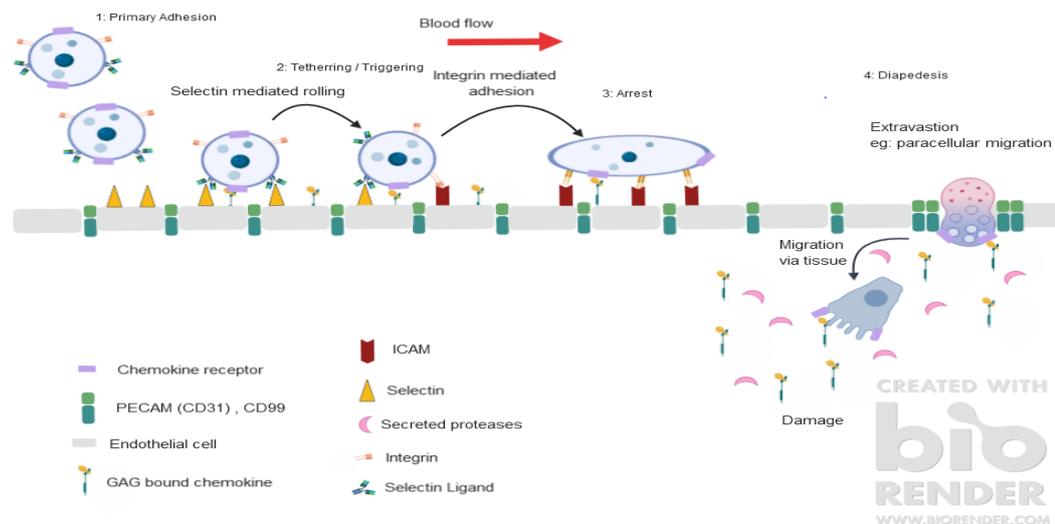


Figure 4. 3: Cellular recruitment of leukocytes/chemokine trans-endothelial migration.

Recruitment of leukocyte from circulation to stimuli position or inflammatory site through a series discrete event: tethering, rolling, adhesion and diapedesis, including chemotaxis through the endothelial membrane to the affected area.

4.1.4 CCL5/RANTES

Human CCL5 protein is an 8kDa protein, also known as (C-C motif) chemokine ligand 5, which is encoded by the CCL5 gene and categorized as a cytokine or chemokine protein. The CCL5 gene, also referred to as RANTES (regulated on activation, normal T cell expressed and secreted) are part of the chemokine genes localized to the q-arm12 of the human chromosome 17 (Donlon *et al.*, 1990). The CCL5 was first discovered in 1998 (Schall *et al.*, 1988), and first isolated and characterized in 1990 (Schall *et al.*, 1990). It belongs to a superfamily of proteins

that are vital in inflammatory and immune-regulatory activities and are expressed in over 100 human diseases. CCL5 proteins exhibit a chemo-attractant characteristic by responding to chemical signals for blood monocytes, basophils, T cells, and eosinophil. It also aids in the active mobilization of leukocytes into inflammatory sites, enabling angiogenesis, stimulating growth or survival, and triggering tumor cell metastasis (Coussen and Werb, 2002; Taichman *et al.*, 2002, Manes *et al.*, 2003).

4.1.5 CCL5 and Inflammatory Diseases

Chemokines have been reported for their important role in inflammatory diseases, including promoting stroma genesis and carcinogenesis (Cambien *et al.*, 2011). The CCL5 are classified among the C-C chemokine subfamily, with its chemokines present on adjacent cysteine position, and consist of most of the chemokines (Murphy, 2002; Allen, 2007). The CCL5 does induce the in vitro migration and mobilization of eosinophils, T cells, dendritic cells, basophils, mast cells, and NK cells. The original name of CCL5, was RANTES as it thought to have been a T cell-specific cytokine (Schall *et al.*, 1988). The blood platelets, macrophages, endothelium, fibroblasts, epithelial, endometrial cells, and eosinophils all produce the CCL5. The most relevant three chemokine C-C motif receptors (CCRs) linked with CCL5 are CCR1, CCR3, and CCR5, similarly known as inflammatory chemokine receptors (Murphy, 2002; Sarvaiya *et al.*, 2013). CCL5 had also been revealed to bind the G protein-coupled receptor 75 (Ignatov *et al.*, 2006), thereby potent in inducing G protein signalling, independently. Upon the activation of the G protein-dependent signalling, the CCL5-receptor complex is co-opted through the clathrin-

mediated endocytosis associated to the Adaptin-2 and b-Arestin adapter molecules. Additionally, the recycling of receptors and the degradation of viable chemokine receptor complexes are a function of mobilizing cytoplasmic organelles and vacuoles. These mobilizations are important to regulate chemokine receptor levels at the extracellular membrane, thereby presenting a channel for the regulation and reduction of the cellular activities of CCL5 and other chemokines extracellular level, a feature important to maintaining homeostasis (Cardona *et al.*, 2008; Mantovani and Locati, 2008). The CCL5 protein activities are implicated in multiple biological processes, such as inflammatory disorders, HIV, and cancer (Levy, 2009), due to the variety of cells that express and mediate the protein.

The inflammatory CCL5 chemokine variably binds to these receptors. For instance, the CCL5 demonstrates a greater binding affinity towards CCR5 and CCR1 receptors, whereas its binding affinity towards the CCR3 receptor is lesser (Blanpain *et al.*, 2001). The interaction between the CCL5 chemokine and its receptor and other chemokine receptors are illustrated in figure 1.

On a broader perspective Elias *et al.*, (2013), had extensively investigated and reported the role of CCL5 in various diseases such as viral infections, respiratory tract infection, neurotropic infection, liver infections, helminth infection, asthma, atherosclerosis, angiogenesis and cancer, fibrosis, and transplant rejection. The figure below previews the list of current CCL5 new investigative tools presently available under different studies.

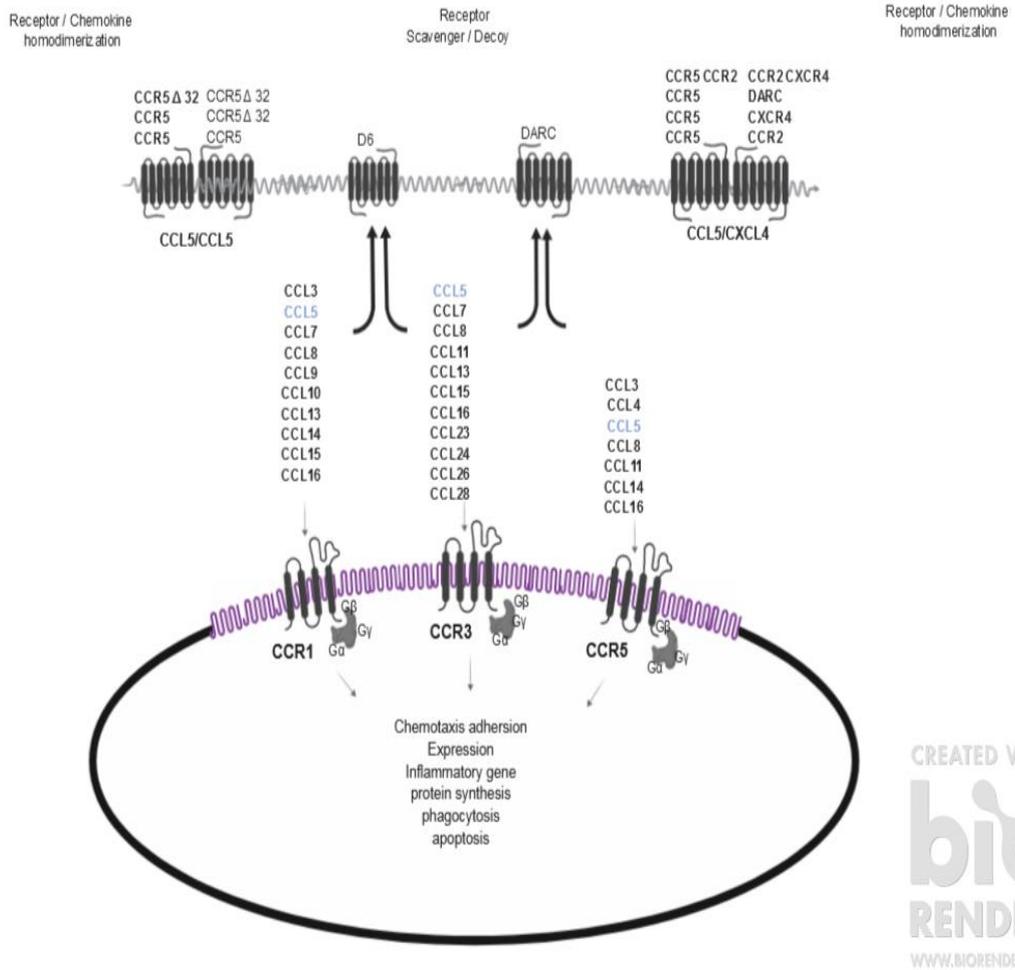


Figure 4.4: The Biology of CCL5.

Schematics of known chemokine receptors for CCL5 and their commonly shared agonist. Also illustrating the effects of chemokine activation on the leukocyte and non-leukocyte populations. The CCL5 biological effects are often modulated by the activities of a non-functional CCR5Δ32 or Atypical (decoy) receptors (D6 and DARC)

4.1.6 Structure of CCL5

The structure of CCL5/RANTES (PDB codes 1RTN and 1HRJ) was first determined as a dimer by using the NMR spectroscopy method at pH 3.7 in 1995 (Repeke *et al.*, 2018, Jin-Ye *et al.*, 2020). The CCL5 was first found in the T cell-specific cDNA, gaining attention for the chemokine as a new subfamily of chemotactic proteins with potential, in influencing a major biological and pathological processes (Levy, 2009). The identification of CCL5 has assisted in collaborating on the functional role of its residues.

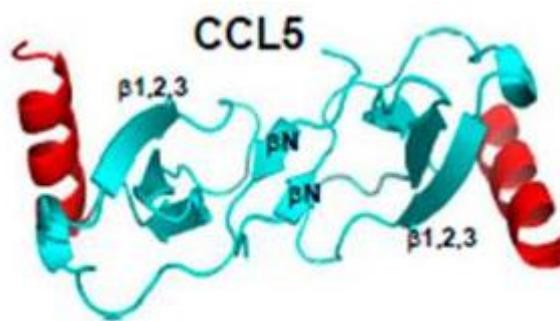


Figure 4.5 A Dimer structure of CCL5/RANTES (Proudfoot and Borlat, 2000)

4.1.7 Glycosaminoglycan

Glycosaminoglycans also known as GAGs, are long complex linear polysaccharides of high molecular mass up to few million Dalton, existing in a varied range of sizes and often highly heterogeneous sulphated glycans (Murata *et al.*, 1985, Jackson *et al.*, 1991, Yeung *et al.*, 2002). They are present on the extracellular matrix of cell surface and consist of over 1000 repetitive disaccharide moieties, each made up of (i) an acetamido amino sugar (*N-acetylglucosamine* or *N-acetylgalactosamine*) in conjunction with (ii) uronic sugar i.e.

(glucuronic acid or galactose) or iduronic acid (Jeffrey *et al.*, 2009). GAGs vary in two essential forms, such as (i) the type of hexosamine and hexose acid attachment they contain, (ii) the geometry of its glycosidic bonding (covalent linkage that joins the sugar molecule to another group).

A few interesting and favourable features are peculiar to the Glycosaminoglycan in the human body. They also have distinguished GAG as a molecule of interest for futuristic therapeutic investigation and interventions. These features include but are not limited to, (i) GAGs are vastly polar and attract water, hence essential to the body as a lubricant or as a shock absorber, (ii) GAGs characterized as a part of the extracellular matrix (ECM) (Hay 1991), and play essential role in cell signalling processes and cell recognition (Mende *et al.*, 2016), (iii) the ability of GAG structure to act as a polymer with multiple anionic charges at several sites (Grand *et al.*, 2014), and (iv) GAGs characteristic potential to interact with cationic charged moieties like cytokines, growth factors, amino acids or plasma proteins (Bernfield *et al.*, 1999, Osterloh *et al.*, 2002).

The glycosaminoglycans are classified into four functional groups among which are (i) heparin (HP) /heparin sulphate (HS), designated as HSGAGs (ii) chondroitin sulphate (CS) / dermatan sulphate (DS) designated as CSGAGs (iii) keratin sulphate (KS), these constituents occur as proteoglycans (Oohira *et al.*, 1983) and (iv) hyaluronic acid (HA), occurring as oligosaccharide chain itself (Willen *et al.*, 1991).

These classifications are essentially according to their respective core disaccharide structures (Ram *et al.*, 2006). In contrast to glycoproteins of

predominate carbohydrates side chain, most GAG functional molecules (except hyaluronic acid) are known to be covalently bonded to serine residues on their proteins of interest via tetra saccharide linkages to form proteoglycans (Williams, 1998). These proteoglycans are essentially responsible for most of the GAGs physiological functions (Victor *et al.*, 2009), and are structurally categorized based on their attached GAG sub-type group into heparan sulfate PG (HSPG), chondroitin sulfate PG (CSPG), dermatan sulphate proteoglycans (DSPG), keratan sulfate PG (KSPG) and hyalectans which are aggregates of HA and extracellular PG (Day and Prestwich 2002). For this chapter, only heparin will be investigated and discussed as a major GAG in interaction with CCL5.

4.1.8 Heparin and Heparan sulphate

Heparan sulphates and its fully sulphated subtype, heparin, are long linear sulphated, heterogeneous polysaccharides, and have the most complex GAG structure (Olczyk *et al.*, 2014). More specifically, heparan sulphate is known to be resident on the surface of the cell membrane and in the extracellular matrix (ECM) of a protein core to form proteoglycans as part of heparan sulphate proteoglycans (HSPGs) of GAG (Maria *et al.*, 2015).

These Heparin and HSPGs are negatively charged carbohydrates involved in interactions with many HSBP; interestingly, sulphation appears to be a key part of driving the biological functions of heparin and heparan sulphate in roles as diverse as angiogenesis, cell growth, adhesion, and blood coagulation

(Weiss *et al.*, 2017). Heparan sulphate and Heparin are made of alternating saccharide units of N-acetylated or N-sulfated D-glucosamine that are α (1-4) - or β (1-4) - linked to uronic acids (Iduronic or D-glucuronic acid). More precisely, the main disaccharide of Heparan sulphate is composed of glucuronic acid (GlcA) and N-acetylated glucosamine (GlcNAc), whereas the main disaccharide of Heparin consists of iduronic acid, which is sulfated at the carbon 2 (IdoA2S) and N-sulfated glucosamine which is additionally sulfated at C6 (GlcNS6S) (Kowitsch *et al.*, 2018). Interestingly, HSPS is synthesized in almost all mammalian cells, in different forms (Rabenstein 2002; Varki 1999), unlike Heparin which is synthesized specifically only in the connective-tissue of mast cells or basophils as a part of the serglycin proteoglycan (Gallagher and Walker, 1985). The conformation of heparan sulphate differs unevenly and temporally in various cell types, and during development, more so, the factors involved in regulating the biosynthesis of heparan sulphate largely remains poorly defined (Esko and Lindahl, 2001)

Interestingly, the structural similarity, existing between Heparan Sulphate and Heparin, does suggest they might possess similar regulative effects (Meneghetti *et al.*, 2015). This is due to their common and unique ability to bind a variety of growth factor (GF) proteins, and thus, induce GF-mediated signal transduction (Eswarakumar *et al.*, 2005; Lindahl *et al.*, 1994). Depending on its specific sulfation motifs, HS can activate fibroblast growth factor proteins (FGF) - and fibroblast growth factor receptor (FGFR) -signaling pathways (Guimond and Turnbull 1999; Xu and Esko 2014). Furthermore, Heparan sulphates can inhibit or activate chemokines, cytokines like interleukins, and tumor necrosis factor α (TNF- α) (Whitelock and Iozzo 2005; Kowitsch *et al.*, 2018). From a clinical perspective,

Heparan sulphate has been reported to be involved in embryonic development, defence of the immune system, cell growth, and inflammation (Xu and Esko, 2014). More precisely, a few studies have collaborated on the participation of Heparan sulphate in wound healing and inflammatory repairs (Olczyk *et al.*, 2105). Including its ability to couple covalently with syndecan and glypican PGs, when present on the cell surface, hence interacting with integrin (Farrugia *et al.*, 2017).

4.1.9 CCL5 Interaction with Heparin and HSPGs during Inflammation

During inflammation, programmed leukocyte recruitment from circulation and transmigration across the endothelium into the tissue and to the site of inflammation are essential in resolving or reversing inflammatory stimuli (Schimmel *et al.*, 2017). The chemokine-mediated transmigration is facilitated by the interaction of CCL5 protein with Heparan sulphate proteoglycans (HSPGs). The HSPGs are abundantly distributed on the wall of the endothelia (Farrugia *et al.*, 2018) and aids the transportation of CCL5-mediated transmigration, across the endothelial wall, through a process called transcytosis (Christopher, 2006).

The Heparin is reported to have multifunctional roles during inflammation, mainly due to the negative charge of heparin multiple sequence polysaccharides subunits. This allows heparan sulphate to interact with a vast range of positive chemokine ligands (Laura and Linda, 2018). The chemokine CCL5 is a positively charged protein that binds the Heparin by electrostatic interaction (Proudfoot *et al.*, 2003). The binding is not a simple electrostatic interaction but appears to have specificity for sulphated Heparin and HSPGs (Carter *et al.*, 2003). This leads to an increase of

CCL5 concentration and bioactivity at the site of production (Whitelock and Iozzo, 2005). Hence the formation necessary haptotactic CCL5 gradient needed for *in vivo* chemokine function, such as leukocyte recruitment during an inflammatory response. The CCL5 protein, like most chemokines, has several highly conserved structural regions, including basic domains for GAG binding interactions (Seegerer *et al.*, 2009). The WT-CCL5 protein consists of two major clusters of basic residues, a BBXB motif on the “40s loop” of its amino acid sequence precisely at the (⁴⁴RKNR⁴⁷). The second binding cluster being the BBXB motif on the “50s loop” of the WT-CCL5 protein, precisely at the (⁵⁴KKWVR⁵⁹), as illustrated in (Figure 4.6) (Martin *et al.*, 2001). These two regions of WT-CCL5 matrix-binding activities have been widely studied previously (Proudfoot *et al.*, 2003; Liang *et al.*, 2016). The 40s cluster domains have been attributed to demonstrate more binding of GAG substrates (Martin *et al.*, 2001; Proudfoot *et al.*, 2001). The amino acid residues essential for GAG-binding activity have been characterized and shown to suggest that wildtype CCL5 protein exerts maximum specificity in the GAG interaction (Hoogewerf *et al.*, 1997, Spillmann *et al.*, 1998, Koopmann *et al.*, 1999, Martin *et al.*, 2001, Laguri *et al.*, 2007, Laguri *et al.*, 2008, Cox *et al.*, 2008). To demonstrate the extent of GAG binding interaction within the CCL5 motif, previous studies had used a neutral amino acid Alanine (A) to substitute the amino acid unit designated as “B” within the 44-47 and 54 – 59 binding regions of the WT-CCL5 protein sequence. Previous studies have collaborated the claim, that a CCL5 mutation at the 40s loop ⁴⁴AANA⁴⁷, was unable to recruit cells in a model of peritoneal cell recruitment (Proudfoot *et al.*, 2003, Proudfoot *et al.*, 2001, Seegerer *et al.*, 2009).

4.1.10 CCL5 Oligomerization in GAG-Binding Interaction

A broader examination of the effects of chemokine oligomerization during GAG binding interaction has been previously investigated and reported ((Johnson *et al.*, 2004; Migliorini *et al.*, 2014; Salanga *et al.*, 2014; Dyer *et al.*, 2016). Chemokines exhibit a range of different measure of affinity for GAGs. While some chemokine oligomerize and demonstrate improved binding functionality, others cannot, and this deficiency has no impact their biological functions and binding interactions (Dyer *et al.*, 2015).

Although the oligomerization of CCL5 and its mutagenic variants during Heparin-GAG interaction, is not the focus on this research, we note however that Johnson and co reported CCL5 formed high molecular oligomeric mass (>600 kDa) in vitro and in vivo during GAG-binding interaction (Johnson *et al.*, 2004). Similarly, a study by Proudfoot and co, investigated the monomeric CCL5 binding interaction with GAG and reported its quaternary structure was non-functional during cell activation and chemotaxis in vitro. On the other hand, the monomeric and dimeric CCL5 were functionally active in vivo, with only CCL5 tetramers was able to attain full cell recruitment activity (Proudfoot *et al.*, 2003).

4.1.0 Research Aim

This chapter aim is to systematically characterize the apparent affinities and binding kinetics of GAG-Heparin against wild type CCL5/ mutant's chemokines using a functional Surface Plasmon Resonance assay as a way of understanding their contribution to leukocyte mobilization during inflammation. The SPR method is used to evaluate under real-time conditions, the binding affinity/kinetics between the wild type CCL5 binding domains, against the GAG carbohydrate compound, as well as the binding interaction of the same GAG against the Alanine substituted binding domain of mutant CCL5 protein, at position ⁴⁴RKNR⁴⁷, ⁵⁴KKWVR⁵⁹, and both respectively.

4.2.0 Results

4.2.1 Introduction

In order to make synthesized CCL5 wildtype and mutant proteins of interest, genomic sequences of CCL5 gene were extracted from an online data-based known as Ensembl tool (<https://www.ensembl.org/>) and aligned using EXPASY ProtParam tool (<https://web.expasy.org/protparam/>), the Integrated DNA Technology CodonOpt tool (<https://eu.idtdna.com/site>) were applied to optimize plasmid codons since *E.coli* was going to utilize as the host system for protein synthesis. To achieve the required neutralization of the two GAG- binding regions of the CCL5 protein sequence, an amino acid of neutral charge Alanine was used to substitute the ⁴⁴RKNR⁴⁷ cluster to ⁴⁴AANA⁴⁷ and the ⁵⁵KKWVR⁵⁹ cluster to ⁵⁵AAWVA⁵⁹. These CCL5 constructs were designated CCL5MutLOW and CCL5MutHIGH; the third CCL5 mutant sequence had the Alanine substitution at both the RXNR and the KKWVR residues of the 40s and 50s cluster position respectively and designated as CCL5MutBOTH in this work. The Alanine substitution attenuates the conventional electrostatic interaction between negatively charged CCL5 binding domains and a positively charged GAG-heparin. The designed wild type and mutant CCL5 sequences were outsourced from IDT UK. To achieve a robust and successful protein synthesis, the pMAL-c5x affinity tag, which expresses MBP were used to construct the CCL5 plasmids and functioned as a fusion protein tag during overexpression. The fusion tag as described chapter 1, was to ensure the ease purification of recombinant protein. Figure 4.17 below, shows the GAG binding

region on CCL5 wild sequence, including point mutations of all the mutants under investigation.

NAME	CODON OPTIMISED SEQUENCE (5' – 3') and PROTEIN SEQUENCE
CCL5 Wild Type	<p>Bam H1 ggatccAGCCCCTATTCAAGTGACACCACTCCTTGCTGTTTCGCTTATATCGCTCGCCCTTACCACGCGCCACATTAAGGAGTATTTTACACTAGCGGTAATGTAGTA ACCCCGCGTGGTCTTCGTTACCGCTGCAAATGCCAGGTCTGCGCGAATCCAGAAGCAGCATGGGTAGCCGAGTATATCAATTCCTTAGAAATGTCCTAGgaattc</p> <p>CCL5 Wild Type Protein Sequence: SPYSSDTPCCFAYIARPLPRAHIKEYFYTSKGKCSNPAVVFVTRK⁴⁴NRQVCANPK⁵⁵KWVREEYINSLEMS N-terminal → 44BBXB47 55BBXB59 → C-terminal EcoR1</p>
CCL5MutBOTH Affinity	<p>Bam H1 ggatccAGCCCCTATTCAAGTGACACCACTCCTTGCTGTTTCGCTTATATCGCTCGCCCTTACCACGCGCCACATTAAGGAGTATTTTACACTAGCGGTAATGTAGTA ACCCCGCGTGGTCTTCGTTACCGCTGCAAATGCCAGGTCTGCGCGAATCCAGAAGCAGCATGGGTAGCCGAGTATATCAATTCCTTAGAAATGTCCTAGgaattc</p> <p>CCL5MutBOTH Affinity Protein Sequence: SPYSSDTPCCFAYIARPLPRAHIKEYFYTSKGKCSNPAVVFVTAANAQVCANPEAAWVAEYINSLEMS EcoR1</p>
CCL5MutHIGH Affinity	<p>Bam H1 ggatccTCACCGTATAGCAGTGATACTACACCTTGCTGTTTTCGCTACATCGCACGCCCTCTGCCCGCGCCCATATTAAGGAGTATTTTACACTTCGGGGAAGTGTAGT AACCCGCTGTCGTATTTCGTTACAGCCGCTAACGCACAAGTGTGTGCAAACCCGGAAAAAATGGGTGCGTGAGTACATCAACTCCCTGGAGATGTCATAGgaattc</p> <p>CCL5MutHIGH Affinity Protein Sequence: SPYSSDTPCCFAYIARPLPRAHIKEYFYTSKGKCSNPAVVFVTAANAQVCANPEKKWVREYINSLEMS EcoR1</p>
CCL5MutLOW Affinity	<p>Bam H1 ggatccAGTCCCTATTTCGCTGACACAACGCCCTGTTGCTTCGCTACATCGCTCGCCCTGCTCGCGCACATATTAAGGAATACTTTTACACTTCTGGGAAATGCAGC AACCCGGCAGTTGTGTTTGTGACTCGCAAAAACCGTCAGTTTTGTGCAAATCCTGAAGCCGCTTGGGTAGCCGAATACATCAATTCGCTGGAAATGTCATAGgaattc</p> <p>CCL5MutLOW Affinity Protein Sequence: SPYSSDTPCCFAYIARPLPRAHIKEYFYTSKGKCSNPAVVFVTRKNRQVCANPEAAWVAEYINSLEMS EcoR1</p>

Figure 4.6: Heparin Binding Amino Acid Region of Wild-type CCL5 and CCL5 Mutant Sequence
 Showing regions of Alanine substitution (in red) of all CCL5 protein constructs, where GAG's (heparin) exerts its electrostatic affinity on RANTES

4.2.2 CCL5 DNA Extraction

The pUC57 carrier vector sequences containing full-length of both wild-type CCL5 and other three mutants (CCL5MutLOW, CCL5MutHIGH and CCL5MutBOTH) under investigation were purchased from IDT UK, the pUC57 here, functioned as a cloning vector which cannot be used for overexpression, was pre-encoded with *Bam*HI and *Eco*RI enzyme to allow for easy digestion and extraction of interest gene (CCL5). Restriction digestion reaction as explained in the method section 2.2.1, was used to obtain all CCL5 genes. The WT-CCL5 DNA band migrated in a consistent pattern with expectations for WT-CCL5 (210 bp) as shown in Figure 4.7(A). The entire CCL5 mutant DNA bands were also observed to have migrated in a consistent pattern with expectations for CCL5 genes. Where lane 3 -5, 7-9 and 11-13 of figure 4.7(B), showed expression band patterns that are consistent with expected sizes of full-length CCL5 mutant genes, of 210bp plasmid length.

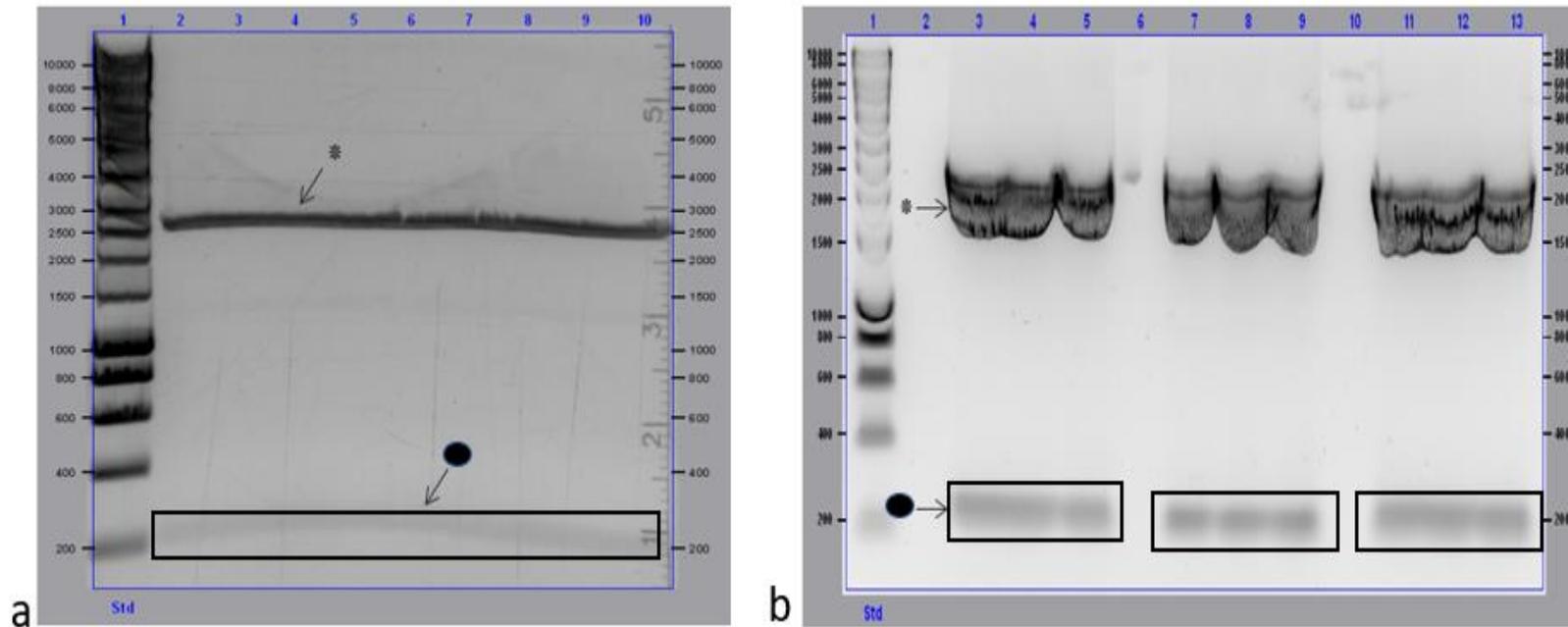


Figure 4.7: DNA Extraction Analysis of CCL5 Genes from IDT pUC57 Carrier Vector

(A) Lane 1: showing hyper ladder 1kb, lane 2-10, showing digested Wild-Type CCL5-pUC57 construct. (B) Lane 1: showing hyper ladder 1kb, lane 3-5, showing digested CCL5MutLOW- pUC57 construct, lane 7-9, digested CCL5MutHIGH- pUC57 construct and lane 11-13, and digested CCL5MutBOTH- pUC57 construct. All plasmid construct was digested with BamH1 and EcoR1 restriction endonuclease enzyme and resolved by 1% Agarose Gel Electrophoresis. Where * designates pUC57 IDT carrier vector and O designates respective CCL5 fragments.

4.2.3 Wild Type CCL5 Cloning in Different *E. coli* Vectors

Thereafter, the extracted and digested WT-CCL5 gene were ligated to several *Bam*H1 and *Eco*R1 pre-digested vectors (pGEX-6p-1, pMAL-p5x, and pMAL-c5x) using sticky end ligation protocol as described in the method section (2.2.5) to obtain recombinant WT-CCL5 DNA plasmids. Recombinant WT-CCL5-pGEX-6p-1 plasmids were used to transform competent *E. coli* DH5 α strain, and transformants were selected in solid culture with ampicillin. The recombinant WT-CCL5-pGEX-6p-1 plasmids figure 4.8(A), CCL5-pMAL-c5x plasmids figure 4.8(B) and CCL5-pMAL-p5x plasmids figure 4.8(C), were all extracted from selected transformants after small scale 2YT broth culture and analysed through restriction digest profiling with same enzymes used for plasmid DNA cloning (4.2.4) and resolved by agarose gel electrophoresis.

The amplicon patterns generated in the agarose gel screening of figure 4.8(A) (Lane 3, 4, 5, 7 and 8), figure 4.8(B) (Lane 3-8), and figure 4.8(C) (Lane 3-7) were all compatible with expectations. The hype ladder 1kb band that migrates between 5000 and 6000 bp markers showed to be consistent with pGEX-6p-1, which is a plasmid of 4984 bp in length, which reflects the migration of 5677 bp length for pMAL-c5x and 5725 kb length for pMAL-p5x respectively. The WT-CCL5 gene transformants except in figure 4.8(A) (lane 6) generated bands that migrated consistently with expectations as defined in section 4.2.1., which implied that the WT-CCL5 gene cloned into the chosen vectors.

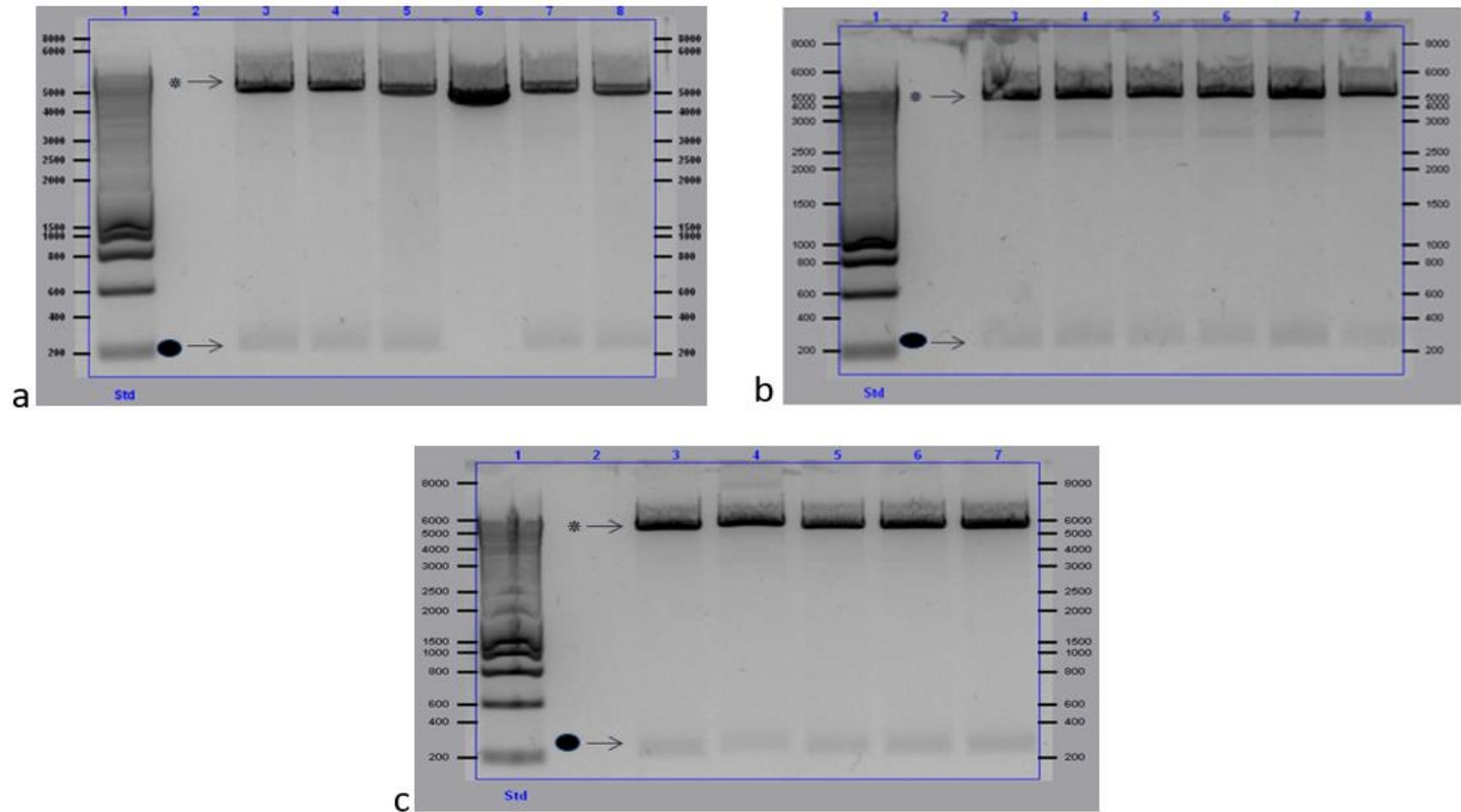


Figure 4.8: Restriction Digest Analysis of Recombinant Wild-Type-CCL5 Plasmid

(A) Showing hyper ladder 1kb and lane 3-8: various Wild-type-CCL5-pGEX-6p-1 constructs (B) Lane 1: showing hyper ladder 1kb and lane 3-8: various Wild-type-CCL5-pMAL-c5x constructs (C) Lane 1: showing hyper ladder 1kb and lane 3-8: various Wild-type-CCL5-pMAL-p5x constructs. All digested with BamH1 and EcoR1 restriction endonuclease enzyme and resolved by 1% Agarose Gel Electrophoresis. Where *, designates respective vector plasmids, and O Wild-type-CCL5 fragment.

Nucleotide sequencing was used to validate that WT-CCL5-pGEX-6p-1, WT-CCL5-pMAL-c5x, and WT-CCL5-pMAL-p5x recombinant plasmids were all successfully cloned as sequencing results came back without any mutations.

4.2.4 Cloning of CCL5 Mutants into pMAL-c5x *E. coli* Vector

Upon the successful extraction and nucleotide sequencing of all mutant CCL5 genes, the pre-digested extracted genes were ligated into the *Bam*HI and *Eco*RI restriction sites of pre-digested pMAL-c5x vectors. The sticky end ligation protocol as described in the method section (2.2.5) was applied to obtain digested recombinant DNA-plasmids. The recombinant mutant CCL5 plasmids were used to transform competent *E. coli* DH5 α strain, and transformants were selected in solid culture under ampicillin resistance. Single colonies of these recombinant mutant CCL5 genes were selected from the transformants. To confirm successful cloning a small scale 2YT broth cell culture was grown using selected single colonies and analysed through restriction digest analysis with the same enzymes used for plasmid DNA cloning (4.2.4) and resolved by agarose gel electrophoresis. The amplicon patterns generated in the agarose gel screening of figure 4.9 (A) lane 3, 4, 6, and 8, figure 4.9(B) lane 3, 5, 6, and 8 and figure 4.9(C) lane 2-6, were all compatible with expectations. The hyper ladder 1kb band that migrates between 5000 and 6000 bp markers showed to be consistent with the pMAL-c5x vector, which is a plasmid of 5677 bp length.

Most CCL5 mutant gene transformants except for those in figure 4.9 (A) lanes 5 and 7, figure 4.9 (B) lanes 4 and 7 and 4.9 (C) lane 7 generated bands that migrated consistent with expectations as defined in section 3.2.1., which indicated that mutant CCL5 genes were effectively cloned into the pMAL-c5x vector. DNA nucleotide sequencing was used to validate that CCL5MutLOW-pMAL-c5x and CCL5MutBOTH-pMAL-c5x recombinant plasmids were all successfully cloned as sequencing results came back without no mutations except those we had deliberately introduced.

4.2.5 Protein Overexpression and Solubilisation of Wild-Type CCL5-MBP and CCL5 Mutants Fusion Protein

Although WT-CCL5 had been successfully cloned into several vectors, as reported in previous publications (Proudfoot and Borlat, 2000). Those cloning results had similarly suggested that the solubility of CCL5 was superior if overexpressed as an MBP fusion protein. Therefore, further protein synthesis was done using MBP-tagged selected constructs. The CCL5/MBP plasmid DNA constructs were transformed and overexpressed in *E. coli* C43 DE3 strain. For protein overexpression 2YT media was supplemented with 0.5M Sorbitol and 1mM Betaine was applied (figure4.14). Prior to the induction of expression, cells were grown to mid-log phase at 37°C. Overexpression samples were taken for analysis at one-hour post 1mM IPTG induction at 37°C (T1), two hours post IPTG induction (T2) and overnight (18 hours) post IPTG induction (to/n), followed by analysis using the SDS-PAGE. After initial expression analysis, it was qualitatively assessed that two hours of expression gave the

best ratio of soluble to insoluble CCL5-MBP recombinant fusion protein. Examples of the expression profiles and the solubility of the CCL5-MBP fusion proteins for all constructs of interest can be seen in figure 4.11, 14.12 and 14.13. A review of the SDS-PAGE analysis of clarified WT-CCL5-MBP, CCL5-MutLOW-MBP, CCL5-MutHigh-MBP and CCL5-MutBOTH-MBP fusion protein overexpression profile, through the analysis of its lysates, showed that synthesis of these recombinant proteins of interest was induced under these conditions in a 2-hour period post IPTG induction. Furthermore, the precision plus molecular weight bands showed a migration between 37 and 50kDa marker, with an expected protein size of about 50.5kDa for all the recombinant CCL5-MBP fusion proteins of interest for this work.

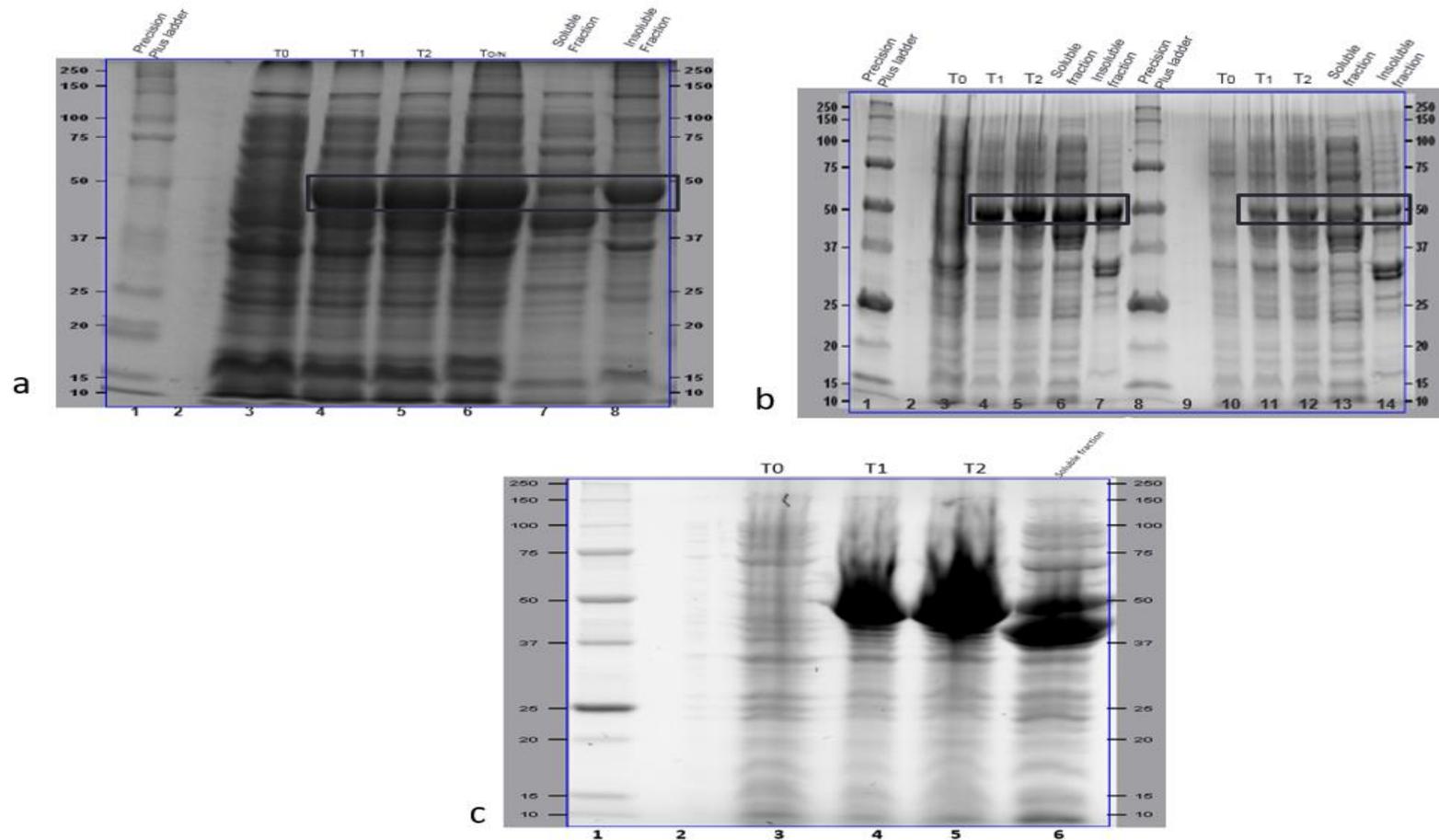


Figure 4.11: Protein Overexpression Analysis of Wild-Type CCL5-MBP and Mutant CCL5-MBP fusion overnight expression in C43 DE3 *E. coli* strain at 37°C in 0.5M Sorbitol and 1mM Betaine 2YT Media

(A) Lane1: precision plus marker, Lane 3 – 8 Wildtype CCL5-MBP expression profile showing post-IPTG (1mM) induction and overnight expression. T0 (L3), T1 (L4), T2 (L5), T0/N (L6), the soluble fraction (L7), the insoluble fraction (L8). **(B)** Lane1: precision plus marker, Lane 3 – 7 CCL5-MutLOW expression profiles showing post IPTG(1mM) induction. T0 (L3), T1 (L4), T2 (L5), soluble fraction (L6), insoluble fraction using (L7), and Lane 8: precision plus marker, Lane 10 – 14 CCL5-MutBOTH expression profile showing post IPTG(1mM) induction. T0 (L10), T1 (L11), T2 (L12), soluble fraction (L13), insoluble fraction (L14). **(C)** Lane1: precision plus marker, Lane 3 – 6 CCL5-Mut-HIGH- MBP expression profile showing post IPTG (1mM) induction and overnight expression. T0 (L3), T1 (L4), T2 (L5), the soluble fraction (L6). Resolved by SDS-PAGE method.

4.2.6 Solubilisation of WT-CCL5-MBP Protein

An appreciable amount of WT-CCL5-MBP fusion protein was observed to reside in the soluble fraction; this appeared qualitatively to be more of an issue with the wildtype construct. To maximize the recovery of soluble protein attempts were made to re-solubilize the insoluble fraction (lanes 7 of figure 4.12 and lanes 8 of figure 4.13). To recover these proteins, the application of solubilizing detergents was incorporated. Two different solubilizing detergents of 6M Urea and 8M Urea, 0.5% Tween 20, 10mM DTT in 20mM Tris pH 9 were applied respectively using the methods described in section 4.3.5. After lysis of insoluble WT-CCL5-MBP fusion protein lysates by sonication. These detergents resulted to a recovery of a slightly increased proportion of recombinant WT-CCL5-MBP fusion protein soluble fraction after refolding (figure not shown) for 6M Urea and a significant release of recombinant WT-CCL5-MBP fusion protein soluble fraction after refolding, when 8M Urea, 0.5% Tween 20, 10mM DTT in 20mM Tris pH 9 solubilizing detergent was applied.

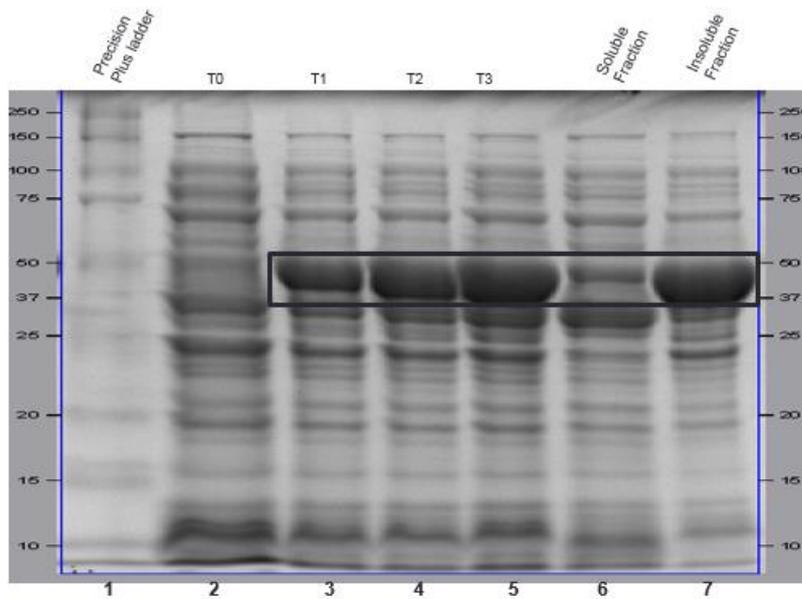


Figure 4.12 SDS-PAGE Protein Overexpression and Solubilisation Analysis of Wild-Type CCL5-MBP fusion protein expressions in DH5 α *E. coli* strain at 37°C in 0.5M Sorbitol and 1mM Betaine 2YT Media.

Lane1: precision plus marker, Lane 2 – 7 Wildtype CCL5-MBP expression profile showing post-IPTG (1mM) induction and overnight expression. T0 (L2), T1 (L3), T2 (L4), T0/N (L5), the soluble fraction (L6), insoluble fraction using 6M Urea (L7).

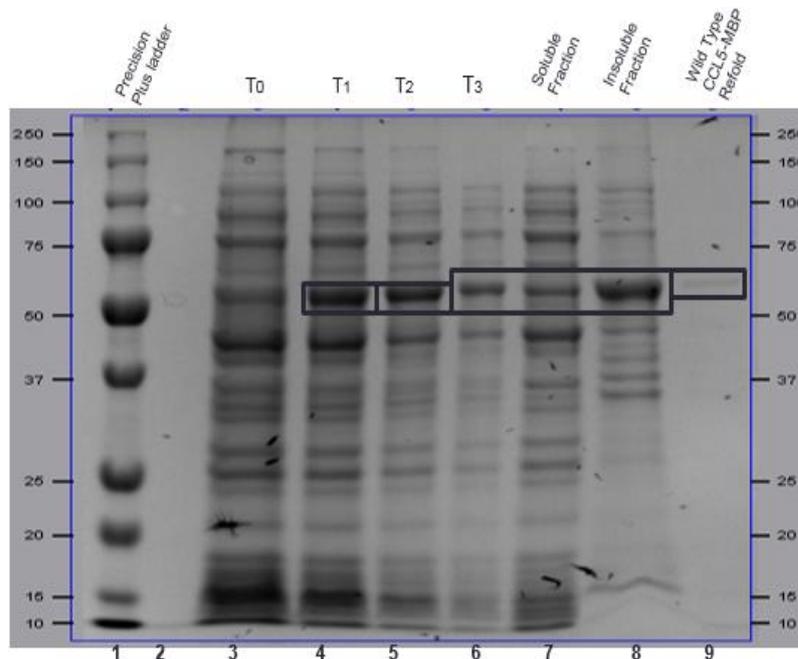


Figure 4.13 SDS-PAGE Protein Overexpression Analysis of Wild-Type CCL5-MBP fusion protein expressions in DH5 α *E. coli* strain at 37°C in 0.5M Sorbitol and 1mM Betaine 2YT Media.

Lane1: precision plus marker, Lane 3 – 7 Wildtype CCL5 expression profile showing post IPTG(1mM) induction. T0 (L3), T1 (L4), T2 (L5), T3 (L6), soluble fraction (L7), insoluble fraction using 8M Urea 0.5% Tween 20, 10mM DTT in 20mM Tris pH 9 (L8), Wild-Type CCL5-MBP Refolded Protein.

4.2.7 Batch Purification of WT-CCL5 and Mutant CCL5-MBP Fusion Protein

From an already prepared 50% slurry stock, 200 μ L of MPB resin slurry (Expedeon UK) was used to re-suspend 10 mL of soluble protein CCL5-MBP fusion lysate and was transferred into a universal tube. The MPB resin and CCL5-MBP fusion protein mixture were incubated for 1-hour on a roller at 4°C. After the 1 hour, the sample mixture was removed, centrifuged and the supernatants discarded. CCL5-MBP fusion batch purification, MBP resin pellet, was washed using 10 ml of PBS and incubated for 10 minutes at 4°C. The mixture was centrifuged at 6000rpm and supernatant stored, while the pellet was washed using 10 mL PBS, 10 min intermittently for a period of third washes. About 100 μ L supernatant of all samples was stored at 4°C.

Final elution was done using 200 μ L of 10mM Maltose, which was added to the re-suspended pellet by incubating on a roller mixer at 4°C for 15 minutes thereafter, centrifuged at 6000rpm for 5 min. The supernatant volume was stored at 20°C as the pure CCL5-MBP fusion protein.

All CCL5-MBP constructs were purified by batch purification, as described in section 2.6.1. Representative gels for batch purification were analysed by SDS-PAGE and are shown in figure 4.14A and 4.14B.

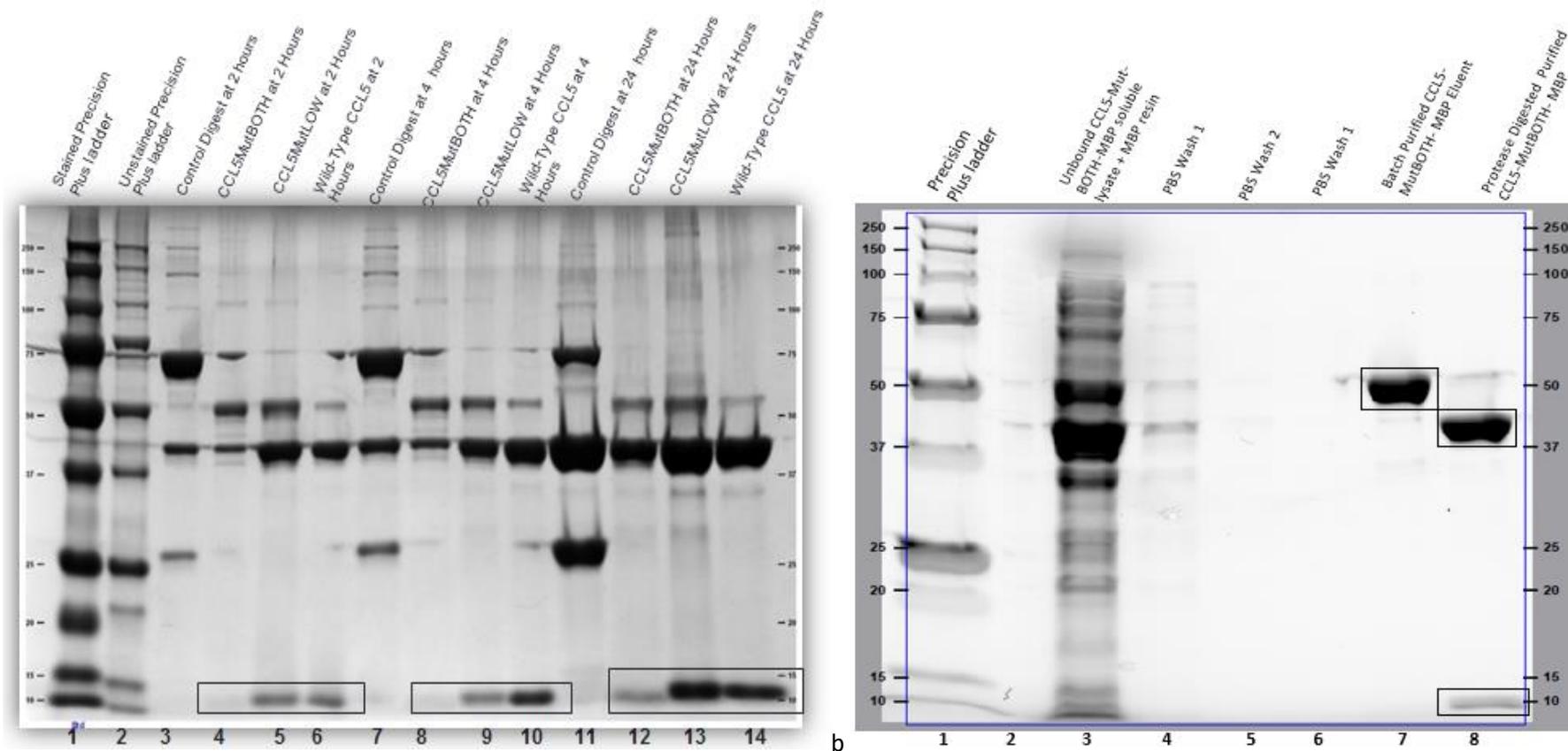


Figure 4.14 Cleavage Removal Analysis of CCL5 fusion proteins

(a) Time-Dependent Protease Cleavage of WT-CCL5-MBP, CCL5-MutLOW –MBP and CCL5-MutBOTH-MBP fusion proteins Lane 1: precision plus unstained marker, Lane 2: precision plus unstained marker, Lane 3 – 6: 2 hours MBP-cleavage digestion profile of Control digestion, CCL5-Mut-BOTH, CCL5-Mut-LOW and WT-CCL5-MBP. Lane 7 – 10: 4 hours MBP-cleavage digestion profile of Control digestion, CCL5-Mut-BOTH, CCL5-Mut-LOW and WT-CCL5-MBP. Lane 11 – 14: 24 hours MBP-cleavage digestion profile of Control digestion, CCL5-Mut-BOTH, CCL5-Mut-LOW and WT-CCL5-MBP. **(b)** MBP Resin Batch Purification and 24 hours Protease Cleavage of CCL5-MutBoth-MBP fusion protein: Lane 1: precision plus unstained marker, Lane 3: Unbound CCL5-MutBoth-MBP soluble lysate + MBP Resin, Lane 4: 1st PBS wash, Lane 5: 2nd PBS wash, Lane 6: 3rd PBS wash, Lane 7: Batch Purified CCL5-MutBoth-MBP recombinant fusion protein, Lane 8: Factor-Xa Digested CCL5-MutBoth-MBP recombinant fusion protein. Resolved by SDS-PAGE methods at 150 Volts.

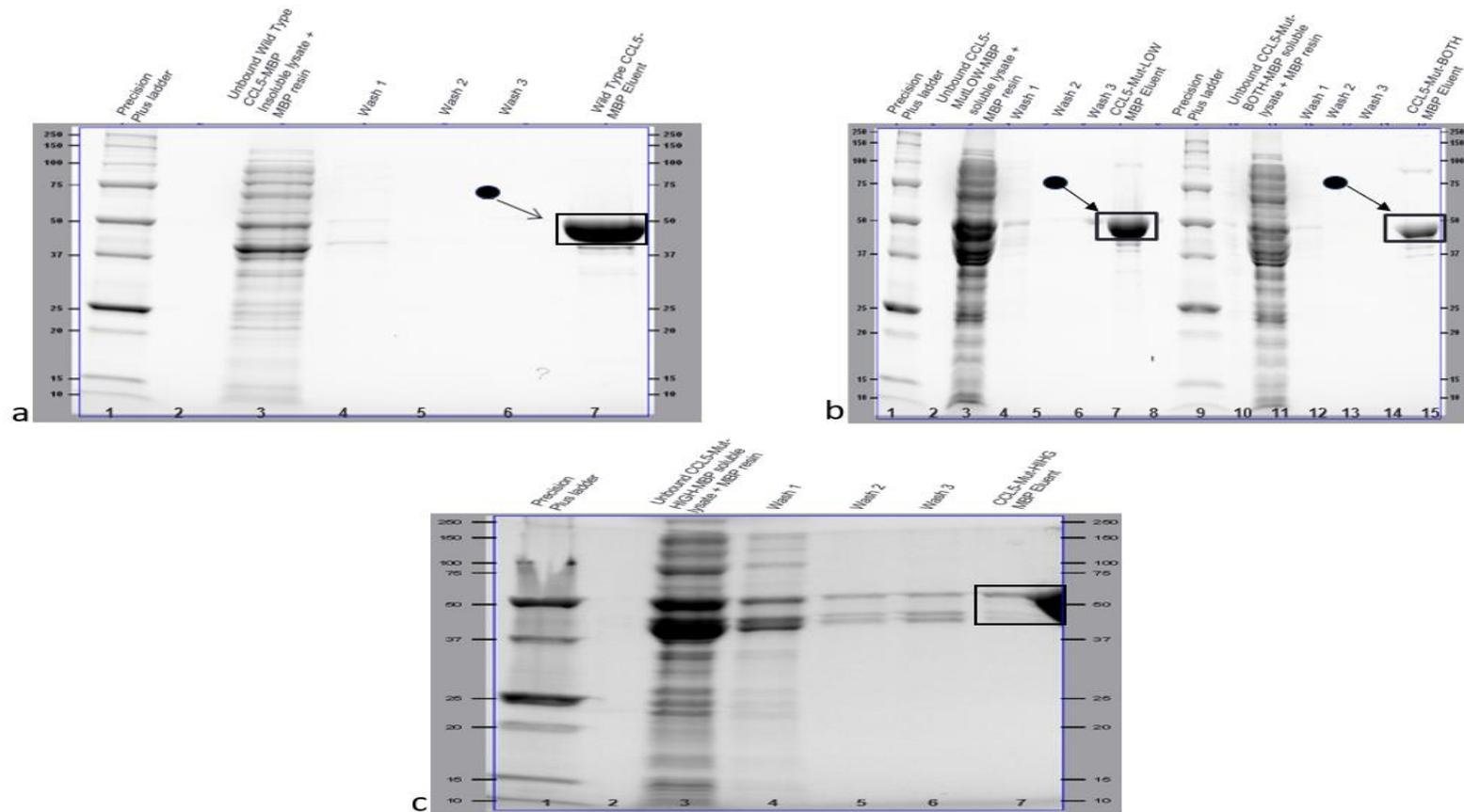


Figure 4.15: Batch Purification Analysis of Wild-type CCL5- MBP and Mutant CCL5-MBP fusion protein.

- (a) Lane 1: precision plus marker, Lane 3 – 7 Wild-type CCL5 batch purification profile showing. Unbound Wild-type CCL5-MBP + MBP resin (L3), 1st PBS wash (L4), 2nd PBS wash (L5), 3rd PBS wash (L6), and Wild-type CCL5-MBP bound eluent (L7). (b) Lane 1: precision plus marker, Lane 3 – 7 CCL5-MutLOW batch purification profiles showing. Unbound CCL5-MutLOW MBP resin (L3), 1st PBS washes (L4), 2nd PBS wash (L5), 3rd PBS wash (L6), and CCL5-MutLOW MBP bound eluent (L7) and Lane 9: precision plus marker, Lane 11 – 15 CCL5-MutBOTH batch purification profiles showing. Unbound CCL5-MutBOTH MBP resin (L11), 1st PBS washes (L12), 2nd PBS wash (L13), 3rd PBS wash (L14), and CCL5-MutLOW MBP bound eluent (L15). (c) Lane 1: precision plus marker, Lane 3 – 7 CCL5-Mut-HIGH-MBP batch purification profile showing. Unbound CCL5-Mut-HIGH-MBP + MBP resin (L3), 1st PBS wash (L4), 2nd PBS wash (L5), 3rd PBS wash (L6), and CCL5-Mut-HIGH-MBP eluent (L7). Resolved by SDS-PAGE methods at 150 Volt

After the batch purification of recombinant CCL5/MBP fusion-protein, a review of the SDS-PAGE analysis of molecular bands of purified eluents, demonstrated that the MBP sepharose batch purification method, recovered high yield of purified WT-CCL5-MBP, CCL5-MutLow-MBP, CCL5-MutHigh-MBP and CCL5-MutBoth-MBP fusion proteins during the batch purification process. The extent of the purity was to however, measured by UV-spectrophotometer at 280nm wavelength. All protein purification band pattern showed a gradual elution of protein of interest increased against incubation time and PBS wash. The precision plus molecular weight bands on the SDS-PAGE analysis attached in figure 4.15b – 4.16 showed migrations between 37 and 50 MW kDa marker, with the expected protein size and concentration of all recombinant CCL5/MBP fusion proteins. Their respective concentrations are summarised and presented in table 4.1 below.

Table 4.1: Experimental recombinant CCL5/MBP fusion protein molecular weight and protein concentration per 2 μ L (mg/mL) before protease cleavage and dialysis.

CCL5-MBP Fusion Proteins	Experimental CCL5-MBP Fusion protein	CCL5-MBP Fusion Protein Concentration (per μ L)
WT-CCL5-MBP	50.5kDa	3.54 mg/mL
CCL5-MutBoth-MBP	50.5kDa	0.28 mg/mL
CCL5-MutHigh-MBP	50.5kDa	14.63 mg/mL
CCL5-MutLow-MBP	50.5kDa	1.7 mg/mL

4.2.8 Cleave Removal of WT-CCL5-MBP and Mutant CCL5-MBP Fusion Protein

For structural and functional studies, the MBP Factor Xa protease was used to digest the CCL5-MBP fusion protein by cleavage removal, as explained in the Methods section above. This CCL5-MBP cleavage removal result showed three

lots of digested CCL5-MBP proteins, which included the wild type CCL5-MBP and two mutants CCL5-MBP (Mut-BOTH and Mut-LOW). The cleavage result indicated correct digestion of these CCL5-MBP proteins at different time intervals which include 4hours, 6hours and 24 hours period independently. The concentrations of these recombinant CCL5-MBP fusion batch purified proteins were measured, using a nanodrop (Wild-type CCL5-MBP = 4.2mg/mL, CCL5-MutLOW- MBP = 1.199mg/mL, CCL5-MutLOW-MBP = 3.12mg/mL). The sample concentrations per mg/mL were diluted to 1mg/mL concentration and used to calculate the total volume of pure CCL5-MBP lysates required for cleavage removal. About 6 μ L of PBS was used to supplement the volume of pure WT-CCL5-MBP lysate, which was 14 μ L, makeup to 20 μ L. Thereafter, 1 μ L of 200 μ g/mL of factor Xa (NEB, UK #P80105) was added to the pure WT-CCL5-MBP lysate. 5 μ L of the CCL5-MBP sample mixture was used against a positive control called MBP5 paramyosin. The WT-CCL5-MBP lysate sample cleavage removal was digested for a 24hours period and samples were taken at the 4th hour, 6th hour and 24th hours respectively. 2.5 μ L of the loading buffer was added after cleavage was completed at every stage of sample withdrawal (4H, 6H, and 24H). MBP restricted sample had to be heat shocked for 5min and analyzed using SDS-PAGE analysis. The evaluation of the pure WT-CCL5-MBP fusion protein cleavage removal result was successful and indicated optimal cleavage of all pure CCL5-MBP fusion protein to be about 24 hours post-digestion (figure 3.16, lane 12,13 and 14).

4.3.0 Isolation of Recombinant CCL5-MBP Protein from Cleaved Fusion Protein

After a successful batch purification and protease (Factor Xa) cleavage digestion of CCL5/MBP expression lysates, recombinant CCL5/MBP fusion proteins were obtained. Based on the objective of using recombinant WT- CCL5 and mutagenic variant proteins for functional assays, we made systematic attempt, to isolate the interest proteins (WT-CCL5, CCL5MutBoth, CCL5MutHigh and CCL5MutLow) from its fusion MBP tagged complex. The focus of eliminating maltose binding protein, is approached here by four different techniques. Firstly, by reverse MBP-resin batch purification as described 2.7.1, secondly by ultracentrifugation using a size differential Amicon Ultra-15 centrifugal device as described in section 2.7.3 and thirdly by a modified application of hypotonic lysis by buffer exchange. Here, 50mM Tris pH 8.5 150 mM NaCl and 20mM Tris, 1mM EDTA, 1mM DTT, 5mM MgCl₂ (sample lysis buffer) were applied to rid the fusion protein of its previous PBS buffer, this was to be achieved by modifying method for Ultra-centrifugation as describe in section 2.7.4 (Spada *et al.*, 2019). Finally based on the principal of electrostatic affinity, a negatively charged Heparin resin was used via a modified batch purification method to attempt isolating chemokine-MBP fusion protein.

The SDS-PAGE analysis of the flow-through samples for the two CCL5 isolation methods presented similar outcome. Although a standard 30kDa ultracentrifugation was applied as a means of retrieving the native CCL5 protein isolate, the result at figure 4.15, showed that WT-CCL5-MBP in unite (a) failed to produce a recovery of native CCL5 protein on lane 9 and 10, when incubate with MBP-resin for 1 hour via reverse-batch purification protocol although a successful

protease cleavage which had isolated the fusion MBP-tag from the recombinant CCL5 was achieved in lane 8. Additional step of incubating the protease cleaved WT-CCL5-MBP recovery with MBP- resin via reverse batch purification for 6 hours and 12 hours respectively before ultracentrifuge separation using 30kDa amicon device, showed that both the flow-through and supernatant recovery of the experimental product, on lanes 9, 10, 11 and 12 had no recovery of the recombinant native CCL5 protein.

The MBP-resin reverse batch purification of CCL5-MutBoth mutant (figure 4.16a) and CCL5-MutLow mutant (Figure 4.16b) showed that after factor xa protease cleavage in figure 4.16 (unit a), the 30kDa ultracentrifugation recovery (supernatant and flow-through) of MBP-resin reverse batch purification on lane 6, 7 and lane 14, 15 respectively, showed the absence of expected 8kDa recombinant native CCL5 mutant protein size, the SDS-PAGE however, indicated the presence of the 42kDa MBP.

Similarly, after substitution of PBS buffer with 50mM Tris pH 8.5 150 mM NaCl on reverse batch purification attempt as shown on CCL5-MutHigh mutant (figure 4.17a) and WT-CCL5 (figure 4.17b). The buffer exchange, alongside the MBP-resin reverse batch purification and 30kDa ultracentrifugation showed in figure 4.17 (unit b) lane 6, 7 and 14, 15 the absence of expected 8kDa recombinant native CCL5-MutHigh mutant and WT-CCL5 protein size respectively. The SDS-PAGE gel however indicated presence of the 42kDa MBP. The inability to isolate the CCL5 protein warranted the application of other techniques that were thought to be able to achieve the desired recombinant native CCL5 protein isolation objective.

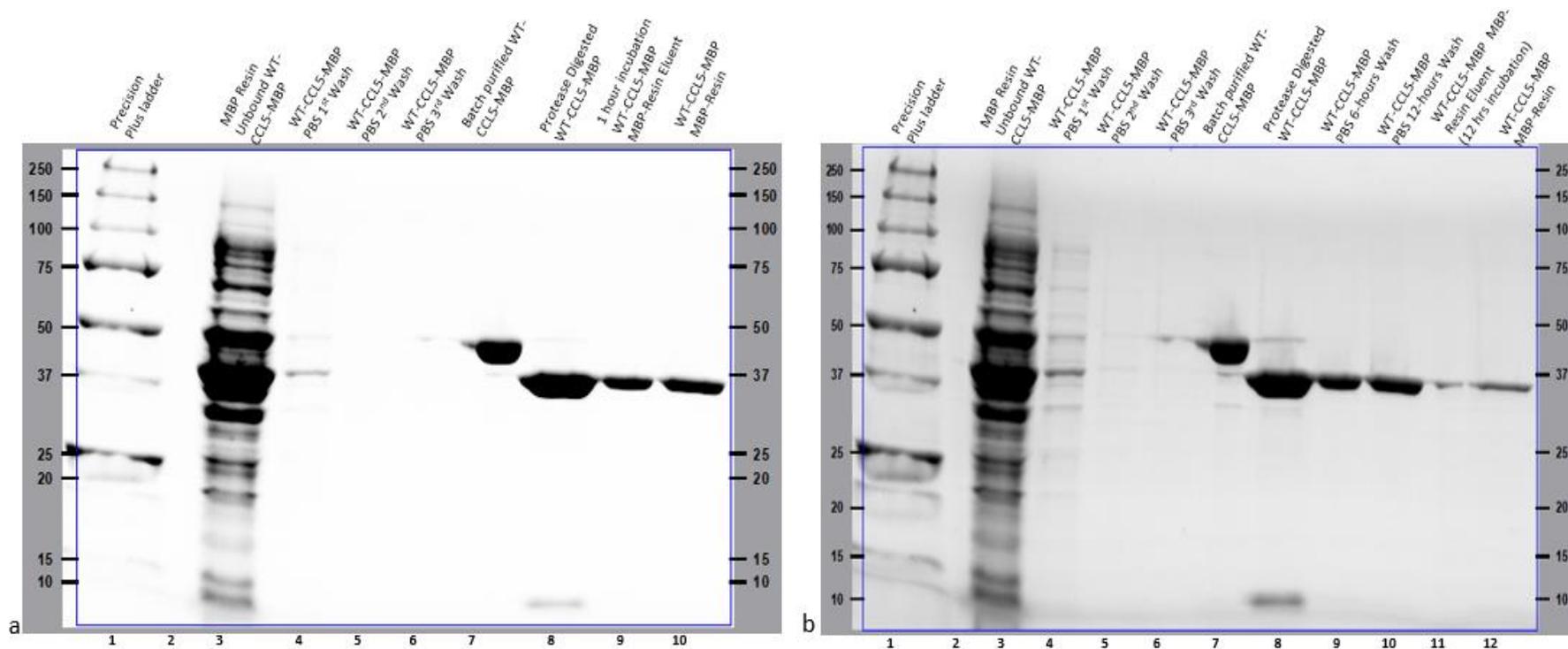


Figure 4.16 CCL5 Proteins Isolation Analysis Using Reverse MBP-Resin

- (a) Lane 1: precision plus marker, Lane 3-6: WT-CCL5-MBP, MBP batch purification profile, Lane 7-10: MBP isolation profile showing reverse batch purified WT-CCL5-MBP eluent (L3), Unbound WT-CCL5- MBP sample with MBP resin (L4- L6), 1st, 2nd and 3rd PBS wash of WT-CCL5- MBP sample (L7) batch purified WT-CCL5- MBP sample. (L8) factor Xa protease digested WT-CCL5- MBP sample (L9) 1 hours reverse MBP-resin incubated supernatant sample of digested WT-CCL5- MBP sample (L10) MBP-resin recovery from WT-CCL5- MBP reverse batch purification after amicon separation (b) Lane 1: precision plus marker, Lane 3 – 6: WT-CCL5-MBP, MBP batch purification profile, Lane 7-12: MBP isolation profile showing reverse batch purified WT-CCL5-MBP eluent (L3), Unbound WT-CCL5- MBP sample with MBP resin (L4- L6), 1st, 2nd and 3rd PBS wash of WT-CCL5- MBP sample (L7) batch purified WT-CCL5- MBP sample. (L8) factor Xa protease digested WT-CCL5- MBP sample (L9 -L10) 6 hours and 12 hours MBP-resin reverse batch purification of WT-CCL5- MBP sample (L11) supernatant sample of 12 hours MBP-resin reverse batch purification of WT-CCL5- MBP sample (L12) MBP-resin recovery from 12 hours MBP-resin reverse batch purification of WT-CCL5- MBP sample after amicon separation. Resolved by SDS-PAGE methods at 150 Volts.

4.3.1.1 Differential Size Exclusion Amicon Separation of CCL5/MBP Mutants

The separation of native CCL5 protein was attempted using 30kDa and 50kDa amicon-ultra-15-centrifugal-filter procedure like previous application as above. however, a different lysis buffer (20Mm Tris, 1Mm EDTA, 1mm DTT, 5Mm MgCl₂) was applied via buffer exchange and observed for its effect in effect in facilitating the separation of soluble native CCL5 protein fraction from the recombinant fusion protein.

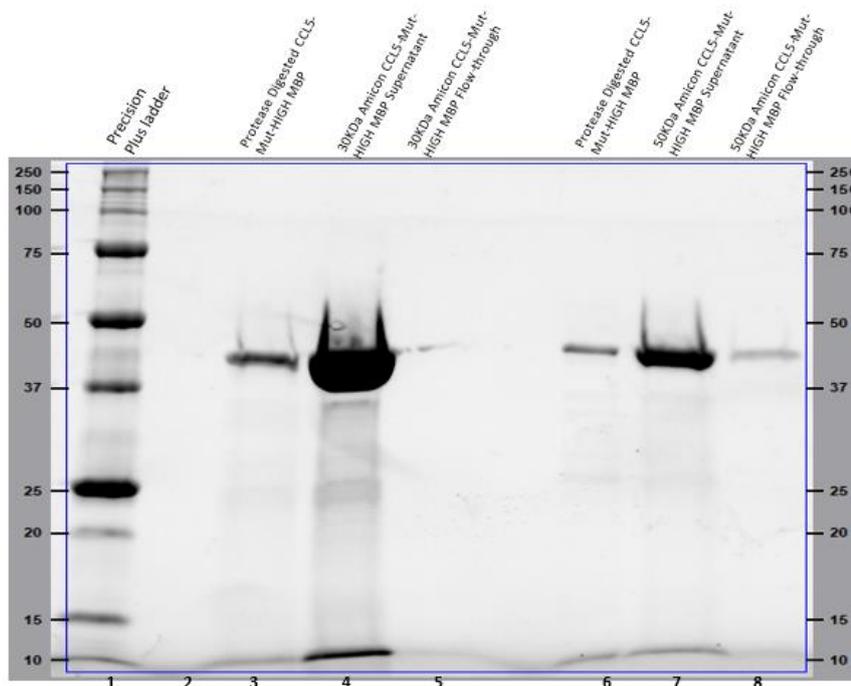


Figure 4.17 Differential Size Exclusion Amicon Ultra-Filtration Separation of CCL5-MBP Mutants and 2nd Buffer Exchange

Lane 1: precision plus marker, Lane 3 – 5 CCL5-MutBOTH isolation profile showing: Factor-Xa digested CCL5-MutBOTH – MBP sample (L3), 30 kDa Amicon centrifugation supernatant CCL5-MutBOTH – MBP sample (L4), 30kDa Amicon Flow-through CCL5-MutBOTH – MBP sample (L5), Lane 6 – 8 CCL5-MutHigh isolation profile showing: Factor-Xa digested CCL5-MutHigh – MBP sample (L6), 30 kDa Amicon centrifugation supernatant CCL5-MutHigh – MBP sample (L7), 30kDa Amicon Flow-through CCL5-MutHigh – MBP sample (L8), applying **20mM Tris, 1mM EDTA, 1mM DTT, 5mM MgCl₂** as buffer. Resolved by SDS-PAGE methods at 150 Volts.

Result showed a consistent loss of the native CCL5 protein in both 30kDa and 50KDa differential size exclusion amicon separation. Showing also that the application of 20mM Tris, 1mM EDTA, 1mM DTT, 5mM MgCl₂ buffer, may not have been effective in this context.

4.3.1.2 Heparin Resin Batch Purification

Heparin-resin based separation was applied to a sample of protease digested WT-CCL5-MBP via a modified heparin binding assay to isolate native WT-CCL5 protein, leaving behind MBP-fusion which ordinarily should have some affinity for heparin. The ultimate loss of WT-CCL5 protein and MBP tag after the second PBS wash is indicative that heparin resin batch purification method has not worked and not a viable technique for recovering native wild type CCL5 protein. There is currently, a lack of understanding within the context of this experiment as to the rational for the disappearance of the WT-CCL5 protein and MBP-tag.

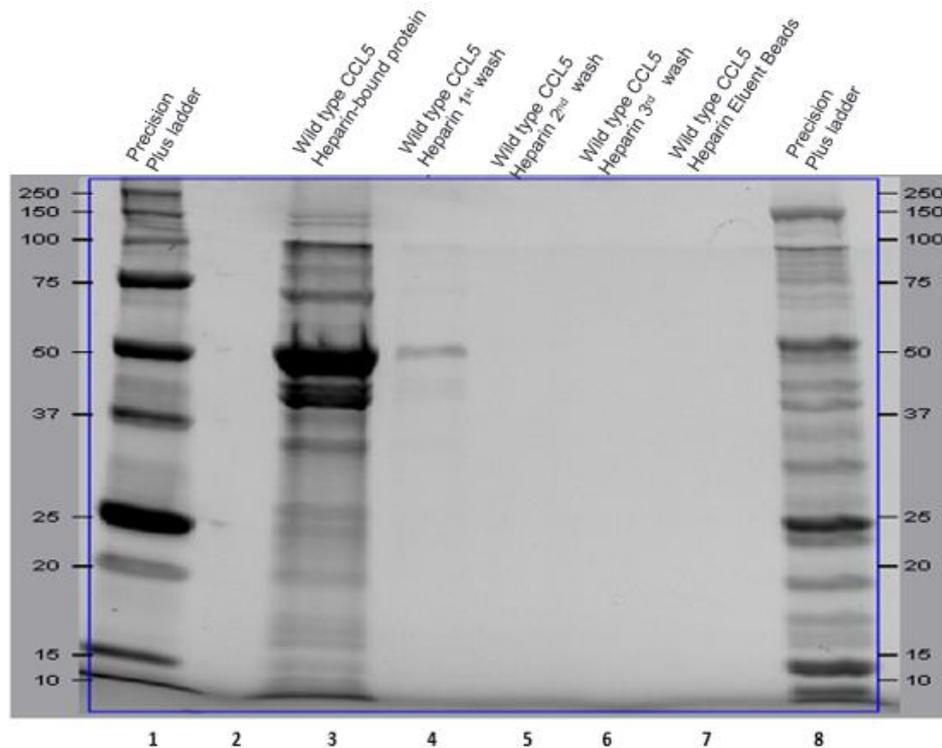


Figure 4.18: CCL5 Proteins Isolation Analysis Using Amicon Ultra Filtration and Buffer Exchange.

Lane 1: precision plus marker, Lane 3 – 8 WT-CCL5-MBP, heparin batch purification profile shows. Bound WT-CCL5-MBP + Heparin resin (L3), 1st buffer wash (L4), 2nd buffer wash (L5), 3rd buffer wash (L6), and WT-CCL5-MBP eluent Beads (L7), precision plus marker (L8), applying **20mM Tris, 1mM EDTA, 1mM DTT, 5mM MgCl₂** as a buffer. Resolved by SDS-PAGE methods at 150 Volts.

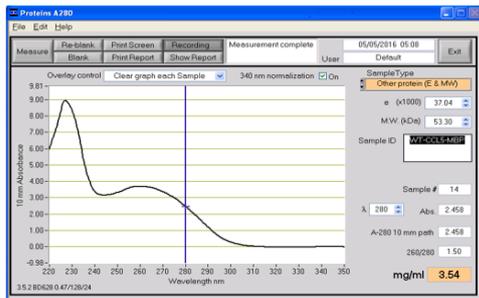
4.3.2 Sample Validation of Pure Recombinant CCL5-MBP Fusion Protein

After a successful protein overexpression, care was taken to ensure the absence of genomic DNAs and RNAs in the sample product. Ideally, cell lysis was introduced as the first step in cell fractionation (Walker, 2009; ThermoFisher.com, 2020). The mechanical lysis of protein expression lysates, being the final stage of physical disruption of the overexpression cell lysate, was lysed in this work via sonication to break down cells, hence leading to the release of nucleic acids. The presence of these cellular components may reduce the sample purity of the recombinant protein products (Hebron *et al.*, 2009; Tan and Yip, 2009). Although all experimental cells were lysed by sonication, no step was taken to guarantee the elimination of nucleic acids from the expressed recombinant protein sample, which

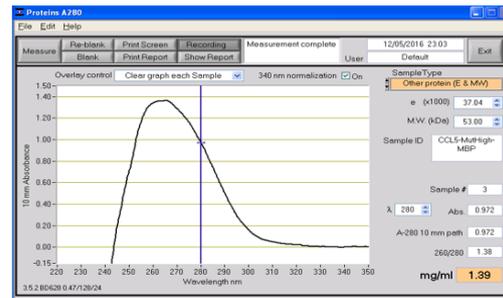
ultimately would reduce the risk of shearing genomic DNAs and RNAs, and the ultimate concerns of low protein yield (embl.de, 2020). A preliminary purity concentration assessment of batch purified wild type CCL5-MBP/mutants at 280nm absorbance under the ultraviolet spectrum, confirmed the presence of these nucleic acids, as the recombinant protein sample peaked at 230nm and 260nm absorbance wavelength using methods described in section 2.8.0 and results shown in figure 4.18.

A subsequent enzymatic digestion by adding DNase and RNase into a purified recombinant wild type CCL5-MBP/mutants protein sample and treating them using the dialysis method described section 2.8.0. showed the wild type CCL5-MBP/mutants protein samples to recovered from genomic contamination as sample demonstrated peak absorbance of ultraviolet light at 280nm, as shown in figure 4.19 and consistent with expectation protein absorbance ultra-violet spectrum for protein (Anthis and Clore, 2013).

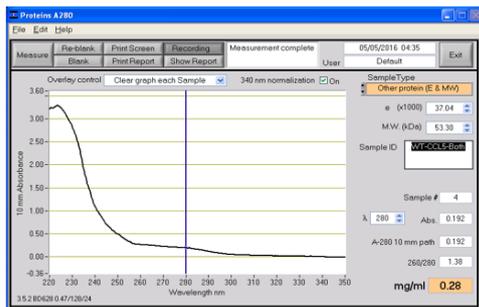
The application of contaminated proteins may affect the robustness of functional assays (Wingfield, 2017). Because we aimed to achieve a few functional assays in this work, we attempted to ascertain the purity and protein integrity of both the control-CCL5 protein and the synthesized recombinant CCL5/MBP fusion proteins that have been synthesized.



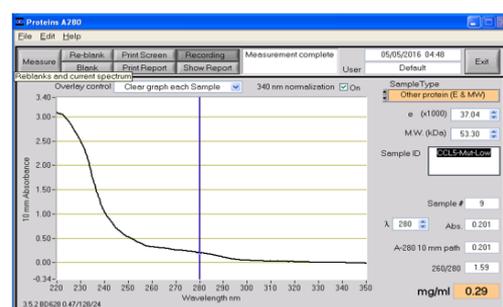
a



b



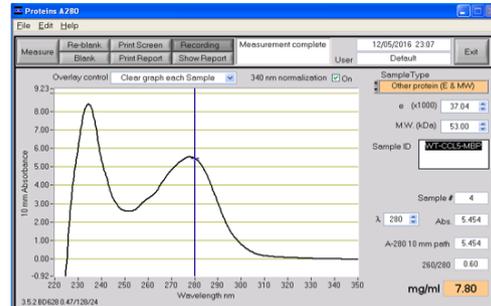
c



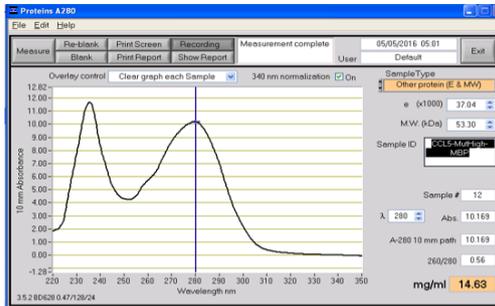
d

Figure 4.19: Crude and Non-dialysed Wildtype and Mutant CCL5 –MBP fusion proteins

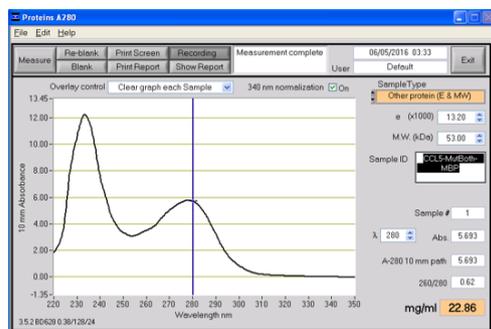
Showing clarified crude protein lysate sample of (a) WT-CCL5-MBP (b) CCL5-MutHigh-MBP (c) CCL5-MutLow-MBP (d) CCL5-MutBoth-MBP showing to absorb UV-spectrum light at A260. Protein concentrations were measured using a Nanodrop Spectrophotometer ND-1000 USA.



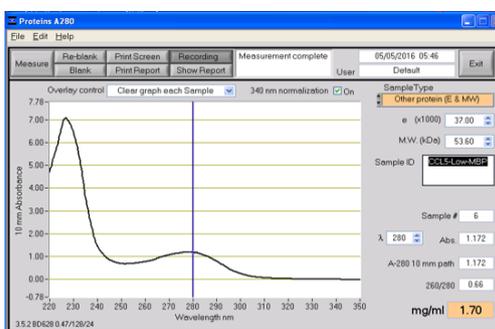
a



b



c



d

Figure 4.20: Purified and Dialysed Wildtype and Mutant CCL5 –MBP fusion proteins

Showing extracted protein lysate and Dialysed sample of (a) WT-CCL5-MBP (b) CCL5-MutHigh-MBP (c) CCL5-MutLow-MBP (d) CCL5-MutBoth-MBP showing to absorb UV-spectrum light at A280. Protein concentrations were measured using a Nanodrop Spectrophotometer ND-1000 USA.

Although we could now confirm that all experimental fusion proteins absorbed UV-light at 280nm wavelength, we proceeded to resolve the recombinant protein sample via SDS-PAGE analysis as shown in figure 4.20, which confirmed a clear presentation of all pure CCL5/MBP recombinant fusion proteins, alongside the control-CCL5. Other essential information that may determine the extent of robustness of the protein synthesis, could be the degree of marginal difference between the theoretically size of the fusion protein, in comparison with the eventual experimental size of the fusion recombinant protein expressed in the laboratory. The table 4.1 below compares the expected protein sizes either as contained in literature (protein databases) (www.expasy.org/resources/randseq) and the eventual experimental recombinant protein sizes. This comparison attempts to rationalize how much of success was achieved during the CCL5 protein synthesis and purification recovery of native recombinant CCL5. With the size of CCL5 known to 8kDa (Frauenschu *et al.*, 2007) and MBP size as 43kDa (Kapust and Waugh, 1999), the Wild type CCL5-MBP fusion protein size based on literature should be 51kDa. The result from table 4.1 showed that Wild type CCL5 and mutant CCL5 recovery to differ in sizes WT-CCL5 = 7.85kDa, Both-CCL5 = 7.42kDa, High-CCL5 = 7.62kDa and Low-CCL5 = 7.65kDa. It is observed however from result in figure 4.21 that the molecular weights of the experimented synthesized MBP-tag remained consistent at 43kDa. These observations conclude that the Alanine based mutation at the GAG binding site of the CCL5 proteins resulted in change of molecular size of the proteins and a high protein recovery after purification of the wild type CCL5 and its mutants.

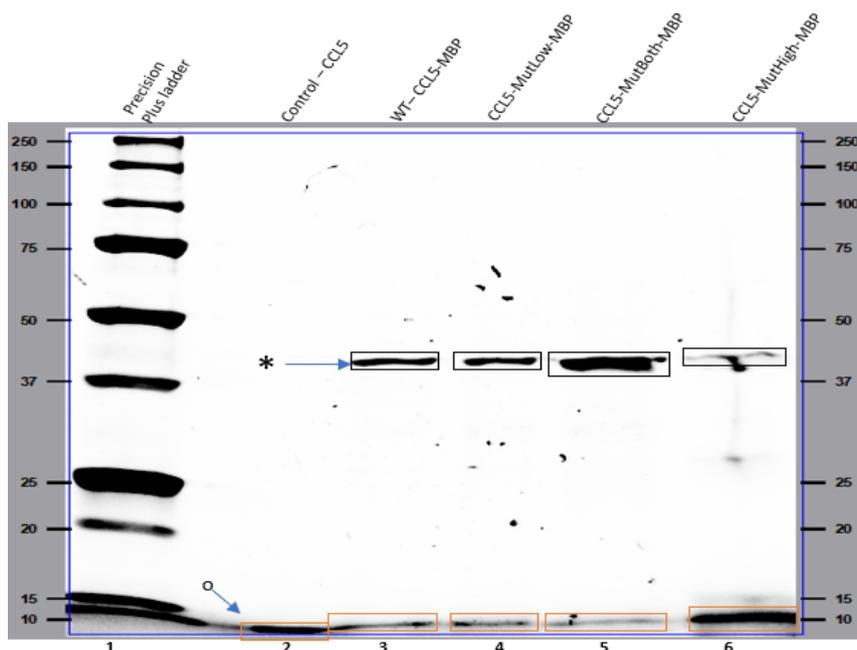


Figure 4.21: SDS-PAGE Analysis of Extracted and Dialysed Recombinant WT-CCL5 and Mutant MBP-Fusion Protein

Lane1: precision plus marker, Control-CCL5 (L2), WT-CCL5-MBP(L3), CCL5-MutLow-MBP (L4), CCL5-MutBoth-MBP (L5), CCL5-MutHigh-MBP (L6) after protein extraction and dialysis using the dialysis cellulose tubing. Resolved by SDS-PAGE methods at 150 Volts.

Table 4.1: Experimental recombinant CCL5/MBP fusion protein molecular weight and protein concentration (2mg/mL) post protease cleavage and dialysis.

CCL5-MBP Fusion Proteins	Experimental CCL5 Fusion Proteins (Size)	Protein Ext. coefficient (x1000) $M^{-1} cm^{-1}$	Protein Concentration
WT-CCL5-MBP	50.85 kDa	37.04	7.80 mg/mL
CCL5-MutBoth-MBP	50.42 kDa	37.04	22.86 mg/mL
CCL5-MutHigh-MBP	50.62 kDa	37.04	14.63 mg/mL
CCL5-MutLow-MBP	50.65 kDa	37.04	1.70 mg/mL

4.3.3 Biotin-Heparin Immobilization

The first step of this SPR experiment was to identify the ligand molecule and have them immobilized onto the surface of the sensor chip. Biotinylated heparin (Biotin-Heparin Sigma Aldrich UK) was chosen as the ligand and was immobilized via indirect coupling as presented in section 1.2.3b. The injection of the Biotinylated heparin was absorbed on a BIAcore SA sensor chip via non-covalent coupling using the method described in section 2.9.1.2 and previously established (Dyer *et al.*, 2016, Martínez-Burgo *et al.*, 2019). After which a signal of 517 RU was obtained after the reference flow cell was subtracted (fc 2-1), see figure 4.22. This immobilization signal was stable after completion of the blocking protocol to prevent non-reactive capture of the analytes by the activated streptavidin dextran matrix.

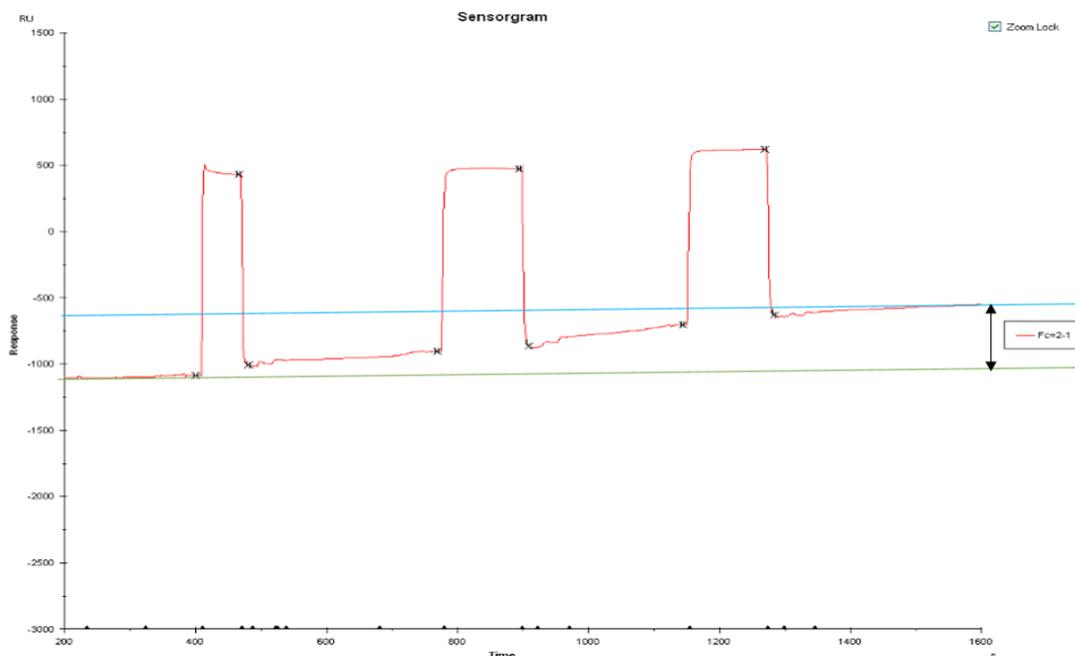


Figure 4.22: Surface Plasmon Resonance of Dialyzed Biotinylated Heparin Immobilisation

Sensorgram plot showing immobilization of dialyzed biotinylated heparin to streptavidin (SA) coated chip. Three injections were captured, which resulted in an increased immobilization signal of 517 RU. The increase in RU from baseline is shown by the green and blue lines respectively.

4.3.4 Binding Analysis of Recombinant CCL5-MBP Protein to GAG- Heparin

To assess the heparin-binding profile and apparent affinity of the synthesized WT-CCL5-MBP fusion protein and its mutants as earlier presented in this study, steady-state SPR binding kinetics studies were performed. The binding of each synthesized recombinant CCL5/MBP fusion protein was then studied, to evaluate affinity and kinetic (binding level association and dissociation) behaviour to GAG-heparin. A control experiment of heparin – control-CCL5 SPR manual mode injection, where a control outsourced pure-CCL5 (Sigma Aldrich UK) was injected over the heparin immobilized SA chip evaluate and subsequently confirm binding activity (figure 4.22).

However, revealed, as previously established by similar chemokine binding assay experiments, that the CCL5 binding interaction with the heparin surface demonstrated a fast K_{on} (association) and K_{off} (dissociation) constant (Segerer *et al.*, 2009, Dyer *et al.*, 2015, Dyer *et al.*, 2017). Therefore steady-state affinity was used as the method of determining

the apparent affinity of the WT-CCL5-MBP fusion protein and its available mutants against heparin. Apparent affinities were therefore calculated for association (k_a) and dissociation (K_d) rates using the model $KD = K_a / K_d$ and a 1:1 Langmuir model of steady-state affinity fitting dataset however, showed CCL5 approached saturation at high concentration. The propensity however for chemokines including CCL5 to oligomerize both at presentation to GAG and in solution, may mean that this model may not provide an accurate description of the CCL5-MBP/Heparin interaction. For analytical runs, solutions of control-CCL5 and all CCL5-MBP fusion proteins, were passed over the GAG-Heparin SA-chip for 120 seconds in an association phase at more than four concentrations (0.01 μ g/mL, 0.05 μ g/mL, 0.1 μ g/mL, 0.5 μ g/mL and 1 μ g/mL in buffer 10mM HEPES, 2M NaCl, 50 mM EDTA, 0.005% Tween-20) and (0.01 μ g/mL, 0.1 μ g/mL, 0.5 μ g/mL and 1 μ g/mL in buffer 10mM HEPES, 2M NaCl, 50 mM EDTA, 0.005% Tween-20) respectively. The analysis was continued in a dissociation phase of 500 seconds per cycle. Experimental binding data were adjusted by applying reference subtraction (fc 2-1), with all complied sensorgrams shown in figure 4.23 – 4.27.

The result showed an apparent successful immobilization of the biotinylated heparin on surface of the Biacore SA-chip, however not without figure 4.21 sensogram, indicating the presence of free-flowing biotin on the dialysed biotinylated heparin sample which had interacted with activated streptavidin dextran matrix surface, therefore resulting to competitive binding and a reduced immobilisation of the heparin. Although the biotinylated heparin had been dialysed using Slide-A-lyzer 2K (2,000 MWC) dialysis cassette (Thermo Scientific USA). A polypropylene device Zebra Spin Desalting Column 2K (2,000 MWC) made up of high-performing size -exclusion chromatography resin, have been shown to provide better free-biotin decontamination, as demonstrated when biotinylated heparan sulphate was de-salted and high yield recovery in a similar work (Clausen *et al.*, 2020). The limitation of the Slide-A-lyzer 2K (2,000 MWC) dialysis cassette,

is that the biotin had diffused back into the Slide-A-lyzer 2K (2,000 MWC) dialysis cassette permeable membrane.

Similarly, for an immobilisation level of 517RU, the binding association level observed with the control – CCL5 binding interaction after the heparin immobilisation on the SA-chip showed binding interaction at 380.7RU, which is a lower RU value than expected. Similar works have achieved better chemokine (CXCL8) R_{max} (RU) values (Martinez *et al.*, 2019). It is widely noted that electrostatic interaction becomes reduced in environments of high salt concentration (Lindman *et al.*, 2006; Seo, Girard, de la Cruz and Mirkin, 2019). The application of a 2M salt regeneration buffer (10mM HEPES pH 7.4, 2M NaCl, 50mM EDTA, 0.005% Tween - 20) after control -CCL5 was flown over the SA-chip, may have further reduced the expected binding level, hence apart from the control-CCL5, the result shows no binding for all other recombinant CCL5 wildtype and mutant proteins.

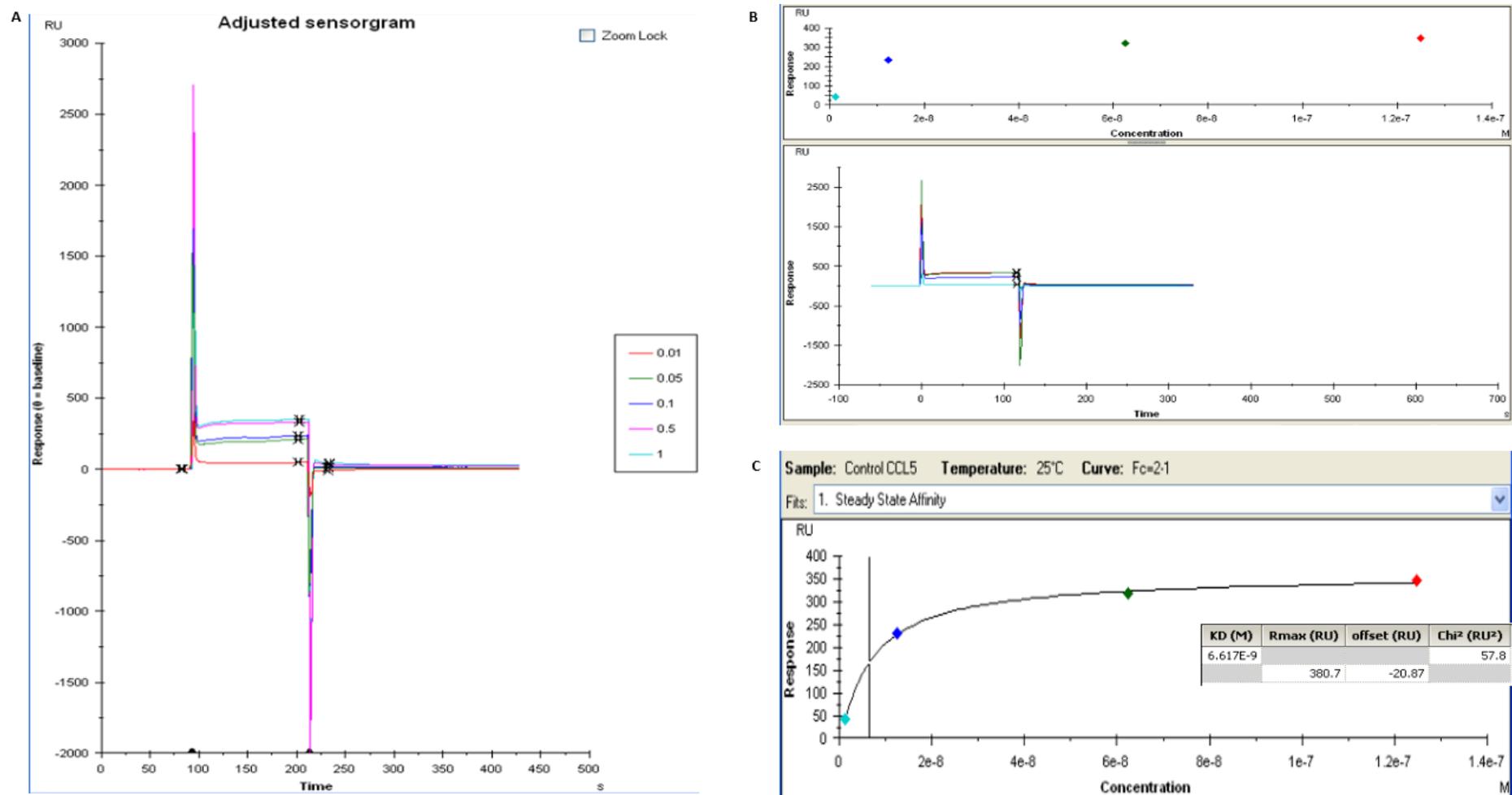


Figure 4.23: Surface Plasmon Resonance Steady State Affinity of Control CCL5 – Heparin Binding

- (a) Surface plasmon resonance sensorgram showing GAG-heparin: controlCCL5 binding. Each concentration (0.01 μ g/mL, 0.05 μ g/mL, 0.1 μ g/mL, 0.5 μ g/mL and 1 μ g/mL) of control-CCL5 chemokine were flowed at 30 μ L/mL over the heparin immobilized SA-sensor chip. (b) Prior to fitting, values of blank (buffer) were subtracted from the real experimental value and the steady-state affinity fittings previewed. (c) A steady-state affinity 1:1 Langmuir model fitted curve of control-CCL5 shows the measure of steady-state fit confidence.

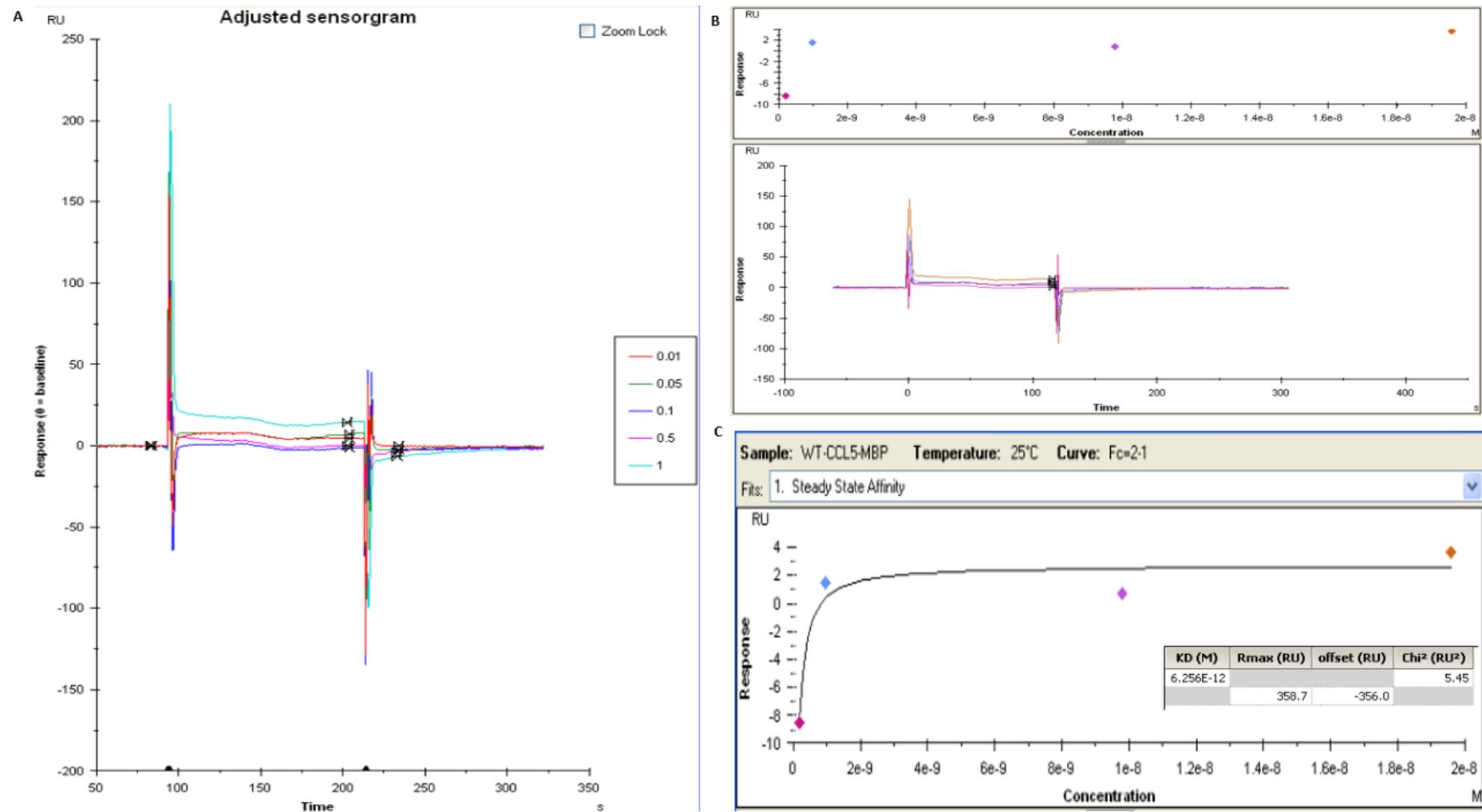


Figure 4.24: Surface Plasmon Resonance Steady State Affinity of WT-CCL5-MBP: Heparin Binding

- (a) Surface plasmon resonance sensorgram showing GAG-heparin: WT-CCL5-MBP binding. Each concentration (0.01 μ g/mL, 0.1 μ g/mL, 0.5 μ g/mL and 1 μ g/mL) of WT-CCL5-MBP chemokine were flowed at 30 μ L/mL over the heparin immobilized SA-sensor chip. (b) Prior to fitting, values of blank (buffer) were subtracted from the real experimental value and the steady-state affinity fittings previewed. (c) A steady-state affinity 1:1 Langmuir model fitted curve of WT-CCL5-MBP shows the measure of steady-state fit confidence.

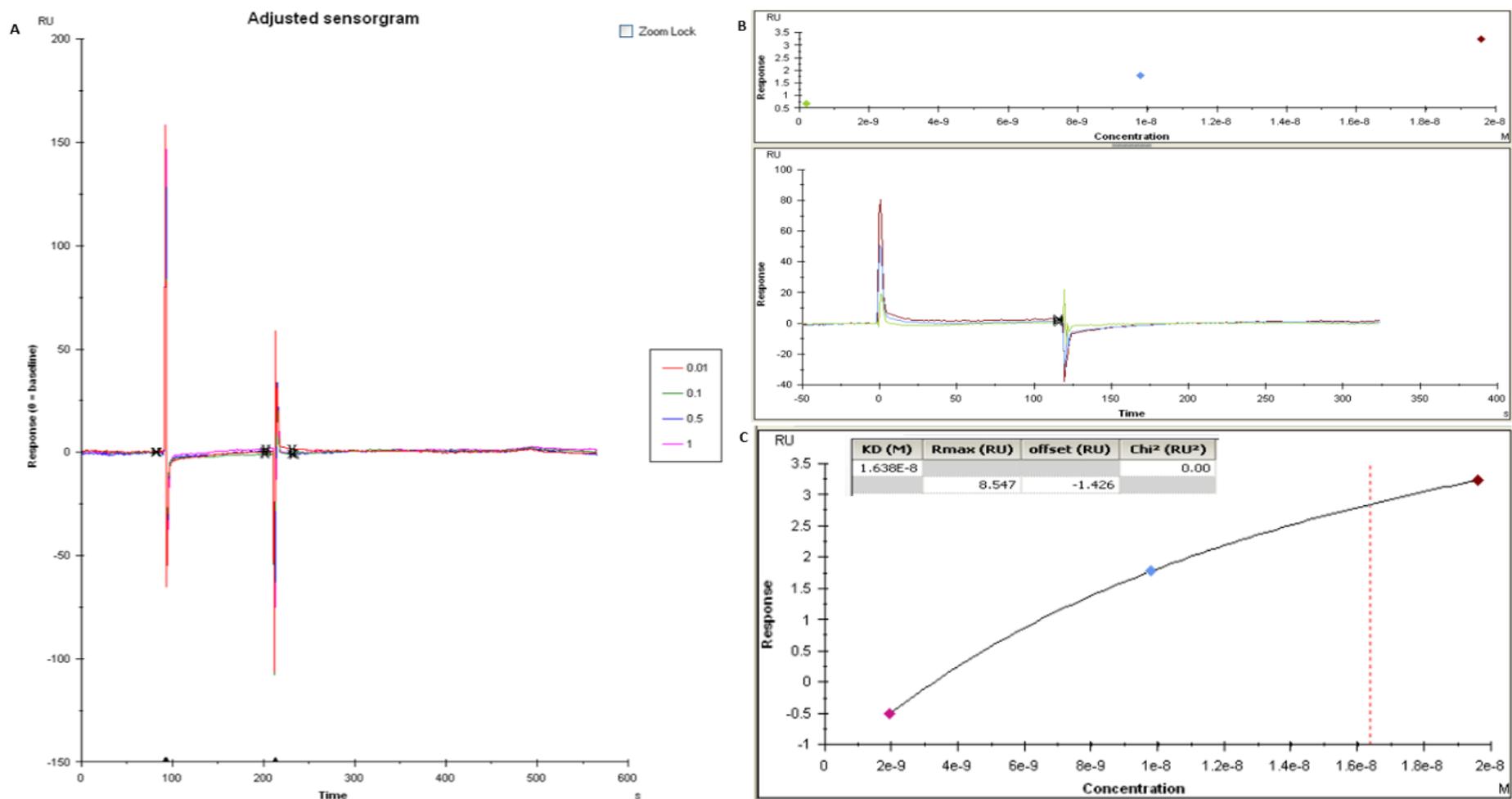


Figure 4.25: Surface Plasmon Resonance Steady State Affinity of CCL5-MutHigh-MBP – Heparin Binding

(a) Surface plasmon resonance sensorgram showing GAG-heparin: CCL5-MutHigh-MBP binding. Each concentration (0.01 μ g/mL, 0.1 μ g/mL, 0.5 μ g/mL, and 1 μ g/mL) of CCL5-MutHigh-MBP were flowed at 30 μ L/mL over the heparin immobilized SA-sensor chip. (b) Prior to fitting, values of blank (buffer) were subtracted from the real experimental value and the steady-state affinity fittings previewed. (c) A steady-state affinity 1:1 Langmuir model fitted curve of CCL5-MutHigh-MBP superimposed in black and the straight vertical red dotted line shows the measure of steady-state fit confidence.

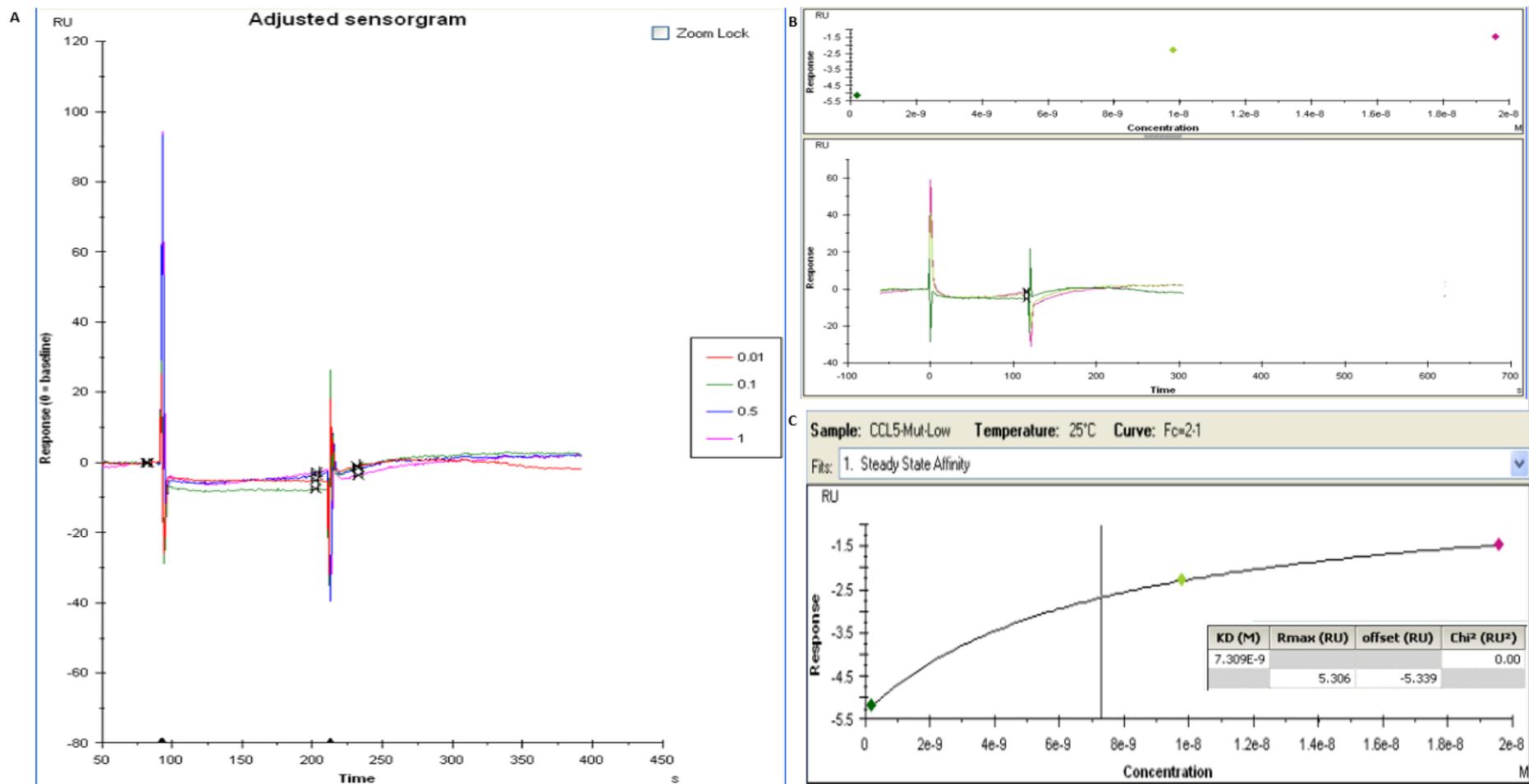


Figure 4.26: Surface Plasmon Resonance Steady State Affinity of CCL5-MutLow-MBP – Heparin Binding

- (a) Surface plasmon resonance sensorgram showing GAG-heparin: CCL5-MutLow-MBP binding. Each concentration (0.01µg/mL, 0.1µg/mL, 0.5µg/mL and 1µg/mL) of CCL5-MutLow-MBP were flowed at 30µL/mL over the heparin immobilized SA-sensor chip. (b) Prior to fitting, values of blank (buffer) were subtracted from the real experimental value and the steady-state affinity fittings previewed. (c) A steady-state affinity 1:1 Langmuir model fitted curve of CCL5-MutLow-MBP superimposed in black and the straight vertical black line shows the measure of steady-state fit confidence.

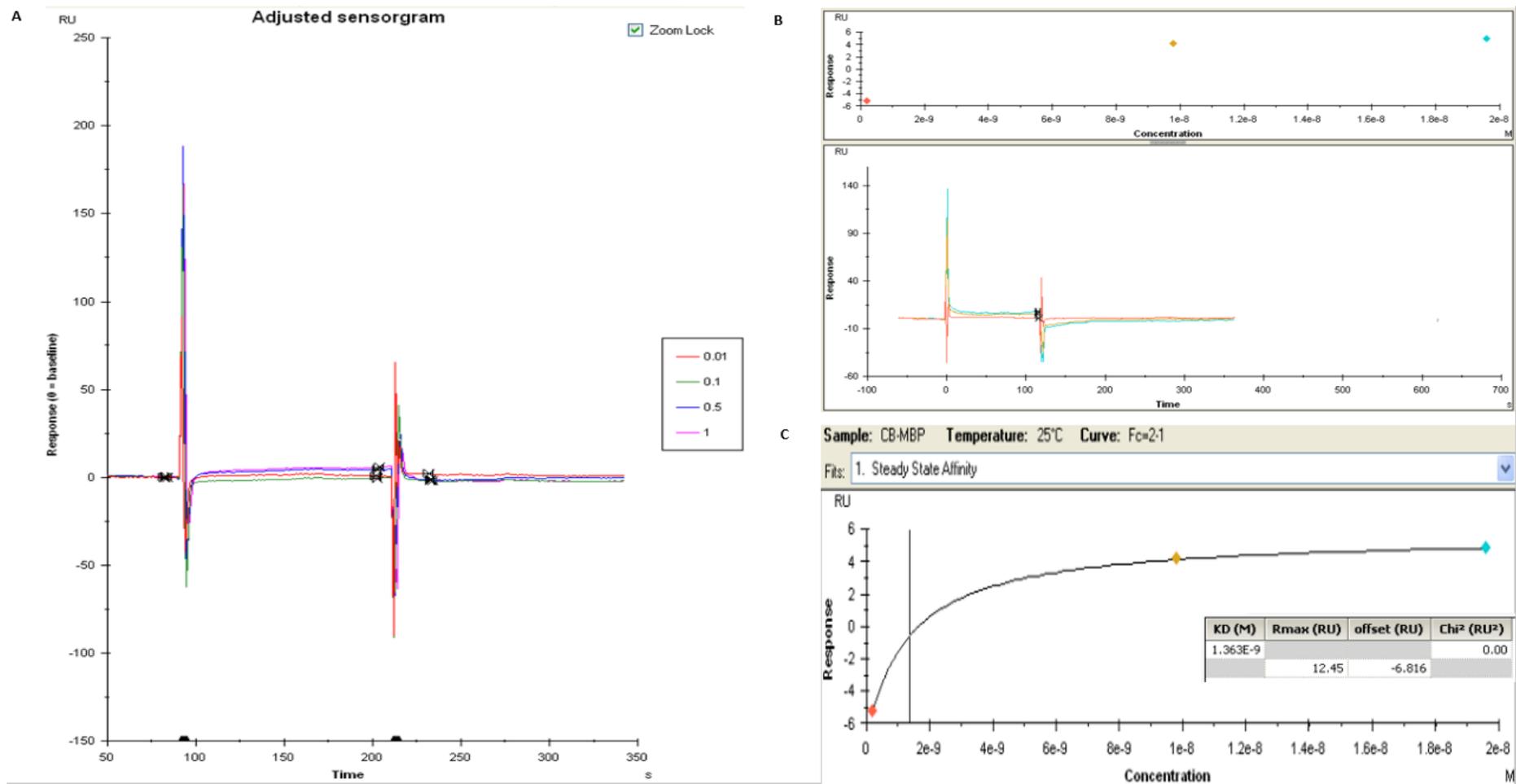


Figure 4. 27: Surface Plasmon Resonance Steady State Affinity of CCL5-MutBoth-MBP – Heparin Binding

- (a) Surface plasmon resonance sensorgram showing GAG-heparin: CCL5-MutBoth-MBP binding. Each concentration (0.01 μ g/mL, 0.01 μ g/mL, 0.5 μ g/mL, and 1 μ g/mL) of CCL5-MutBoth-MBP were flowed at 30 μ L/mL over the heparin immobilized SA-sensor chip. (b) Prior to fitting, values of blank (buffer) were subtracted from the real experimental value and the steady-state affinity fittings previewed. (c) A steady-state affinity 1:1 Langmuir model fitted curve of CCL5-MutBoth-MBP superimposed in black and the straight vertical black line shows the measure of steady-state fit confidence.

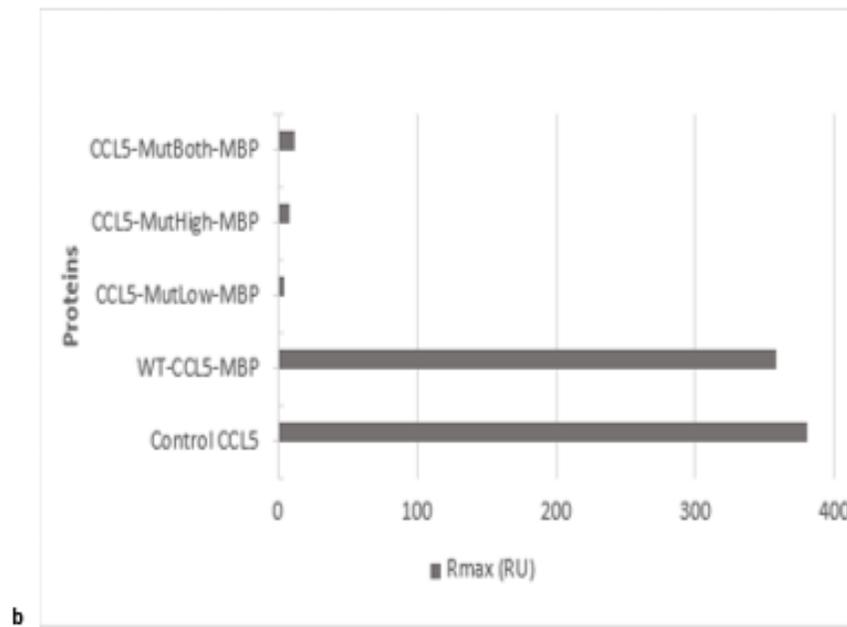
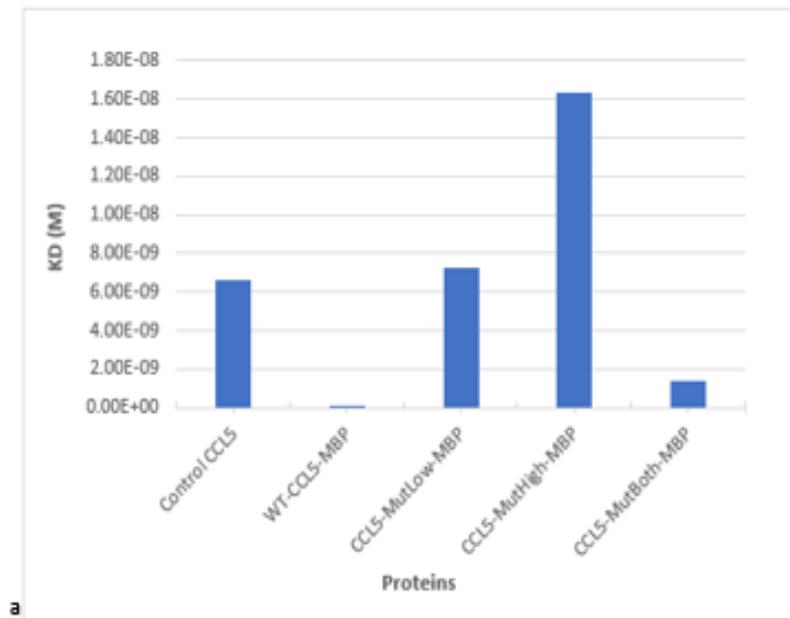


Figure 4.28 Summary of SPR CCL5/MBP – Heparin Binding Level Interaction Data

- (a) Bar chart show the measure of binding affinity KD(M) value of all CCL5-Heparin steady-state binding interaction, (B) Bar chart show over sample functional binding level activity at equilibrium binding interaction Rmax (RU).

4.3.5 Endothelial Migration Flow assay videos

A schematic representation of the endothelial biochip seeded with THP-cells, and subsequent THP-leucocyte cell flow-based adhesion is shown in Figure 4.29 and 4.30. Primary leukocyte adhesion in response to TNF- α and IFN- γ stimulated, human monocytic cell (THP-cell)-treated HMECs was used as positive control. Preliminary CCL5 mediated leukocyte flow-based adhesion to investigate the binding of endothelial GAG to two CCL5 peptides were assay in the presence of 500ng ml⁻¹ of wild-type CCL5-MBP and CCL5MutLow-MBP fusion protein. At 5 minutes the potential of the two peptides to block trans endothelial leukocyte migration was evaluated. We observed reduced leukocyte flow-based adhesion on the CCL5MutLow-MBP fusion protein mediated flow chamber. The number of cells adhesion were about 130 (figure 4.30) in the CCL5MutLow-MBP fusion protein mediated flow chamber, while the wild-type CCL5-MBP fusion protein mediated flow chamber had about 69 counts (see figure 4.29). The cell counting was done visually, by spotting presence of cells on flow chamber of different CCL5 peptide at time 5mins, that were not on the flow chamber at time 0 second. Cells were identified by their black outer membrane layer.

These experiments are considered preliminary works/data, towards understanding CCL5 behaviour, during leucocyte migration in the endothelia and only significant to the extent of more elaborate experimentation. The assay results can be viewed through the link below.

<https://www.dropbox.com/sh/3ubj6zI5c3ruhkl/AABOP4276syxZmb5iDXBB-QHa?dl=0>

<https://tinyurl.com/sr348pf>

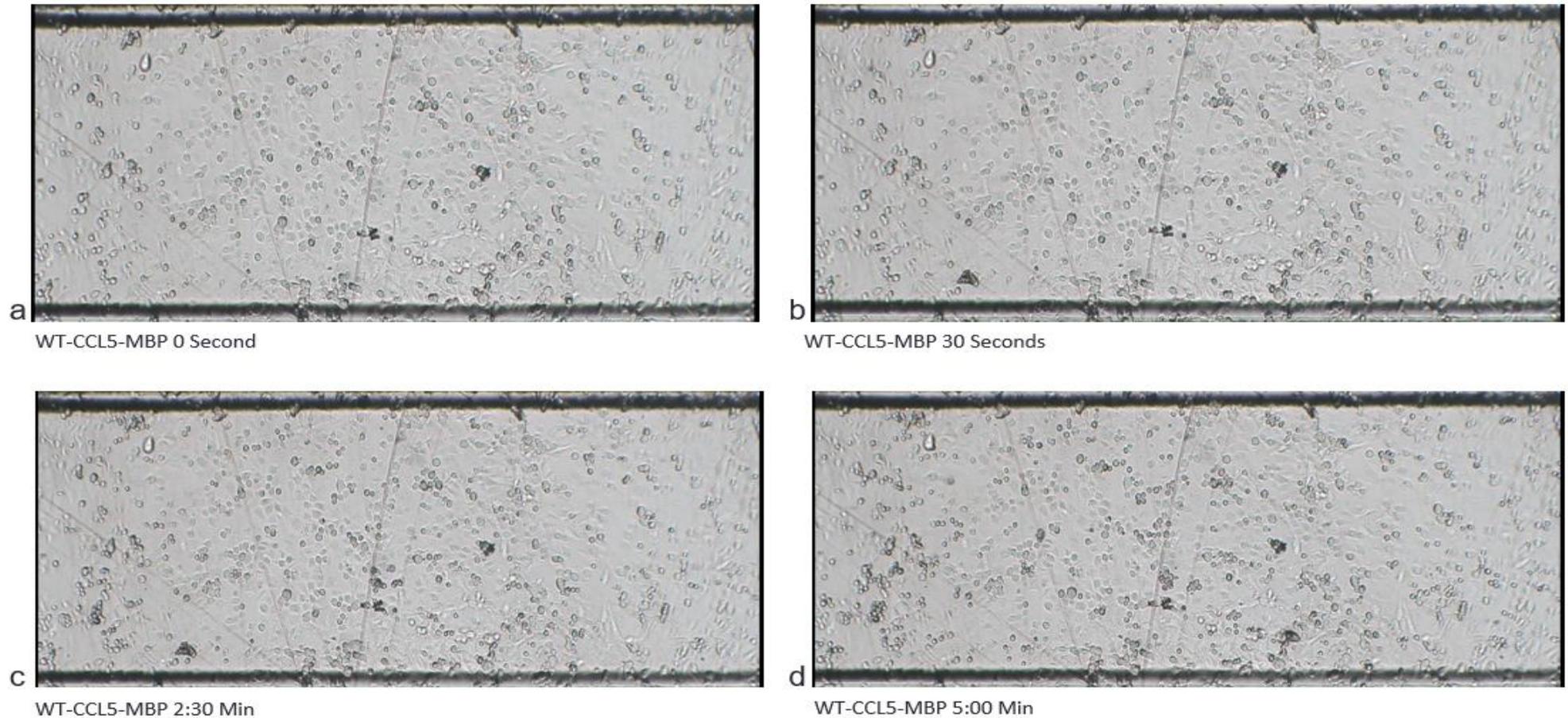


Figure 4.30: Schematic representation of THP-cell perfusion and adhesion over primary human microvascular endothelial cell line-1(HMEC-1) seeded with recombinant WT-CCL5-MBP

The HMEC-1 were first seeded over THP-1 monocyte coated biochip. The THP- leucocyte cells were loaded onto the HMEC-1 endothelial cells stimulated for 24 hours with TNF- α (10 ng ml⁻¹) and IFN- γ (10 ng ml⁻¹). Positive control is human monocytic cell (THP-cell)-stimulated HMEC-1 with 100 ng ml⁻¹ for the peprotech CCL5 or at 500 ng ml⁻¹ of the recombinant CCL5-MBP fusion protein 10 μ l of Bovine Fibronectin (5mg ml⁻¹ in 0.5M TBS pH 7.5, Sigma Aldrich, UK). The recombinant WT-CCL5-MBP (500 ng ml⁻¹) were added over THP-stimulated HMEC-1 and leucocyte adhesion was analysed after 5min of treatment. (a) At time 0 second of the THP-stimulated HMEC-1 flow on a TNF- α and IFN- γ stimulated endothelia cell surface, were the recombinant WT-CCL5-MBP had been seeded into a chamber. (b) At time 30 seconds (c) time 2.30 Minutes and (d) at time 5:00mineuts of flow analyses. Recorded in real-time using a GXCAM hichrome-lite digital camera (GT Vision, UK).

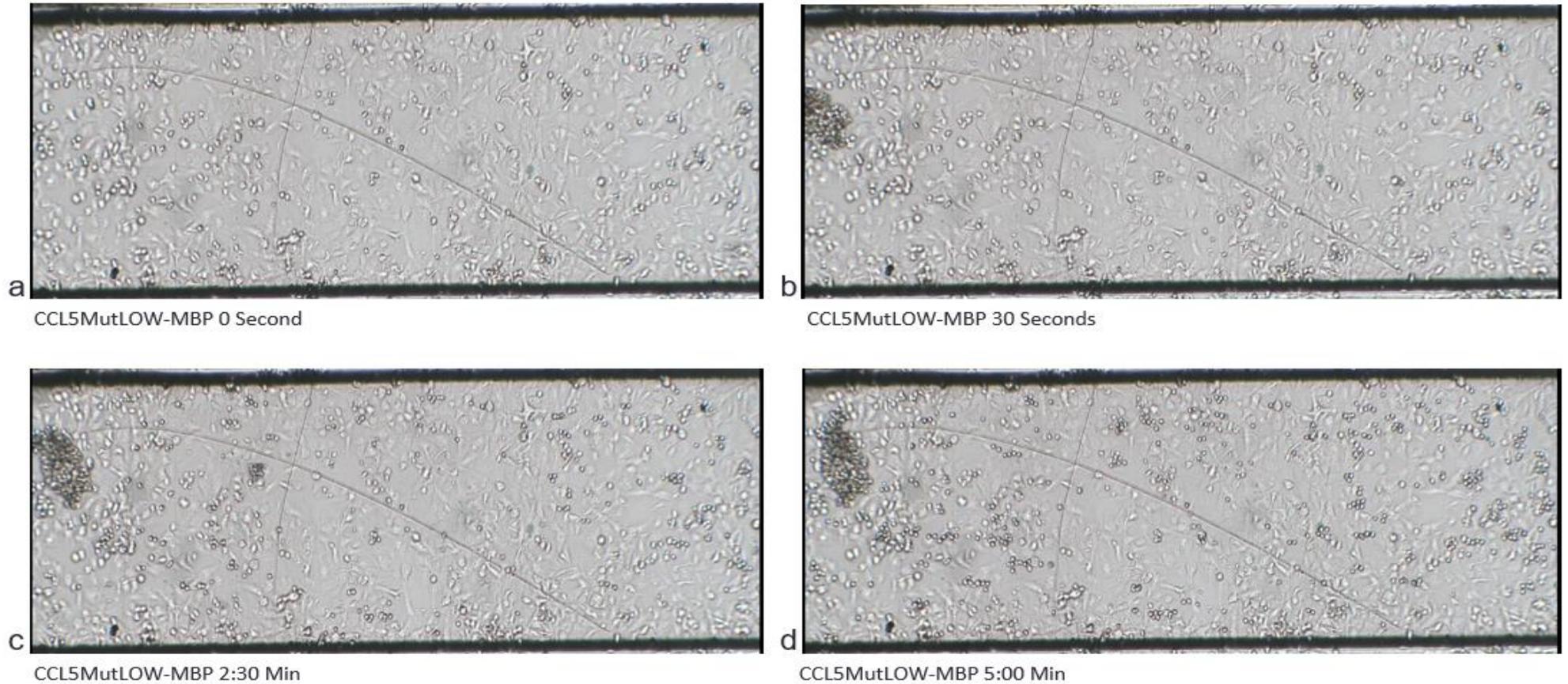


Figure 4.31: Schematic of THP-cell perfusion and adhesion over primary human microvascular endothelial cell line-1(HMEC-1) seeded with recombinant CCL5-MutLow-MBP (mutant peptide)

The HMEC-1 were first seeded over THP-1 monocyte coated biochip. The THP- leucocyte cells were loaded onto the HMEC-1 endothelial cells stimulated for 24 hours with TNF- α (10 ng ml⁻¹) and IFN- γ (10 ng ml⁻¹). Positive control is human monocytic cell (THP-cell)-stimulated HMEC-1 with 100 ng ml⁻¹ for the peprotech CCL5 or at 500 ng ml⁻¹ of the recombinant CCL5-MBP fusion protein 10 μ l of Bovine Fibronectin (5mg ml⁻¹ in 0.5M TBS pH 7.5, Sigma Aldrich, UK). The recombinant CCL5-MutLow-MBP (500 ng ml⁻¹) were added over THP-stimulated HMEC-1 and leucocyte adhesion was analysed after 5min of treatment. (a) At time 0 second of the THP-stimulated HMEC-1 flow on a TNF- α and IFN- γ stimulated endothelia cell surface, were the recombinant CCL5-MutLow-MBP had been seeded into a chamber. (b) At time 30 seconds (c) time 2.30 Minutes and (d) at time 5:00minutes of flow analyses. Recorded in real-time using a GXCAM hichrome-lite digital camera (GT Vision, UK).

4.4.0 Discussion

4.4.1 Introduction

Chemokines are critical mediators and have many important protective and developmental functions in a range of diseases such as asthma (respiratory disease), atherosclerosis and inflammatory bowel disease (inflammatory diseases), and multiple sclerosis (autoimmune diseases) (McNaughton *et al.*, 2018). These functions include cell trafficking and cell migration and have leukocyte non-directional mobilization and accumulation as a major component of chemokine activities in resolving tissue injury or inflammation (Sokol and Luster, 2015). McNaughton and co-worker submitted that the involvement of chemokines in regulating and activating leukocyte mobilisation becomes more characterized based on the necessities of chemokines ligands interacting with glycosaminoglycans on the surface of the endothelial cells (Metzemaekers *et al.*, 2018). A phenomenon that creates an environment of haptotactic gradient, previously verified to occur both in vitro and in vivo (Proudfoot *et al.*, 2017; McNaughton *et al.*, 2018). The CCL5 interaction with GAG-Heparin are essential towards understanding the potentials of CCL5-mediated anti-inflammation process. Dyer and co-workers have further validated that the formation of chemotactic gradients on endothelial cell surfaces and on the extracellular matrix (ECM), as results of the CCL5-GAG-Heparin interactions are in fact, effective in providing directional leukocytes mobilisation towards resolving inflammation (Dyer *et al.*, 2017).

The CCL5-GAGs binding affinity is mediated by strong electrostatic interactions, although the molecular and biological details of CCL5 – GAGs have been summarised by many researchers, as complex and are currently poorly/not understood (Dyer *et al.*, 2016; Thompson *et al.*, 2017; Proudfoot *et al.*, 2017).

What is presently known, is that the disruption of CCL5-GAG-Heparin interactions has shown to be effective in attenuating the unidirectional migration of leukocytes *in vivo* (Carter *et al.*, 2003; Proudfoot *et al.*, 2010). Previous studies have characterised the ⁴⁴BBXB⁴⁷ and the ⁵⁵BBXB⁵⁹ motifs as the GAG-heparin binding regions on the CCL5 protein structure (Johnson *et al.*, 2004; Segerer *et al.*, 2009; Dyer *et al.*, 2015). The standard methods for investigating these GAG-binding sites on the chemokine structure have typically involved mutagenesis of the chemokine domains. Previously studies have centred the investigation of CCL5 GAG-Chemotactic evaluation assay on the concept of neutralizing the physiological non-specific electrostatic interaction, naturally occurring between a positively charged CCL5 protein and the negatively charged polysaccharide GAG-Heparin molecule (Johnson *et al.*, 2005).

Segerer and co-workers, demonstrated, that the amino acid Alanine (Ala) being a neutrally charged amino acid was suitable for substituting the “B” amino acid unit of both the ⁴⁴BBXB⁴⁷ and the ⁵⁵BBXB⁵⁹ motifs on the CCL5 protein structure (Segerer *et al.*, 2009). The two codons optimised mutagenic sequences designed from wildtype CCL5 proteins by alanine substitution at 40s and 50s region respectively (CCL5-MutLow and CCL5-MutHigh) as described in section 4.2.1 are consistent with previous works (Chen *et al.*, 2018). An additional CCL5 mutagenic sequence the CCL5-MutBoth, bearing a simultaneous alanine substitution on both

the 40s and 50s GAG-Heparin binding region of the CCL5 protein structure was analysis in this work for its binding interaction with GAG-Heparin.

A fundamental question that we set out to address in this study, was whether, three CCL5 mutagenic proteins would demonstrate differential binding interaction with Heparin in vitro using the Surface plasmon resonance biophysical technique, and other functional assays despite the neutralisation of GAG-Heparin/CCL5 electrostatic interaction and if the results were similar to those previously reported (Dyer *et al.*, 2017, Martinez-Burgo *et al.*, 2019). And secondly, how these binding interaction data can inform the biology of chemokine function in resolving inflammation (Johnson *et al.*, 2004, Segerer *et al.*, 2009). Johnson and co-worker have previously demonstrated that the CCL5-⁴⁴AANA⁴⁷ mutant, lost 80% GAG binding capacity as compared to WT-CCL5 when inflammatory cells were induced into the peritoneal cavity and bronchoalveolar air space of a mice using a combination of in vivo and in vitro assays which are not similar to SPR (Johnson *et al.*, 2004). Similarly, Segerer *et al.*, (2009), collaborated that CCL5-MutLow-⁴⁴AANA⁴⁷ using a Boyden chamber assay, exhibited a chemotaxis migration potential, while CCL5-MutLow-⁵⁵AAWVA⁵⁹ mutation did not exhibit chemotaxis migration effect in a modified Boyden chamber.

The use of surface plasmon resonance as a major technique, among other assays, for the investigating of CCL5/GAG-Heparin binding interaction has been validated by other previous works on chemokine-GAG interaction (Salanga *et al.*, 2014, Dyer *et al.*, 2015). Martinez-Burgo and co-workers, had used surface plasmon resonance to investigate whether C-terminal CXCL8 peptide was implicated in GAG binding during neutrophil recruitment (Martinez-Burgo *et al.*, 2019).

A range of biophysical techniques apart from SPR, have been designed to investigate the interactions between GAG and chemokines. These techniques include heparin column-based affinity chromatography (Bolten *et al.*, 2018), isothermal fluorescence titration (Vanheule *et al.*, 2015), isothermal titration calorimetry (Sepuru, *et al.*, 2018), mass spectrometry techniques (Soe *et al.*, 2013), and NMR spectroscopy (Pomin, 2014; Deshauer *et al.*, 2015).

These experiments generally require cloning of wildtype CCL5 and mutant proteins, expression, and purification to evaluate the contribution of specific mutagenic CCL5 residue to GAG binding as previously demonstrated (Proudfoot *et al.*, 2017). The application of flow chamber assay was used to investigate a chemotactic flow-based control CCL5 and CCL5/MBP adhesion on stimulated endothelial surface under flow conditions.

4.4.2 Overexpression of recombinant Chemokine MBP fusion proteins

The experimental wildtype CCL5 and mutagenic proteins were designed to be cloned using a plasmid vector pMAL-C5x, which expresses an N-terminal fusion tagged maltose binding protein (MBP) periplasmic protein encoded by the *malE* gene in a *E. coli* expression system (Liu *et al.*, 2018). The final constructs of MBP-WT-CCL5, MBP-CCL5-MutLow, MBP-CCL5-MutHigh and MBP-CCL5-MutBoth were also predesigned using restriction endonuclease enzymes as described in section 4.2.4 which enabled success ligation. Wildtype CCL5 and mutagenic protein expression were induced in *E. coli* C43 (DE3) cells at 37°C for 2 hours. WT-CCL5 and mutagenic MBP fusion proteins were expressed with satisfactory soluble yields at 2 hours post-IPTG expression time after cell disruption via sonication, and removal of cellular debris by ultracentrifugation at 20,000 rpm for 20min at 4°C. Although

inclusion bodies were formed in the insoluble fractions (figure 4.10), the appreciable yield of soluble chemokine protein recovery, meant that further optimisation of the insoluble fraction was discontinued. Typically, the ratio of inclusion bodies observed when WT-CCL5 and mutagenic MBP fusion proteins were overexpressed overnight was higher (figure not shown). Similarly, a previous *E. coli* expression strategy adopted by Cho *et al.*, (2008) and Picciocchi *et al.*, (2014) utilised pMAL-C4x expression plasmid and *E. coli* BL21 (DE3), at 20°C post 1mM Isopropyl β -D-Thiogalactoside induction for 16 hours in producing a satisfactory quantity of recombinant MBP-CCL5-Strep protein. Deshauer and co, also synthesised a mature CCL5 with an E66S mutation using a pET23a expression plasmid and transformed into *E. coli* BL21 (DE3)pLysS, at 28°C post 0.5mM IPTG induction for 18 hours (Deshauer *et al.*, 2015). An investigation of CCL8 functional expression has successfully utilised pET28a-His₁₂-CCL8 plasmid for bacteria protein expression (Ge *et al.*, 2017).

The bacteria utilize a favourable intrinsic system to mitigate WT-CCL5 and mutagenic MBP fusion proteins from denaturation when in osmotic shock. The osmotic shock is induced by high external sugar concentrations (Barth *et al.*, 2000). Hence protein overexpression media (yeast extract media) had been supplemented with osmotic shock buffers 0.5M sorbitol and 0.5 mM betaine which increases the solubility of recombinant proteins in bacterial expression (Hofmann, 2015). Hence the reason, for the increased recovery of soluble recombinant WT-CCL5 and mutagenic MBP fusion proteins under the described overexpression condition.

4.4.3 Chemokine Protein Purification

Protein purification is an essential step towards characterizing recombinant proteins that have been synthesised for functional studies. To determine the extent of purity soluble recombinant WT-CCL5 and mutagenic MBP fusion proteins, an MBP sepharose batch purification technique with some modifications, was applied. The MBP sepharose, reacting with the N-terminally fused MBP-WT-CCL5 and mutagenic proteins, achieved high level recombinant protein purification from bacterial genomic DNAs. The result presented a series of actively purified recombinant chemokine and mutants' fusion proteins, which were verified by SDS-PAGE. Hofmann (2015) has described protein purification by batch experiment as an analytical chromatographic method. Purification of MBP- tagged proteins have previously been achieved satisfactory by amylose column purification (Karav, *et al.*, 2017). Similarly, Cho and co-worker purified MBP-RANTES in two steps, firstly via MBP affinity chromatography using MBP Excellose, and eluting fusion protein with 10 mM maltose, 20 mM Tris-HCl, 0.2 M NaCl, 1 mM EDTA, pH 7.4 and secondly through dialysis and IDA Excellose column purification via immobilized metal affinity chromatography (Cho *et al.*, 2008).

4.4.4 Removal of MBP-Tag

There is evidence to show however that the removal of MBP-tag during protein purification may result to protein to misfolding, hence, decreasing protein solubility (Zhao *et al.*, 2013). On the other hand, Wingfield (2015) insisted that tagging proteins with MBP necessitates increased expression levels and higher solubility of the target protein. Another CCL5 and CXCL12 functional

characterization assay successfully purified MBP-CCL5-Strep using StrepTactin column affinity chromatography, applying an extensive wash step after supplementing with 2.5mM desthiobiotin (Picciocchi *et al.*, 2014). This further validates the WT-CCL5 and mutagenic MBP fusion proteins batch purification as a robust purification technique.

Subsequently cleavage separation of the MBP tag from the recombinant chemokine fusion proteins were successful after convention application of Factor Xa protease digestion for 24 hours. The Factor Xa was found to cleave the recombinant WT-CCL5 and mutagenic MBP fusion protein substrates after the arginine residue site Gly – Arg of its preferred cleavage site Ile-Glu/Asp-Gly-Arg (Quinlan *et al.*, 1989, Kunji *et al.*, 2018). Vu and co showed previously that Factor Xa proteases was usefully used to purify and cleave a hCCL2 inflammatory chemokine, which had a maltose binding protein tagged at its N-terminal (Vu *et al.*, 2015).

The Factor Xa being compatible with the all pMAL tagged WT-CCL5 and mutagenic fusion protein separation and subsequent purification. The next step after protease digestion of the purified WT-CCL5 and mutagenic fusion protein, was to ensure a successful removal of the cleaved MBP molecule from the fusion complex. Here, a reverse MBP-resin batch purification as shown in figure 4.15a revealed that after one-hour MBP-resin incubation and centrifugation of purified protease digested WT-CCL5-MBP fusion protein, the 10mM maltose eluent and resin fraction, demonstrated a complete precipitation and absence of WT-CCL5 moiety, leaving only the MBP of 40.5kDa. Similarly, a twelve hours MBP-resin incubation of a purified and protease digested WT-CCL5-MBP fusion protein (figure 4.15b) also

revealed the absence of recombinant WT-CCL5-MBP, while retaining only the MBP fusion tag. There is substantial evidence however to demonstrate that MBP moiety of an MBP-CCL5 fusion protein, had previously been removed successfully via two steps. Firstly, by purification over an amylose column and subsequently by subjecting the unbound chemokine fraction to Superdex 200 gel filtration after the unbound chemokines had been concentrated (Picciocchi *et al.*, 2014). Although the MBP-isolation method adopted by Picciocchi and co-workers differs from the technique applied in this work only by the Superdex 200 gel filtration purification. The precipitation of experimental recombinant WT-CCL5 over many attempts, to remove the MBP-tag, can best be described as a cascade of unexplainable events. Several researchers have collaborated the difficulty in isolating MBP fusion tag from recombinant fusion protein post-protease digestion, insisting that gel filtration, and ion exchange chromatography ultimately could not isolate the MBP moiety because of a closely ranged P.I value of the MBP-tag and the MBP-resin (Adnan, 2020). Although the addition of His-tag to the N-terminally tagged MBP has been observed to satisfactorily eliminate MBP fusion protein after passing recombinant fusion protein over a cobalt or nickel column chromatography. For example, Hong and co-worker cloned a dual affinity pMAL-p2-AqpZ-HIS fusion protein (an N-terminal MBP tag and a C-terminal 8 x His TAG) and purified same using Ni-NTA agarose column chromatographic technique (Picciocchi *et al.*, 2014; Hong *et al.*, 2016). Followed by an IMAC purification were 500mM imidazole increased purity to 70% (Cho *et al.*, 2008). Other techniques employed to isolate MBP-tag from the WT-CCL5-MBP fusion complex are (i) independent ultracentrifugation using 30kDa and 50kDa molecular weight cut-off, (ii)

application of hypotonic lysis buffer via buffer exchange using the Amicon ultracentrifugation simultaneously, and (iii) applying Heparin-resin column batch purification. Ultracentrifugation technique, using a 30-kDa cut-off Amicon Ultra-15 centrifugal device was used through a series of modified applications, the rationale, was to fractionate the protease digested WT-CCL5-MBP fusion proteins based on their respective sizes. It was thought that since MBP was a 42.4kDa protein, under ultracentrifugation conditions, it would be retained on the upper chamber of the Amicon device, while other proteins lower than the 30kDa molecular size like the 8kDa WT-CCL5 will be recovered in the flow through. This concept presents to be easy and straightforward theoretically but did result to the collect loss of the WT-CCL5 in practical.

This perspective was previously proposed by a researcher that the 30 MWCO membrane cut-off centrifugation tube, could isolate any protein of less molecular weight than 30kDa (ResearchGate 2013). The ultracentrifugation device technique was similarly used to concentrate WT-CCL5-MBP and mutagenic fusion protein samples, and the application resulted a good yield of highly concentrated fusion protein products (Hang *et al.*, 2016).

There still was no evidence of MBP-tag isolation and recovery of WT-CCL5 proteins after hypotonic lysis buffer (50mM Tris pH 8.5 150 mM NaCl and 20mM Tris, 1mM EDTA, 1mM DTT, 5mM MgCl₂) were applied to recombinant WT-CCL5-MBP and mutagenic fusion proteins via a combined process of buffer exchange and ultracentrifugation using 30kDa and 50kDa Amicon ultra-filter device respectively. Spada and co-worker had previous demonstrated hypotonic lysis buffers were

efficient in isolating cytoplasmic protein fraction, in the order of (i) membrane disruption and (ii) centrifugation (Leung *et al.*, 2008; Spada *et al.*, 2019).

These two methods utilize the application of the same device which can perform purification of macromolecular compounds via molecular weight cut offs, protein dialysis and buffer exchange (Smejkal *et al.*, 2006; Amicon-Ultra-15-Centrifugal-Filter-Units, 2020). The continued lack of success in isolating MBP-tag necessitated the exploitation of the natural electrostatic interaction between a positively charged WT-CCL5, within the WT-CCL5-MBP fusion complex and the negative charge of heparin. A Heparin-resin based separation was applied to a sample of protease digested CCL5-MutBoth-MBP via a modified heparin binding assay or heparin batch purification. The heparin in this context, immobilised on an agarose resin served as a target carbohydrate used in batch purification mode, and aimed, to isolate WT-CCL5 protein, leaving behind MBP-fusion, which ordinarily should have some affinity for heparin. The heparin-based purification in this case, had no effect, on isolating the native WT-CCL5 protein. The ultimate loss of WT-CCL5 protein and MBP tag after the second PBS wash is indicative that heparin resin batch purification may not be a viable method, recovering native wild type CCL5 protein. There is currently, a lack of understanding within the context of this experiment as to the rationale for the disappearance of the WT-CCL5 protein and MBP-tag. Bolten *et al.*, (2018) however collaborated that heparin-based purification methods such as heparin affinity chromatography, shows characteristic specific interaction and compatibility with many protein purifications, example cytokines/cytokines (Ori *et al.*, 2011) and chemokines (XCL1) (Fox *et al.*, 2016) and PF4 (Platelet factor 4) (Imberty *et al.*, 2007).

The application of heparin-binding affinity tag, a recently introduced technique, to aid the recombinant protein expression of a wide range proteins. Jayanthi and co-workers experimented the production of the HBP-pET22b-tag vectors as a strategic intervention, which can easily be purified via Heparin sepharose chromatograph (Jayanthi *et al.*, 2018). These strategies were thought to facilitate the recovery of independent recombinant WT-CCL5, CCL5-MutLow, CCL5-MutHigh and CCL5-MutBoth. However, the ultimate results were negative and indicative that function assay was to continue otherwise without the removal of MBP-fusion tag. Some researchers have suggested that the presence of MBP fusion tag, may reveal addition information during functional analysis (Reuten *et al.*, 2016). Others insist, that the presence of MBP-tag as part of fusion complex, may interfere during functional assays especially as MBP tag has a significantly higher molecular weight (42kDa) than CCL5 (8kDa) in the recombinant fusion construct (Waugh, 2016). During recombinant CCL5 and heparin binding interaction functional assays, a control CCL5 without MBP was introduced and analysed alongside recombinant CCL5/MBP wild type and mutant proteins. The rationale was to evaluate binding, against the background of a positive control. Secondly to assess if and the extent of MBP-tag interference in during CCL5-MBP and heparin binding studies.

4.4.5 Biotin-Heparin Sample Validation for Surface Plasmon Resonance

Primarily wildtype and mutant CCL5/MBP fusion protein binding interaction with heparin was evaluated using surface plasmon resonance. Due to presence of free reacting biotin contamination in the heparin sodium salt sample (Sigma Aldrich USA), during initial immobilisation (data not shown). This resulted to unreactive

interaction and interferences between the streptavidin coated biosensor chip and the free biotin in solution during immobilization. A dialysis application of the contaminated biotin-heparin sigma stock sample using a Slide-A-Lyzer 2,000 MWC dialysis cassette as described in section 2.8.1 resulted to a recovery of biotinylated-heparin devoid of excess biotin contamination. The rationale here is based on the ability of the dialysis cassette to allow removal of low molecular weight cut-off of contaminant less than 2,000 MWC. In this context, biotin being a 244.32 MWC compound, was satisfactorily cut-off using the experimental protocol applied for dialysis, therefore providing a biotin-heparin of high integrity. Interestingly, Zhu and co-workers had similarly and previously applied this kit and technique, in removing free biotin via dialysis from a bioAAV and bioSoc molecules (Zhu *et al.*, 2019).

4.4.6 SPR interaction of Chemokine proteins in free solution with immobilized Biotin-Heparin

Functional evaluation of heparin binding interaction against wildtype and mutagenic CCL5 were modelled in vitro and in solution using a steady state affinity assay of the surface plasmon resonance, similarly, adopted by previous chemokine-GAG interaction investigations (Saesen *et al.*, 2013; Von Hundelshausen *et al.*, 2017, Chen *et al.*, 2019). Biotin-heparin immobilisation level of 517RU affixed irreversibly on the streptavidin chip dextran was satisfactory as binding interaction sensograms display no evidence of mass transport effect. Which is a key concern when binding is excessively immobilised on the surface of the sensor chip. Martinez and co-workers had previously attained a 200RU

immobilization level for biotin-heparin on a SA-chip, during a CXCL8 binding interaction investigation (Martinez *et al.*, 2019). Also, Saesen and co-worker, had previously immobilised heparin on SA-chip up to 800RU before binding using similar technique (Saesen *et al.*, 2013).

The SPR steady-state apparent affinity model value, for biotin-heparin binding interaction against control-CCL5, WT-CCL5-MBP, CCL5-MutLow-MBP, CCL5-MutHigh-MBP and CCL5-MutBoth-MBP; showed that equilibrium dissociation constant (K_D) value of the binding interaction assay was in the following order ranging from lowest K_D -value to highest K_D -value. WT-CCL5-MBP (6.26E-12), CCL5-MutBoth-MBP (1.36E-09), Control-CCL5 (6.62E-09), CCL5-MutLow-MBP (7.31E-09) and CCL5-MutHigh-MBP (1.64E-08). In principle, the lower the K_D -value during an apparent affinity SPR steady-state binding assay, the higher the measure of affinity.

Steady-state apparent affinity model equation

$$K_D = k_a/k_d$$

applying a 1:1 Langmuir model of steady state fitting

Parameter	Description
K_D	Equilibrium dissociation constant (M)
k_a	Association rate constant (M ⁻¹ s ⁻¹)
k_d	Dissociation rate constant (s ⁻¹)

These results suggest that, the recombinant WT-CCL5-MBP protein, demonstrated the highest apparent affinity with GAG-Heparin in this experiment, followed by recombinant CCL5-MutLow-MBP and CCL5-MutHigh-MBP. Although a control-

CCL5, was evaluated as well, its K_D -value was disproportionately higher than WT-CCL5-MBP by 308%, an indication that the cleaved MBP-tag of the wild type and mutagenic chemokine proteins may be intervening and accounted for within the sample mixture used in this binding affinity study. This would mean that within a sample solution of chemokine-fusion protein (WT-CCL5-MBP) the MBP-tag is present in an unknown proportion and may be competitive in the binding of immobilised Biotin-heparin. Two crucial parameters suggest that the experimental K_D -values may not be entirely accurate, they are (i) the role of MBP-tag interference in the binding affinity assay, (ii) the fact that steady-state affinity evaluation did not attain equilibrium, this can be corrected by evaluating additional units of higher concentrations (example: 2 μ g, 5 μ g and 10 μ g). Nonetheless, previous CCL5-GAG binding affinity interaction studies, have submitted similar conclusions to the results obtained in this experiment (Johnson, *et al.*, 2004; Liang *et al.*, 2016).

Similarly, the steady state fitting results in figure 4.22 - 4.26 indicates that the Biacore X100 SPR instrument can analyse the control-CCL5, WT-CCL5-MBP, CCL5-MutHigh-MBP, CCL5-MutLow-MBP and CCL5-MutBoth-MBP proteins within its limit of detection. Except for CCL5-MutHigh-MBP whose fitting was found to be outside the instruments limit of detection. The results also reveal that the CCL5 - heparin afferent affinity and binding interaction, are concentration dependent as when with control protein and more so at higher CCL5 concentration. Despite these limitations, the K_D - value of CCL5-MutBoth-MBP, has demonstrated an unexpected high affinity for GAG-heparin that need to be further investigated.

4.4.7 Endothelia Flow Chamber Assay

To understand the cell recruitment behaviour of WT-CCL5-MBP and CCL5-MutLow-MBP proteins when interacting with a TNF- α and IFN- γ stimulated endothelial receptor surface environment, we investigated the chemokines adhesion ability using transendothelial chemokine adhesion assay under flow conditions. Early result findings of this novel technique, showed CCL5-MutLow-MBP fusion protein mediated flow chamber, had about double amount of cell adhesion on the endothelial more than the WT-CCL5-MBP mediated endothelial flow chamber. Because only one concentration of 500 ng/ml was assayed, it is not clear whether this endothelial adhesion of leukocyte occurs in a dose-dependent manner. Activated endothelial cells express a range of chemokines that affect the recruitment of monocytes as well as T-cells rolling and adhesion to the vessel wall (Øynebråten *et al.*, 2015).

In previous similar works peritoneal cell recruitment assay has been used to investigate the effect of [44AANA47]-RANTES on CCL5-mediated peritoneal cell recruitment and adhesion across the endothelial. The findings showed [44AANA47]-RANTES was able to inhibit inflammatory cell recruitment in a dose-dependent manner (Johnson *et al.*, 2004). On the other hand, Severin, and co-workers, investigated GAG-analogues as a novel anti-inflammatory strategy utilizing peritoneal cell recruitment assay (Severin *et al.*, 2012). Other transmembrane adhesion assay methods that have been applied towards analysing the role of CCL5 in various endothelial cell migration, recruitment and adhesion behaviour are the Boyden transwell migration chambers (Maillard *et al.*, 2014; Wang *et al.*, 2015; An *et al.*, 2019). The advantage of flow-based chemokine

adhesion via flow chamber assay method over the peritoneal cell recruitment assay and Boyden transwell migration assay is that the cell recruitment and adhesion are evaluated in real-time underflow condition in the flow chamber assay, which mimics the in vivo adhesion, tethering and rolling behaviour of the chemokine – GAG interaction across the endothelium. In other transmigration methods, all vital parameters are evaluated under static or fixed condition.

Chapter 5

Rhodococcus equi: Virulence-Associated Protein

Chapter Five

5.0 Introduction

The first clinical occurrence of pathogenic *Rhodococcus equi* isolate has been traced to 1923, when Swedish foals, were found to be suffering from bronchopneumonia (Magnussun, 1923). *R. equi*, has been described as the major significant cause of a series of pulmonary and extra-pulmonary disorders in young foals. These disorders include but are not limited to, lung abscesses, and a range of sub-acute to acute suppurative bronchopneumonia. These conditions are often linked, with a high mortality rate amongst young foal, hence, interfering with the global growth of the horse-breeding industry (Muscatello, 2012; Vázquez-Boland et al., 2013). This coccobacillus Gram-positive pathogenic bacterium, *R. equi*, thrives in a range of environments, including dry, wet, and dusty soil. The presence of the bacterium in many domestic and farm animals has been a major source of pneumonia condition. Takai et al., (2003) and Riberio et al., (2011), agree that *Rhodococcus equi*, has been isolated from animals such as cats, sheep, dogs, wild boars, domestic pigs (Rzewuskwa et al., 2014), foals (Kalinowski et al., 2016) and cattle (Nakagawa et al., 2018). Increasing evidence suggests that foals within 4- 6 months of age have the highest record of *Rhodococcus equi* disease (*rhodococcosis*), with the greatest clinical manifestation of the disease, obvious at about four months of age (Meijer and Prescott, 2004). Cases of extrapulmonary disorders (EPDs) have been reported in horses and may be associated with or independent of bronchopneumonia (Reuss et al., 2009). The ingestion of *R. equi* pathogen may result in conditions such as diarrhoea, osteomyelitis, hyperthermia, typhlocolitis, lymphadenitis and immune-mediated hemolytic anemia (Letek et al.,

2010). Furthermore, these represent a range of first-line clinical abnormalities that may arise after the treatment of pneumonia. A study has also collaborated that extrapulmonary disorders, caused by *R. equi* pathogen, maybe the ultimate prompt for euthanasia in farm animals (Reuss *et al.*, 2009; Vazquez-Boland *et al.*, 2013).

In the absence of appropriate treatment of *Rhodococcus equi* infection, mortality rates have been found to reach 80% (Elissalde *et al.*, 1980; Cohen *et al.*, 2005; Coulson *et al.*, 2010; Kalinowski *et al.*, 2016). Reports have also indicated that foal deaths from pneumonia, are primarily caused by *Rhodococcus equi*, and accounts for about 3 % of foal mortality per annum (Oldfield *et al.*, 2004). The *Rhodococcus equi* disease poses a major threat to foal's wastage and costs the horse industry millions of dollars towards treatment (Muscatello, 2012 Kalinowski *et al.*, 2016).

Recently, a moderately effective treatment option for *Rhodococcus equi* infection has been by antibiotic administration to host. However, 30% mortality rate incidences, are still recorded, amongst the antibiotic-treated foals. Moreover, an estimated 70-80% mortality rate incidence, have been recorded in cases were *Rhodococcus equi* infection is untreated. Based on data from a series of *in vitro* susceptibility, pharmacokinetic, and retrospective studies by Giguere *et al.* (2011), rifampin was earlier considered and administered as the treatment of choice for *Rhodococcus equi* infections.

However, there are claims that *rhodococcosis* treatment, using rifampin/antibiotic, may have other concerns (Jacks *et al.*, 2003), including evidence of increasing antibiotics resistance (Cisek *et al.*, 2014). Interestingly, Okoko (2014) reported that combination therapy of macrolides and rifampin antibiotics proved to be effective against antibiotic-resistant strains of *Rhodococcus equi*, which were, isolated from

horse-breeding farms (Giguere *et al.* 2012, Burton *et al.*, 2013). Horohov *et al.*, (2011) and Gurel *et al.*, (2013) unanimously presented that streptolysin-O (as adjuvant) antibiotic therapy exerts, a more superior degree of *Rhodococcus equi* pathogen, clearance as against treatment involving rifampin antibiotics alone. A more recent study showed that a combination of erythromycin, azithromycin, or clarithromycin (macrolide) with rifampin demonstrated by far, maximum clearance, and remained a recommended therapy for foals with clinical signs of *Rhodococcus equi* infection (Giguère, 2017). After the recovery of pre-exposed foals, after treatment from *Rhodococcus equi* infection, those foals are often less likely to race as adult horses in the future (Petry *et al.*, 2017).

Although the cost of antibiotic treatment for *Rhodococcus equi* infections is on the rise, the futuristic production of a commercial vaccine may bring about the much-desired effective management of *Rhodococcus equi* infections. Through a robust preventive immunization intervention strategy (Meijer and Prescott, 2004; Giguere *et al.*, 2011; Reuss and Cohen, 2015).

5.1 Genome and Virulence Associated Proteins

A work by Letek *et al.*, (2010), described the *R. equi*, as a facultative intracellular parasite and the only animal pathogen, within the biotechnologically important actinobacterial *Rhodococcus* genus. The *R. equi* genome 103S is made up of a circular chromosome of 5.0Mb, which contains about 4,525 predicted genes, including a circular virulence plasmid of nearly 80 – 90 kb, with evidence that the *Rhodococcus equi* genome may be constituted with an overall G-C nucleotide of 68.76% (Letek *et al.*, 2010; Okoko, 2014). The backbone sequence of *R. equi*, is

significantly smaller than the genome of other environmental rhodococci (*Rhodococcus erythropolis*) (Letek *et al.*, 2008). Though the mechanisms of virulence of *R. equi* are not entirely understood now, literature evidence suggests it may depend mostly on the circular virulence plasmid (Meijer and Prescott, 2004). Interestingly, *Rhodococcus equi* strains, with virulent activity in horses, have been reported to lack these circular virulence plasmid (Takai *et al.*, 2000).

An exciting sequence and annotation analysis have shown that this virulence plasmid contains 73 coding sequences (CDSs), includes 69 open reading frames (ORFs) (Letek *et al.*, 2008). They are divided, among four regions, based on the amino acid sequence similarity and protein function (Letek *et al.*, 2008). Among the four regions, only three regions encode proteins with known virulent functions. One of the functional regions is known for plasmid replication/partitioning, conjugation, and another region described as a region of unknown functions (Meijer and Prescott, 2004; von Bargen and Hass, 2009). The third region of the plasmid as described by Merger and Prescott (2004) but reported by Giguere *et al.* (2011), as the fourth region, is a 27.5 kb domain pathogenicity island (PAI). This domain is important for virulence of *Rhodococcus equi* for foals (Takai *et al.*, 2003) and exemplified by a lower GC nucleotide content than the rest of the plasmid. There are unclear accounts about the origin of the pathogenicity island (PAI). However, a report suggested, it may have been acquired via horizontal gene transfer from a bacterial source whose origin is unknown, with events of insertion occurring in the soil, which may have enhanced the ability of the host bacterium to survive in vivo (Giguere *et al.*, 2011).

The PAI is consist of 26 putative CDSs/ORFs, including the *R. equi*- specific gene family that encodes the virulence-associated protein family (Vap family) (Takai *et al.*, 2003). There are about six full-length genes that encode VapA, - C, - D, - E, - F, - G, and H proteins (Meijer and Prescott, 2004) and other truncated (vap - pseudogenes) namely VapF, pseudo-VapE, VapX and VapI proteins. These are frame-shift mutant proteins, and non-functional (von Bargen and Hass, 2009). The Vap genes *vapA*, *vapC*, *vapD*, *vapE*, *vapF*, *vapG* and *vapH* are found throughout 19,000 base pairs Meijer and Prescott, (2004). The *vapA*, *vapC*, *vapD* are arranged closely to each other while *vapE* and *vapF* are positioned adjacent to each other. However, *vapG* and *vapH* are present as individual genes with *vapG* found in the negative strand (Takai *et al.*, 2003). The virulence of the *R. equi* organism is largely attributed to the pathogenicity island because of the VapA antigen (MW = 17 kDa) (Rofe *et al.*, 2016). Although all *vap* genes have similar homology, each plasmid has a distinct *vap* gene sequence, which may reflect different functional specificities among the bacterial isolates. Comparative analyses of *vap* plasmid DNA sequences revealed that the differences observed are a result of gene duplications, translocations, inversions, and insertion/deletion events (Willingham-Lane *et al.*, 2016). The pVAPA1037 and pVAPB1593 share approximately 75% similarity in the homology of their sequence, which is highly conserved (showing 95% DNA sequence identity). Oldfield *et al.*, (2004), described the VapA and VapB alleles at one locus, which divergently evolved in two different plasmids. This conserved sequence is referred to as the plasmid backbone and includes the domains of unknown functionality, conjugation, and plasmid replication-and-partitioning (Lake *et al.*, 2008).

5.2 *Rhodococcus equi* infections: pathogenesis and pathology

A fundamental pathogenic and virulence determinant is the *vapA* and *vapB* strain characterization in the PAI. Most *Rhodococcus equi* strains have been identified based on plasmid diversity, the organism, however, can be classified into a *vapA*-expressing strains, a *vapB*-expressing strains, and those that neither expresses *vapA* nor *vapB*, which mean they possess no virulence plasmid (Meijer and Prescott, 2004; Poolkhet *et al.*, 2010). Both the *vapA* expressing and *vapB* expressing plasmids have a conserved housekeeping backbone and a PAI that encodes different *vap* genes, which are upregulated when the organism invades the macrophage (Letek *et al.*, 2008). Although several Vap-proteins are expressed by *R. equi*, only the *vapA* is essential for bacterial survival in young horses, by aiding the pathogen rhodococcosis, to survive the destructive cellular lysosome during macrophage (Rofe *et al.*, 2016). The most common manifestation of this is chronic suppurative bronchopneumonia with extensive abscessation and associated suppurative lymphadenitis (Phumoonna *et al.*, 2006). *In vitro* experiments suggest that the infectivity of *Rhodococcus equi* is limited to the monocyte-macrophage lineage (Mosser and Hondalus, 1996); thus, alveolar macrophages could be the main targets of the organism. The early stages of the infection resemble other lung infections and may pose difficulties in early diagnosis coupled with the slow spread of the lung infection (Giguere and Prescott, 1997). The ability of *R. equi* to replicate in macrophages may result in necrosis, followed by the destruction of the lung parenchyma at the progressed stages of the disease (Hondalus, 1997). This may result from the destruction of the cell through lysosomal degranulation after the intracellular multiplication of the organism (Yager, 1987).

5.3 Mechanism of *Rhodococcus equi* infectivity

The activation of macrophages releases chemokines, cytokines, and other pro-inflammatory mediators. This results in the generation of an oxidative burst. An oxidative burst is a process that involves the release of reactive oxygen and nitrogen species (ROS and RNS), culminating in the eradication of the microorganism. Lysozyme, a cell wall degrading enzyme, is an important component of the innate immune system. The *R. equi*, is reportedly resistant to lysozyme (Hebert *et al.*, 2014).

Though the exact mechanism underlying the survival of the *R. equi* organism is unknown, inadequate acidification of the phagosome may be responsible for the inability of the phagocytes to degrade the bacteria (Zink *et al.*, 1987; Toyooka *et al.*, 2005). The pathogenicity of *R. equi* is based on its ability to multiply inside the macrophage, where it ultimately leads to lysosome degranulation and subsequently results in apoptosis (Prescott, 1991). The organism, as previously mentioned, has a complex cell envelope that enables them to survive very harsh conditions such as low pH and oxidative stress. Its ability to survive oxidative stress is perhaps the important mechanism of macrophages defense against foreign bodies (Mohanty *et al.*, 2015, Benoit *et al.*, 2000; Benoit *et al.*, 2002). The virulence of the organism is attributed to the pathogenicity island VapA (a 15 -17 kDa antigen) thought to mediate intracellular growth of organism in the macrophages and disease development via the arrest of phagosome maturation. Research has shown that when *vapA* attenuates the virulence of *R. equi*, to almost the same degree as that of eliminating of the entire *vap* locus (Jain *et al.*, 2003, Okoko, 2014).

Through the analysis of a *vapA* deletion mutant strain, Jain and colleagues demonstrated VapA to be responsible for the intracellular growth and virulence of *R. equi*. The loss of *vapA* resulted in a fully attenuated mutant, no longer able to replicate in macrophages or establish disease in the *in vivo* chronic disease mouse model (Jain *et al.*, 2003). The role of VapA in macrophage infection is not fully understood; however, Bargen *et al.*, (2009) reported, that VapA appears to interfere with normal phagosome maturation (Von-Bargen *et al.*, 2009).

Similarly, a study by Rofe *et al.*, (2016) examine the effect of VapA on lysosome morphology, by inducing uptake of incubated VapA cell via fluid endocytosis over 48 hours. The finding suggested that GST-myc-VapA triggered an increase in lysosome size (swelling). This swelling factor was peculiar to VapA and no other Vap-proteins. The VapA protein shifts intracellular transcription factor (TFEB) to the nucleus, hence more lysosomes get produced. Summarily, the biogenesis of lysosome in macrophage is regulated, therefore when it is disrupted by VapA, the cells begin to produce more lysosome, a phenomenon that is defective for fighting *R. equi* infection.

5.4 *Rhodococcus equi* Virulent Factors

Other factors are known to significantly contribute to the virulence of *R. equi* VapA. These factors include chromosomal-encoded factors, plasmid-encoded factors, and environmental factors and have been concisely, presented below.

5.4.1 Chromosome-encoded Factors

There are two key enzymes of the aromatic amino acid biosynthesis pathway that the *Rhodococcus equi* chromosome encodes. These are chorismite mutases (REQ23860) and anthranilate synthase (REQ23850). According to Letek *et al.*, (2010), the encoding genes are co-induced with the *vap* genes. *R. equi* depends on the *de novo* supply of aromatic acids. This is because the amino acids tyrosine, phenylalanine, and tryptophan occur at limited concentrations in the hypoxic macrophage intracellular environment. Thus, full proliferation and intracellular fitness of the bacterium could be compromised resulting to the independent mutations of the REQ23860 and REQ23850 (Letek *et al.*, 2010)

There are other chromosome-encoded factors to be considered in the exploitation and modification of host lipids. Hondalus (1997) considers the exo-enzymes cholesterol oxidase and phospholipase C as vital virulence factors. The 54 kDa secreted protein; cholesterol oxidase catalyzes the oxidation of cholesterol to 4-cholesten-3-one. It is also implicated in the destabilization of the host cell membrane and may be responsible for tissue damage associated with *R. equi* infections (Linder and Bernheimer 1997). The *R. equi* has four phospholipase C proteins encoded in its genome (Vera-Cabrera *et al.*, 2013). Phospholipase C modifies the activity of cholesterol oxidase and provides access to the target, which may confer the membranolytic activity of *R. equi* through the combined action of cholesterol oxidase and phospholipase. Cholesterol oxidase and phospholipase are called 'equi factors' (Prescott, 1991). The cholesterol oxidase enzyme targets the sterol layer of the lysosomal or cellular membranes of the macrophage. This might contribute to the macrophage degeneration observed *in vitro* with *R. equi*

infectivity (Hondulas, 1997). Van der Geize *et al.* (2011) demonstrated that *R. equi* catabolises cholesterol to products that serve as substrates for other vital biochemical pathways (also called the steroid catabolic pathway). Van der Geize *et al.*, (2011) similarly showed that *R. equi* displayed attenuated phenotypes and limited propagation in an *in vitro* macrophage infection assay, when alterations occur in some key enzymes of the cholesterol biochemical pathway the same pathway. This steroid catabolic pathway is utilized by *R. equi* to enable macrophage membrane sterols. *R. equi* may use this pathway to utilise macrophages membrane sterols possibly due to extra demands for carbon in the hostile environment.

Rhodococcus equi contains other secreted lipases in addition to cholesterol oxidase and phospholipase, which can deteriorate host cells for fatty acid oxidation (Alvarez *et al.*, 2013; Villalba *et al.*, 2013). Another virulence factor is likely nitrate reductase encoded by *narG*. *R. equi* is an aerobic organism and can be facultative since it is capable of anaerobic respiration through the denitrification – reducing nitrates to nitrites (Letek *et al.*, 2010). Consequently, it may be capable of utilizing nitrate in the hypoxic macrophage vacuolar compartment as a terminal electron acceptor in its respiratory chain. Giguere *et al.*, (2011), in a mouse challenge study, observed alterations in *narG* were also significantly attenuated.

5.4.2 Environmental Factors

Temperature and pH are major environmental factors that determine the virulence of *R. equi*. These determinants regulate the transcriptional expression of the *vap* PAI genes with a tendency to upregulate at 37°C, pH 6.5, and downregulate at 30°C, pH 8.0 (Vázquez-Boland *et al.*, 2013). Meijer and Prescott (2004), state that there are environmental stresses that monitor gene expression within the pathogenicity island. These stresses are oxidative stress, magnesium, and iron stress. *R. equi* VapA virulence is also directly related to the ability of an organism to obtain ferric iron from the host (Letek *et al.*, 2010).

5.4.3 Plasmid-encoded Factors

The virulence of *R. equi* is also attributable to VapA. The transcription of VapA is influenced by the regulation of the proteins encoded by the genes *virR* and *virS*, which are part of the *virR* operon (Byrne *et al.*, 2007). As a result, a change in either *virR* or *virS* may minimize the expression of *vapA* and reduce the malignance of *R. equi* infectivity (Giguere *et al.*, 2011; MacArthur *et al.*, 2011). *IcgA* is another gene that is an integral element of the *virR* operon (Wang *et al.*, 2014). Wang *et al.* (2014) observed that the elimination of *icgA* did not influence the transcription of *vapA*. However, it increases the growth of phenotype and causes a significant decrease in the viability of the organism in macrophages.

5.5 Immunological aspects of *Rhodococcus equi* infections

Macrophages are immune cells function in driving primary responses to pathogens, maintenance of tissue homeostasis, including the coordination of adaptive immune responses, tissue repair, inflammation, and resolution of inflammatory stimuli (Hirayama, Iida and Nakase, 2017; Kumar 2019). Based on distinct functions and physiological roles, macrophages are activated in response to different innate or adaptive immune signals. According to Willingham-Lane and co-workers, the ability of *R. equi* to multiply inside the macrophage is majorly responsible for its virulence (Willingham-Lane *et al.*, 2018). Macrophages may be the most convenient hosts for the organism to thrive. This is because *in vitro* experiments have shown that the virulence is constrained to cells that show the complement receptor Mac-1 (CD11b and CD18). These are surface integrin receptors only found to induce macrophage immune responses (Zhou *et al.*, 2012). The intracellular growth of the organism is considered toxic because it impedes the production of the phagolysosome. This inhibition is due to either intracellular growth or as a survival mechanism (Hondalus and Mosser 1994). A phagosome is formed when a phagocyte engulfs a non-self-substance, which includes a bacterium through receptor-mediated endocytosis. The formed phagosome continues to mature into a phagolysosome through an early and a late phagosome stage by connecting with the lysosome (Von Bargen and Haas 2009). The interaction of low pH (4.0 - 5.0), hydrolytic enzymes, the formation of reactive oxygen and nitrogen species result in the elimination of these pathogens (Von Bargen and Haas 2009). Fernandez-Mora *et al.*, (2005) claim that most (*ca.* 90%) of the vacuoles with *R. equi* are arrested at a late endosome stage *en-route* to the lysosome, therefore, preventing

the phagosome from maturing (Figure 5.1). On the other hand, Toyooka *et al.*, (2005) argued that when virulent *R. equi* are incorporated in the macrophages, maturation of the phagosomes, continues initially just like other phagosomes containing harmless microorganisms, but they are tolerant to the bactericidal substances present in the phagolysosomal environment. Toyook *et al.*, (2005) further explain that *R. equi* acts by releasing substances that reduce the acidification of the phagolysosome. This, by implication, prevents the optimum pH required for the activity of phagolysosome. Consequently, cell necrosis occurs because of the continuous multiplication of the organism in the compartment (Figure 5.2).

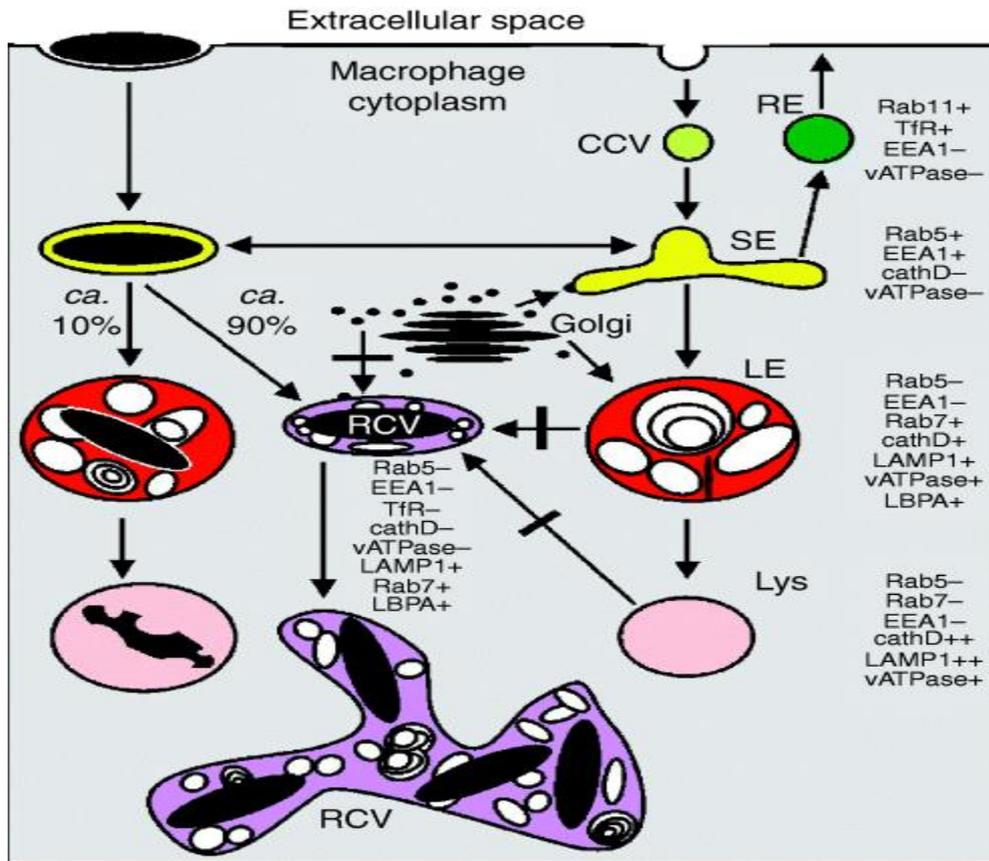


Figure 5.1: A model for *R. equi* infectivity in macrophages.

Clathrin-coated vesicles (CCV) feed extracellular liquid into the endocytic system. Phagosomes are produced, at the plasma membrane and normally interact continuously with early sorting endosomes (SE), late endosomes (LE) and lysosomes (Lys), but not with early recycling endosomes (RE) through which plasma membrane receptors can be recycled from the SE back to the surface. Some 10% of the ingested *R. equi* have moved along the degradative pathway to a (phago) lysosomal compartment and possibly degraded, whereas most (about 90%) is directed to an unusual compartment where the composition is between the SE and the LE. The Golgi compartment communicates via vesicular trafficking with the SE/LE system but not with RCVs. Typical compartmental markers are indicated (Modified from Fernandez-Mora *et al.*, 2005).

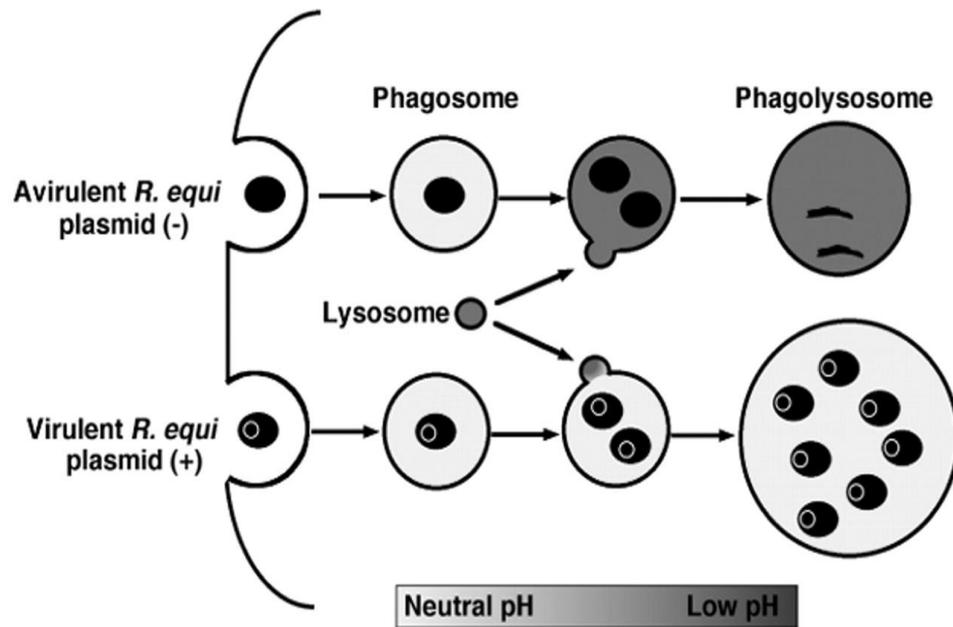


Figure 5.2: Alternative model for *R. equi* infection in macrophages.

After engulfing *R. equi*, phagosomes containing virulent or avirulent *R. equi* integrate with lysosomes. After that, avirulent *R. equi* inadequate of a plasmid cannot thrive in an acidic compartment in the phagolysosomes. Virulent *R. equi* harboring an 85 kb plasmid subdue the acidification. Thus, they thrive and proliferate in the phagolysosomes (Adapted from Toyooka *et al.*, 2005)

A lysozyme that degrades a cell wall is a fundamental component of the innate immune system. *R. equi* is marginally resistant to lysozyme (Hébert *et al.*, 2014). This marginal resistance is expected, as the enzyme does not readily access the peptidoglycan substrate located beneath the mycolic acid membrane.

Nitric oxide (NO) is an important reactive nitrogen species. Nitric oxide is synthesized from arginine in a reaction that is catalysed by inducible nitric oxide synthase (iNOS). Interferon-gamma (IFN- γ) is an initial signal for the transcription of iNOS. However, there are secondary signals such as tumour necrosis factor- α (Darrah *et al.*, 2000). The transcription of IFN- γ mRNA usually decreases when *R. equi* strains gain access into the alveolar macrophage of susceptible foals (Hines *et al.*, 2003). This indicates that *R. equi* may produce factors that inhibit the

production of IFN- γ . Interferon-gamma upregulates phagolysosomal fusion (a component of the Th1 response) in addition to stimulating macrophages. Cytokine, therefore, plays a crucial role in the phagocytotic susceptibility of a host to *R. equi* (Kanaly *et al.*, 1995).

Patton *et al.*, (2005), through experiments in mouse models involving neutralizing antibodies and adoptive transfer, showed that a decrease in *R. equi* led to the formation of Th2 responses characterized by the formation of interleukin-4 as against IFN- γ . The mice developed peculiar pulmonary lesions, which may be attributed to the scarcity of IFN- γ . Macrophages could be stirred to produce other reactive species like hydrogen peroxide. However, *R. equi* is strongly resistant to hydrogen peroxide stress (Meijer and Prescott, 2004). Consequently, hydrogen peroxide activates the induction of *vap* genes specifically *vapA* and *vapG* of the organism (Benoit *et al.*, 2002; Meijer and Prescott 2004). The exact mechanism of a protective immune response in adult horses remains unknown. Notwithstanding, there are shreds of evidence to suggest that pulmonary clearance of *R. equi* is due to the response of Th1, where virulent strains are eradicated, with the increase in CD4⁺ and CD8⁺ T-lymphocytes (Hines *et al.*, 2001). Studies have also shown that T-lymphocytes, obtained from the lungs of adult horses after immune clearance, multiply in culture as a response to the antigens of *R. equi*, including IFN- γ apart from interleukin-4 (Giguère *et al.*, 2011).

Antibody-mediated responses also have an important part in protective immunity. In foals, the minimum levels of colostrum-derived maternal antibody are observed by eight weeks. Following this, they begin to develop antibodies against the organism. According to Prescott (1991), foals with diminished levels of maternal

antibodies are more susceptible to *R. equi* infection. Immunoglobulin G (IgG) plays a significant part in opsonization and phagocytosis of virulent *R. equi*, while mucosal IgA (gotten from infected foals) reacts strongly to specific B-cell epitope on *vapA* (Muscatello 2012).

The vulnerability of foals to *R. equi* is also enhanced by the absence of IgA from the nasal mucosa (which occurs during the first four weeks of life) and the diminished circulating antibody titers (which occurs between one and three months) (Muscatello 2012). Adult horses are typically resistant to *R. equi* infection. This is because most of them have significant antibody titer to the organism – hyperimmune. They are highlighting, the reverse correlation between antibody, disease prevalence, and severity (Walscheid *et al.*, 2014). Mosser and Hondalus (1996), showed that passive immunization of hyperimmune sera to foals exposed to *R. equi* protected them from developing the *R. equi* infection. In addition to humoral immunity, cell-mediated immunity can be transferred through mare's colostrum (Porto *et al.*, 2014).

5.6 Aim of Study

The ultimate arrest of phagosome maturation, because of VapA intracellular mediated disruption of lysosome biogenesis in macrophages during disease development, alters the physiological mechanism through which lysosome destroys *R. equi* or any other bacterium. We have already established that *R. equi* virulence is mostly due to a virulent strain, retained at the surface of the cell that contains a pathogenicity island that encodes VapA. Interestingly there is a hypothesis that when other non- virulent Vap-proteins self-associate with VapA, they could attenuate its virulence factor, which reduces the concerns of *R. equi* infectivity and its ability to destroy lysosome activities during macrophage. The aim of the study was to investigate the hypothesis of VapA's association or binding interaction with VapG, VapE, and VapH, through a heterologous expression of cloned Vap-protein that have Histidine-tag at both N-terminal (pET-28a) and C-terminal (pET-23a) respectively. The objective is to investigate with the use of biophysical or protein-protein interaction techniques and to determine whether non-virulent Vap-protein can associate via binding with Vap A, with the help of Histidine fusion protein at both N-terminal (pET-28a) and C-terminal (pET-23a), of the non-virulent Vap protein.

5.7 Results

5.7.1 Cloning, Purification, and Characterization of Vap proteins

5.7.1.1 Introduction

This result section attempts to characterize the *R. equi* vap-plasmid clone *vapA*, *vapD*, *vapE*, *vapG*, and *vapH*, towards actualizing the synthesis of soluble pure heterologous recombinant protein products. This was achieved through a systematic protein biosynthesis method, included the transformation of the original vap-plasmids into the desired *Escherichia coli* strain C43 (DE3). Vap proteins were overexpressed and purified using Immobilised Metal Affinity Chromatography (IMAC) separation methods. This was followed by cleaning up the heterologous recombinant fusions proteins to ensure they are optimally viable devoid of genomic DNAs and RNAs.

Once this was satisfied, Vap-proteins were applied towards functional assay using Surface Plasmon Resonance and to confirm the outputs of the fitted data, and thermal shift assay was also chosen as an orthogonal technique. These methods were applied, to find answers on affinity and binding interactions, binding kinetics, and a measure of entropy or enthalpy for Vap protein interaction. It is instructive to note that previous research (Okoko, 2014) had identified some of the Vap-protein molecular structures; however, what remains unknown is the mechanism through which other Vaps may enhance or attenuate virulence when bound or interacted with VapA has not been reported.

The principle and rationale for applying these techniques to this work have been captured in section 1.2.1, 1.2.1, and 1.2.3.1 respectively. While the methodology

of these techniques has been concisely described in section 2.9.11, 2.10.1 and 2.11.2 respectively. The *Vap*-plasmid clones that were retransformed into *Escherichia coli* C43 (DE3) and used in this work were found to produce high yield soluble recombinant proteins.

5.8 Amplification of *R. equi* Vap Genes and Cloning

The source of the *R. equi vap*-plasmid clones used in this work was received as a kind gift from Dr. Lynn Dover of Northumbria University UK. This *R. equi vap*-plasmid includes *vap A*, *D*, *E*, *G*, and *H*, which had been successfully cloned into pET23a and pET28a (C and N terminally hexahistidine – tagged) plasmid vectors respectively using *Escherichia coli* Top10 strain. These *vap*-plasmids have been cloned and verified to have high copy number plasmids with a strong T7- promoter, as they were received as a gift. The presence of the pET23a (C-terminal hexahistidine fusion) and pET28a (N-terminal hexahistidine fusion) were to facilitate a metal chelated affinity chromatographic technique, for the purification of recombinant Vaps-fusion protein products.

5.9 Recombinant *R. equi* VAP Protein Expression

Protein overexpression using bacteria cells, as highlighted in chapter 1, remains one of the most viable means of obtaining recombinant proteins for biological and functional assays. The *R. equi vap* recombinant plasmids, the *vapA*-pET23a, *vapE*-pET23a, *vapG*-pET23a and *vapG*-pET28a, *vapH*-pET23a and *vapH*-pET28a were all used to transform *Escherichia coli* C43 (DE3) strain. Heterologous protein synthesis in *Escherichia coli* is found to depend on a few factors such as temperature,

appropriate broth, right IPTG concentration, with temperature having a profound effect on protein stability and folding (Weickert *et al.*, 1996). The appropriate *R. equi* Vap-protein overexpression conditions applied are described in section 2.5.2. Also, SDS-PAGE electrophoresis acquired images of prepared from clarifying recombinant Vap protein lysates showed that the protein synthesis of selected Vap construct was induced under the condition described in section 2.5.2. Large-scale production of recombinant Vap- proteins in *E. coli* was adopted as described in section 2.5.3, to produce high yields of recombinant proteins towards further functional and structural studies.

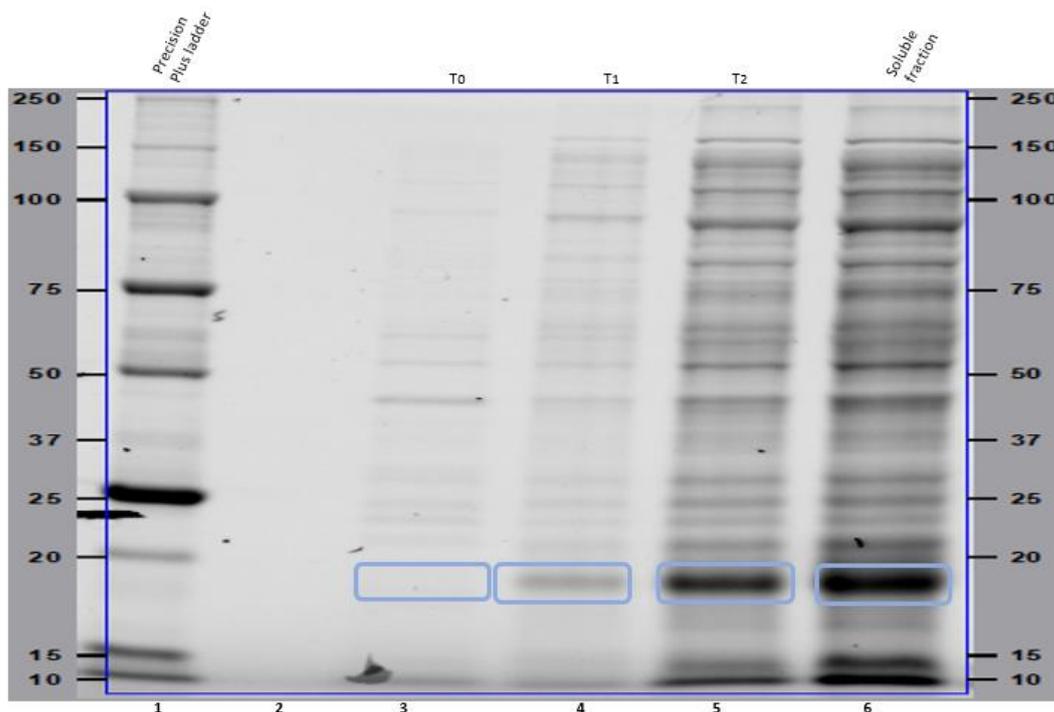


Figure 5.3 Protein Overexpression Analysis of C-terminal-HIS tagged *R. equi* pVapA Fusion Protein expression in C43 (DE3) *E. coli* strain at 37°C in 2YT media. Lane1: Biorad Precision plus MW marker, Lane 3 – 6, C-terminal-HIS tagged *R. equi* recombinant *pET23a- vapA* Fusion Protein expression profile showing post IPTG (1mM) induction. T0 (L3), T1 (L4), T2 (L5), the soluble fraction (L6). All samples were resolved by SDS-PAGE methods at 150 Volts.

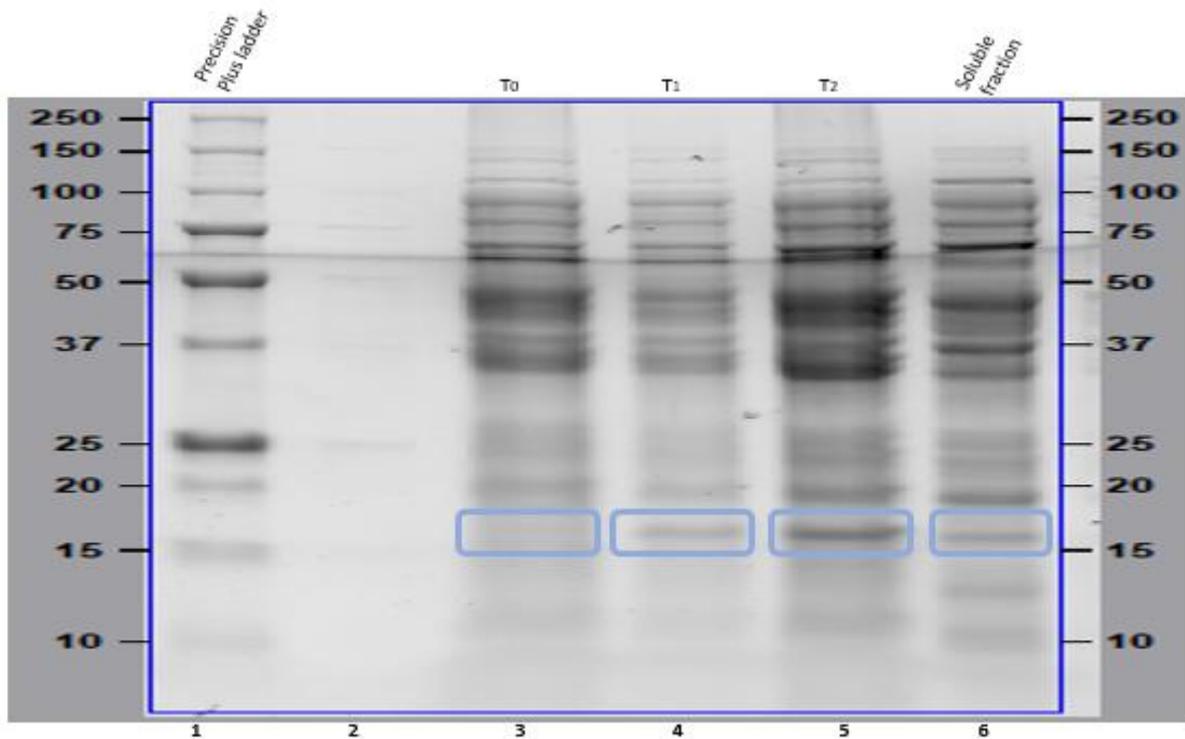


Figure 5.4 Protein Overexpression Analysis of C-terminal-HIS tagged *R. equi pVapE* Fusion Protein expression in C43 (DE3) *E. coli* strain at 37°C in 2YT media. Lane1: Precision plus MW marker, Lane 3 – 6, C-terminal-HIS tagged *R. equi* recombinant *pET23a-pVapE* fusion protein expression profile showing post-IPTG (1mM) induction. T0 (L3), T1 (L4), T2 (L5), the soluble fraction (L6). All samples were resolved by SDS-PAGE methods at 150 Volts.

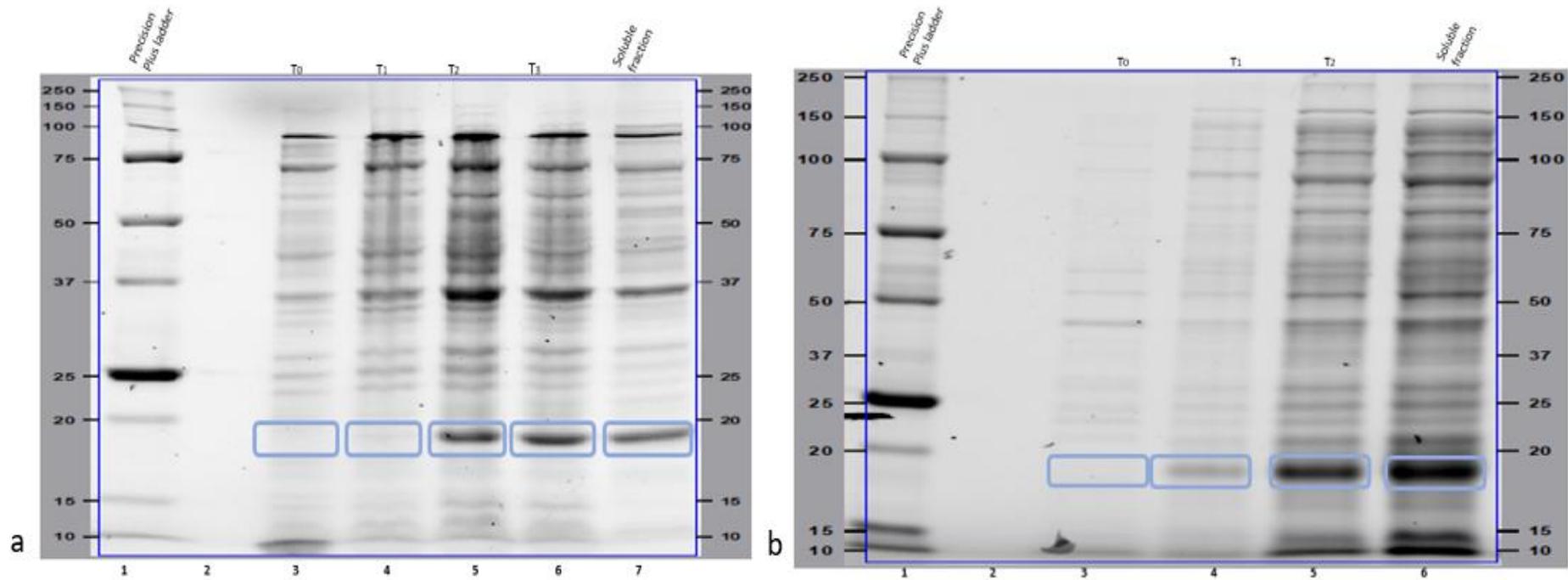


Figure 5.5 Protein Overexpression Analysis of N-terminal-HIS and C-terminal tagged *R. equi* pVapG Fusion Protein expression in C43(DE3) *E. coli* strain at 37°C in 2YT media.

(A) Lane1: Biorad Precision plus MW marker, Lane 3 – 6, N-terminal-HIS tagged *R. equi* recombinant *pET23a-pVapG* fusion protein expression profile showing post-IPTG (1mM) induction. T0 (L3), T1 (L4), T2 (L5), T3(L3) soluble fraction (L7). **(B).** Lane1: Biorad Precision plus MW marker, Lane 3 – 7, C-terminal-HIS tagged *R. equi* recombinant *pET28a- pVapG* Fusion Protein expression profile showing post-IPTG (1mM) induction. T0 (L3), T1 (L4), T2 (L5) (whole expression lysate) the soluble fraction (L6). All samples were induced at optical density of 0.7 nm and resolved by SDS-PAGE methods at 150 Volts.

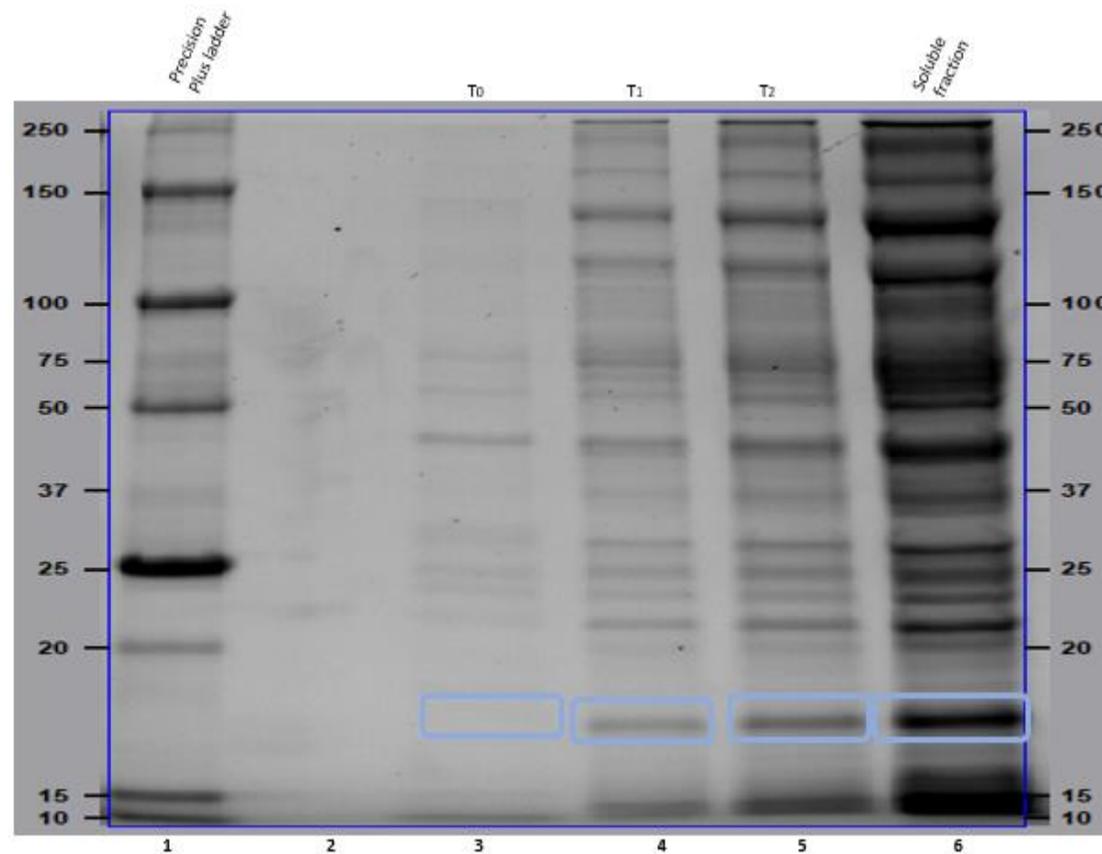


Figure 5.6 Protein Overexpression Analysis of N-terminal-HIS tagged *R. equi* pVapH Fusion Protein expression in C43(DE3) *E. coli* strain at 37°C in 2YT media. Lane1: Bio-Rad Precision plus MW marker, Lane 3 – 6, N-terminal-HIS tagged *R. equi* recombinant *pET28a-pVapH* fusion protein expression profile showing post-IPTG (1mM) induction. T0 (L3), T1 (L4), T2 (L5) (whole expression lysate), the soluble fraction (L6). All samples were induced at optical density of 0.7 nm and resolved by SDS-PAGE methods at 150 Volt.

Before the induction of expression, *R. equi* Vap-protein cells were grown to mid-log phase at 37°C. Overexpression samples were taken for analysis at one-hour post 1mM IPTG induction at 37°C (T1), two hours post IPTG induction (T2), and overnight post IPTG induction (to/n), followed by analysis using the SDS-PAGE. After initial expression analysis, it was qualitatively assessed, that two hours of expression gave the best ratio of soluble to insoluble Vap-His tagged recombinant fusion protein. Representative SDS-PAGE gel for pET23a-*pVapA*, pET23a-*pVapE*, pET28a-*pVapG*, pET23a-*pVapG* and pET28a-*pVapH* protein synthesis and solubility profile as shown, in figure 5.3 - 5.6. A review of the SDS-PAGE analysis of clarified Vap fusion protein overexpression profile, through the analysis of its lysates, showed that synthesis of these recombinant proteins of interest was induced, under these conditions in 2 hours post IPTG induction. With pET23a-*pVapA*, pET23a-*pVapG*, and pET28a-*pVapH* showing a high degree of soluble recovery of recombinant proteins and pET28a-*pVapG* and pET23a-*pVapE* showing a moderate recovery of recombinant soluble protein. Furthermore, the fusion protein band markers migration between 15 and 20 MW kDa marker, with an expected protein size of about 21kDa for pET23a-*pVapA*, 17kDa for pET23a-*pVapG*, 16kDa for pET28a-*pVapG*, and 18kDa for pET28a-*pVapH* and pET23a-*pVapE* respectively.

5.10 Recombinant *R. equi* VAP Protein Purification

To extract purified soluble recombinant proteins required for experimental functional assays. The process of recovering and clarifying all recombinant soluble *R. equi* Vap-proteins under study, by applying the mechanical lysate lysis technique of sonication (as described in section 2.5.6) was followed. Thereafter the processes of recombinant soluble *R. equi* Vap-protein purification via IMAC column affinity chromatography were applied (as described in section 2.6.2), which includes equilibration of the system, sample injection and wash. The purified *R. equi* Vap-protein fractions were eluted in a linear gradient of 1 x NP, Imidazole concentration terminating at 1 x NP 500mM Imidazole as appropriate wash buffer. The *R. equi* Vap-protein elution fractions for all Vap protein samples were monitored at 280nm absorbance and verified by SDS-PAGE analysis to identify the homogeneity of *R. equi* Vap-protein eluents that contain purified recombinant proteins. The *R. equi* Vap-protein purification peak fractions were also resolved using the Bradford method, pooled based on the lysate fraction that turned blue as described in section 2.6.3 and concentrated via Amicon ultra-centrifuge (as described in section 2.7.3) which in this case is N-terminally and C-terminally tagged *R. equi* Vap proteins respectively? Representative chromatographic images of the protein purification profile, including confirmatory SDS-PAGE gel of eluents from the sample (i), C-terminal-HIS tagged pVapA, (ii) C-terminal-HIS tagged pVapG, (iii) N-terminal-HIS tagged pVapG, (iv) N-terminal-HIS tagged pVapH, and (v) C-terminal-HIS tagged pVapE are shown from figure 5.7 – 5.10 below.

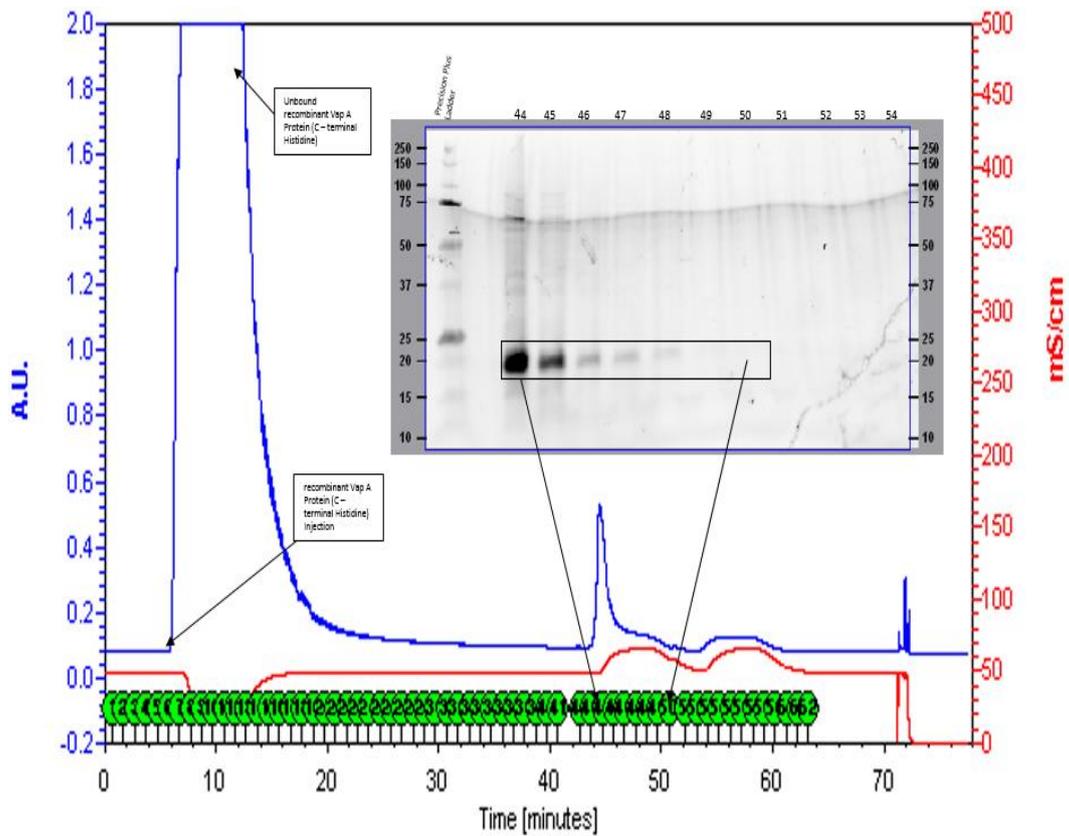


Figure 5.7 Purification of Recombinant C-terminal Vap A-HIS tagged Fusion Protein

(A) Chromatogram showing a single-step BioLogic C-terminal Vap A His-tagged protein purification using Affinity Chromatography. Purification profile showing unbound protein, normalized baseline, and the bound protein. Recombinant proteins were purified at 280nm and eluted with 1 x NP 500mM Imidazole (B) Showing purified C-terminal Vap A-HIS-tagged protein Eluents, (L1) Precision plus marker, Lane 3 – 11, C-terminal Vap A-HIS tagged protein eluent samples 44-48. Resolved by SDS-PAGE methods at 150 Volts.

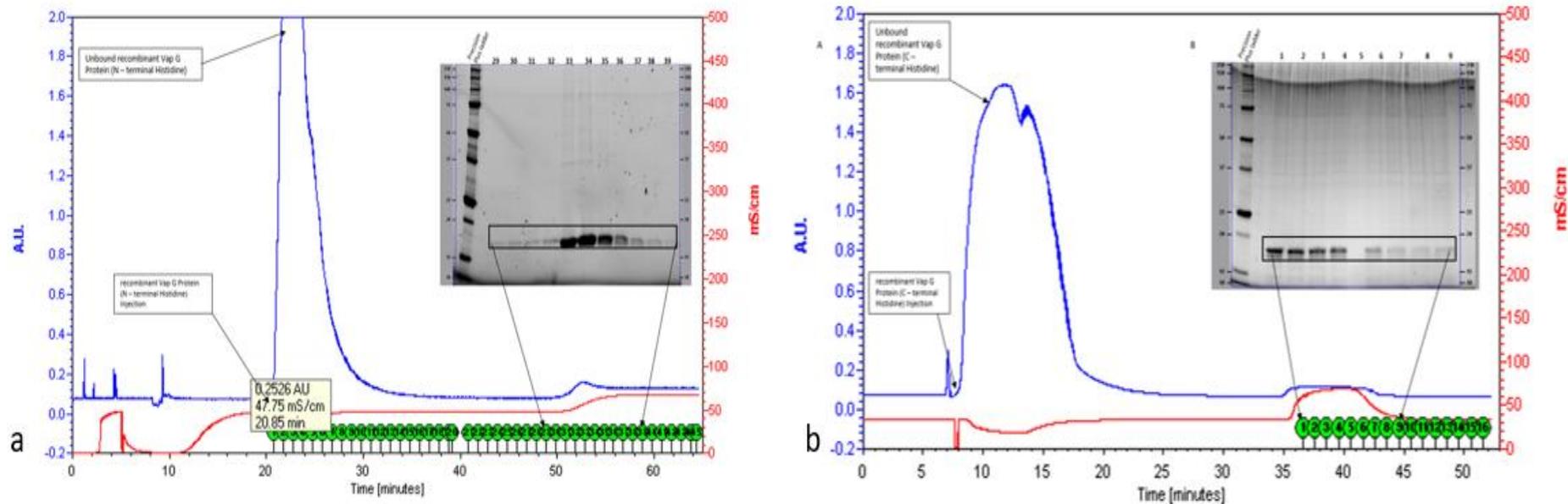


Figure 5.8 Purification of Recombinant N-terminal and C-terminal *Vap G*-HIS tagged Fusion Protein

(A) Chromatogram showing a single-step BioLogic N-terminal *Vap G* HIS-tagged protein purification using Affinity Chromatography. Purification profile showing unbound protein, normalized baseline, and the bound protein. Recombinant proteins were purified at 280nm and eluted with 1 x NP 500mM Imidazole (SDS-PAGE image) Showing purified N-terminal *Vap G*-HIS-tagged protein Eluents, (L1) Precision plus marker, Lane 3 – 11, N-terminal *Vap G*-HIS-tagged protein eluent samples 29 - 39. **(B)** Chromatogram showing a single-step BioLogic C-terminal *Vap G* HIS-tagged protein purification using Affinity Chromatography. Purification profile showing unbound protein, normalized baseline and the bound protein. Recombinant proteins were purified at 280nm and eluted with 1 x NP 500mM Imidazole (SDS-PAGE image) Showing purified C-terminal *Vap G*-HIS tagged protein Eluents, (L1) Precision plus marker, Lane 3 – 11, C-terminal *Vap G*-HIS tagged protein eluent samples 1-9. Resolved by SDS-PAGE methods at 150 Volts.

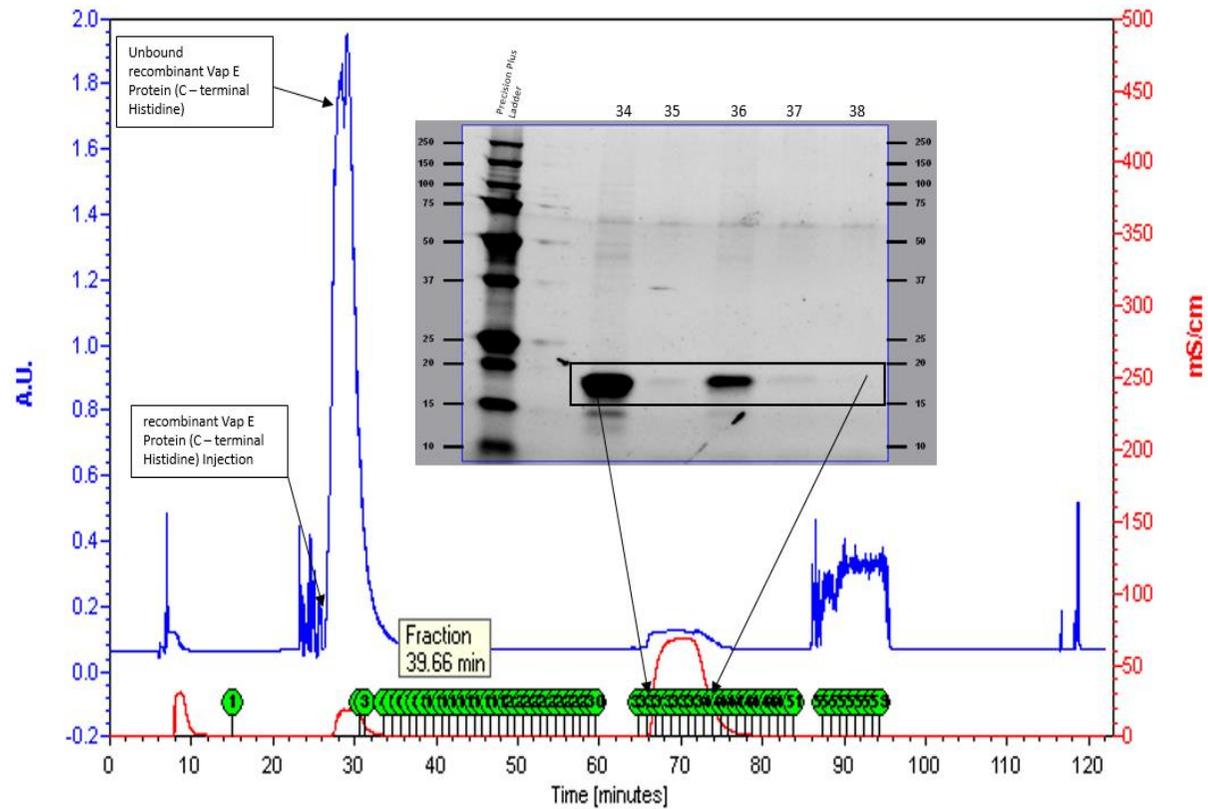


Figure 5.9 Purification of Recombinant C-terminal *vap* E-HIS tagged Fusion Protein

- (A) Chromatogram showing a single-step BioLogic C-terminal Vap E HIS-tagged protein purification using Affinity Chromatography. Purification profile showing unbound protein, normalized baseline, and the bound protein. Recombinant proteins were purified at 280nm and eluted with 1 x NP 500mM Imidazole (B) Showing purified C-terminal *vap* E-HIS tagged protein Eluents, (L1) Precision plus marker, Lane 3 – 11, C-terminal *Vap* E-HIS tagged protein eluent samples 34-37. Resolved by SDS-PAGE methods at 150 Volts.

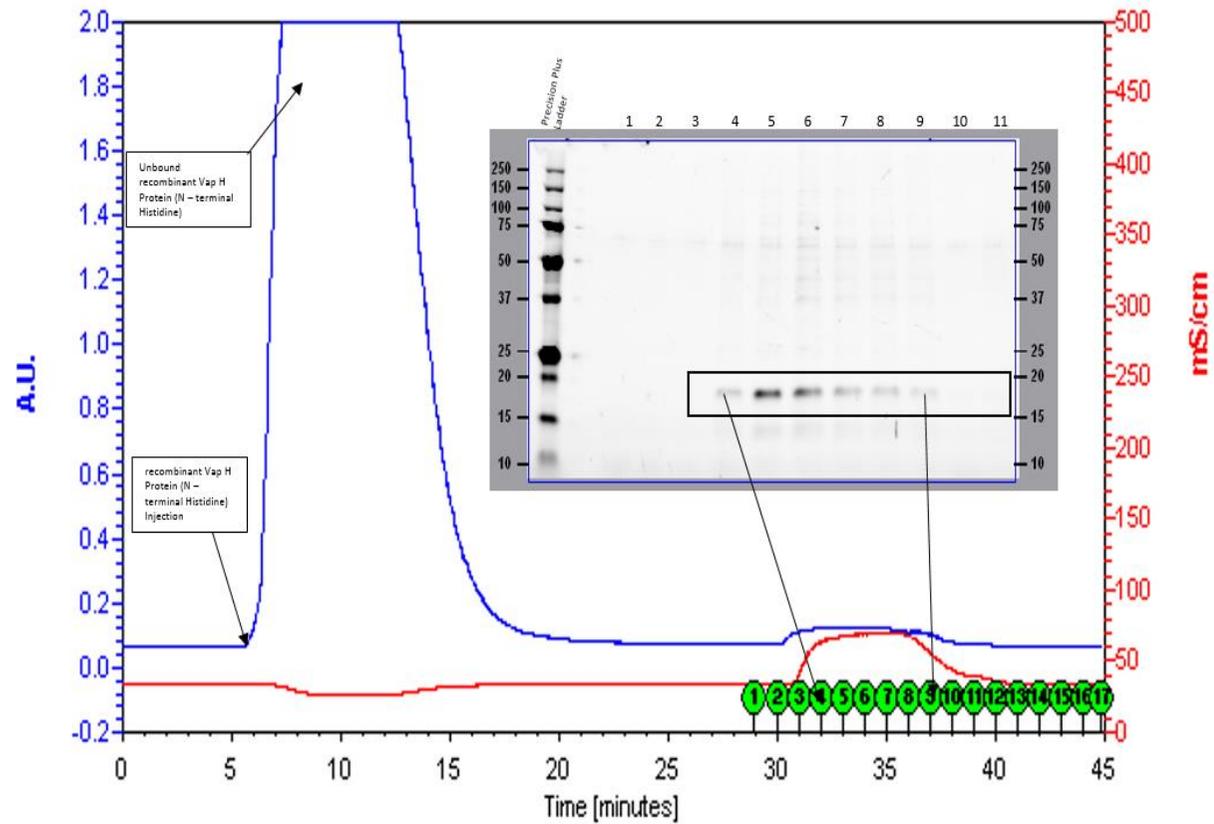


Figure 5.10 Purification of Recombinant N-terminal *vap* H-HIS tagged Fusion Protein

(A) Chromatogram showing a single-step BioLogic N-terminal Vap H HIS-tagged protein purification using Affinity Chromatography. Purification profile showing unbound protein, normalized baseline, and the bound protein. Recombinant proteins were purified at 280nm and eluted with 1 x NP 500mM Imidazole (B) Showing purified N-terminal *vap* H-HIS tagged protein Eluents, (L1) Precision plus marker, Lane 3 – 11, N-terminal *Vap* H-HIS tagged protein eluent samples 4 - 9. Resolved by SDS-PAGE methods at 150 Volts.

After IMAC column affinity chromatography purification of recombinant vap-protein, a review of the SDS-PAGE analysis of purified Vap fusion eluents, showed that the two-step IMAC chromatography had successfully separated pET23a- *pVapA*, pET23a- *pVapG* and pET28a- *pVapH*, pET28a- *pVapG* and pET23a- *pVapE* fusion protein from the *E. coli* genomic nucleic acid. Furthermore, the purification chromatogram showed the pattern of purification consistent with expectation. Also, the molecular weight bands on the SDS-PAGE analysis presented in figure 5.7 – 5.10 showed migrations between 15 and 20 MW kDa marker, with the expected protein size and concentration of all Vap proteins summarised and presented in table 5.1 below.

Table 5.1: Summary of experimental size and concentration per μL of C43DE3 expressed recombinant Vap - proteins.

<i>R. equi</i> Vap proteins	Vap-fusion protein experimental Molecular Weight	Vap-fusion protein Concentration (mg/ μL)
VapA C-terminal	21Kda	1.91
VapG C-terminal	17kDA	0.76
VapG N-terminal	16kDA	1.66
VapH N-terminal	18kDA	0.63
VapE C-terminal	18kDA	0.92

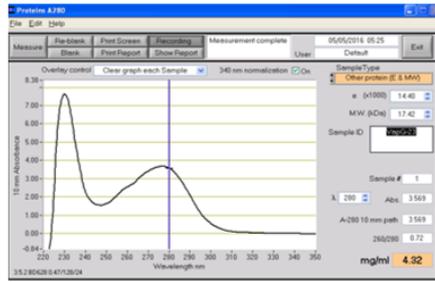
5.11 Isolation of Pure Recombinant Vap - His Fusion Protein

After a successful protein overexpression, care was taken to ensure the absence of genomic DNAs and RNAs in the sample product. Ideally, cell lysis remains the first step in cell fractionation (Walker, 2009; Thermofisher.com, 2020). The mechanical lysis of protein expression lysates, being the final stage of physical disruption of the overexpression cell lysate, was lysed in this work via sonication to break down cells,

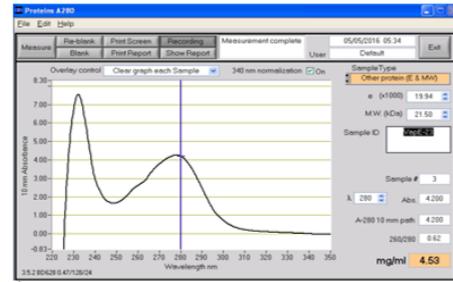
hence leading to the release of nucleic acids. The presence of these cellular components reduces the supreme purity of the recombinant protein products (Hebron *et al.*, 2009; Tan and Yiap, 2009). Although all experimental protein expression lysates were lysed by sonication, no step was taken to guarantee the elimination of genomic DNAs from the protein expression sample. The disadvantage of this omission could mean that over expressed recombinant protein contaminant that interfere during binding interaction assay. A series of purity and concentration assessment of IMAC purified Vap proteins at 280nm absorbance under the ultraviolet spectrum, showed that might be tiny traces of genomic DNAs or RNAs, this is because, the peak absorbance was with 270nm and 280nm spectrum, rather than at 280nm as shown in figure 5.10.

A subsequent enzymatic digestion by adding DNase and RNase into a purified recombinant Vap protein sample and treating them using the methods described section 2.8.0. This clarified that all Vap protein showed peak absorbance to UV light at 280nm, as shown in figure 5.11 and consistent with expectation protein absorbance ultra-violet spectrum for protein (Anthis and Clore, 2013).

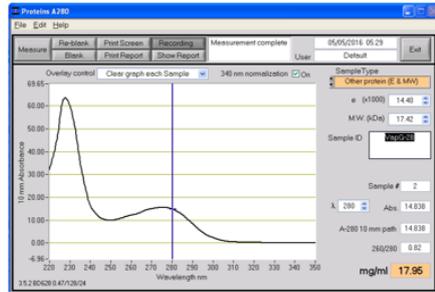
The application of contaminated or impure proteins may affect the robustness of functional assays (Wingfield, 2017). Because we aimed to do some functional assays in this work, we attempted to ascertain the purity and protein integrity of Vap protein and the recombinant Vap-His fusion proteins that have been synthesized in this work.



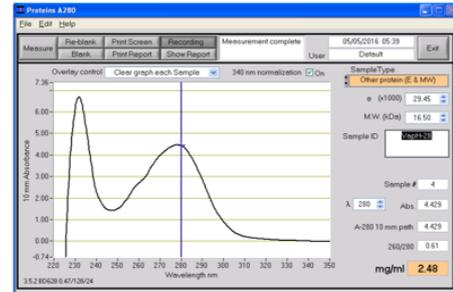
a



b



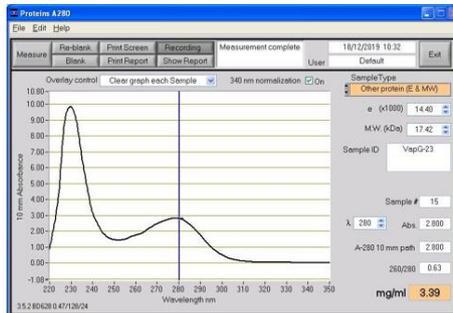
c



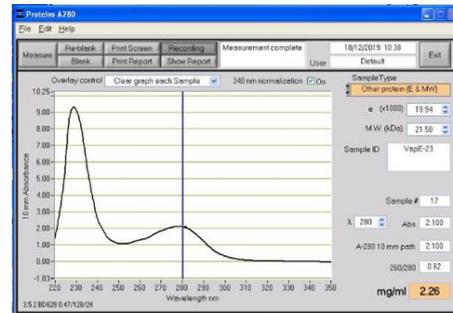
d

Figure 5.11: Crude and Non-dialysed Pure Vap-His fusion proteins

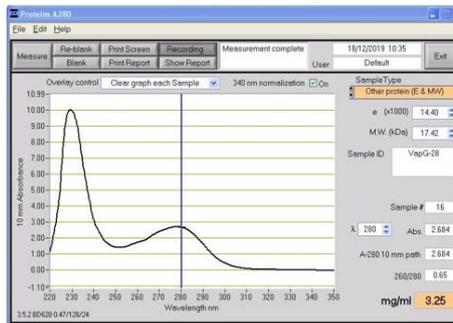
Showing clarified crude protein lysate sample of (a) *pET23a-VapG* (b) *pET23a-VapE* (c) *pET28a-VapG* (d) *pET28a-VapH* showing to absorb UV-spectrum between 270 and 280nm. Protein concentrations were measured using a Nanodrop Spectrophotometer ND-1000 USA.



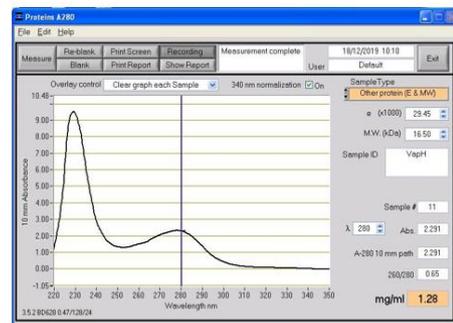
a



b



c



d

Figure 5.12: DNase and RNase Digested and Dialysed Vap-His fusion proteins

Showing clarified crude protein lysate sample of (a) *pET23a-VapG* (b) *pET23a-VapE* (c) *pET28a-VapG* (d) *pET28a-VapH* showing to absorb UV-spectrum between 270 and 280nm. Protein concentrations were measured using a Nanodrop Spectrophotometer ND-1000 USA.

Although we could now confirm that all experimental fusion proteins absorbed UV-light at 280nm wavelength, we proceeded to resolve the recombinant protein sample via SDS-PAGE analysis, as shown in figure 5.13, which confirmed a clear presentation of all pure recombinant Vap-His fusion proteins.

Other essential information that may determine the extent of validity of Vap protein syntheses such as the margin of difference and eventual fusion protein size, when compared to the protein expected size either as contained on protein databases, are summarised in the table below. The concentration, 2mg/mL of the final purified recombinant fusion proteins synthesized in this work, is presented in table 5.2 and establishes a measure of success of protein synthesis.

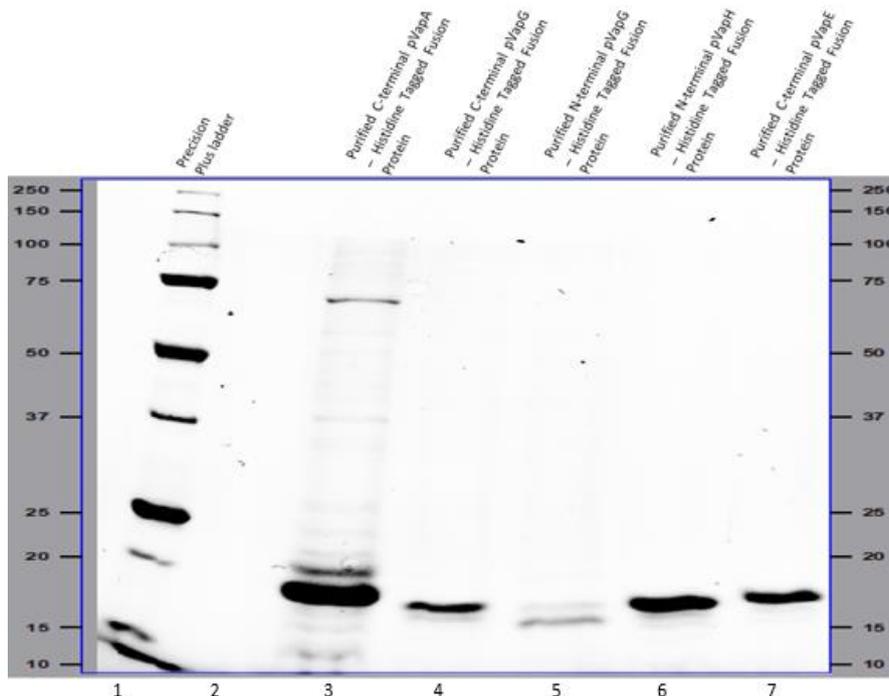


Figure 5:13: SDS-PAGE Analysis of Extracted and Dialysed Recombinant *R. equi* Vap fusion Proteins
 Lane1: Bio-rad Precision plus MW marker, *pET23a-VapA* (L3), *pET23a-VapG* (L4), *pET28a-VapG* (L5), *pET28a-VapH* (L6), *pET23a-VapE* (L7) after protein isolation by protease digestion and dialysis using the dialysis cellulose tubing. Resolved by SDS-PAGE methods at 150 Volts.

<i>R. equi</i> Vap-proteins	Experimental Vap-fusion protein	Protein Ext. coefficient (x1000) M ⁻¹ cm ⁻¹	Vap-fusion protein Concentration (per 2 μL)	Literature Source (Expassy source)
VapA C-terminal	21kDa	19.94	3.611 mg/mL	19.1 kDa
VapG C-terminal	17kDa	14.40	2.26 mg/mL	17.4 kDa
VapG N-terminal	16kDa	14.40	3.39 mg/mL	17.4 kDa
VapH N-terminal	18kDa	29.45	1.28 mg/mL	16.5 kDa
Vap E C-terminal	18kDa	19.94	2.32 mg/mL	21.5 kDa

TABLE 5.2: Experimental recombinant Vap-His tagged fusion protein molecular weight and protein concentration per 2mg/mL post protease cleavage and dialysis.

5.12 SPR Interaction of *VapA* with other *Vaps*

The advantages of surface plasmon resonance as a flexible label-free and preferred technique of investigating protein-protein interactions have been earlier discussed in section 1.2.1. This technique has been applied in this chapter aimed to explore the potential interaction of Vap protein self-association. And also, to understand the binding behaviour (kinetics and affinity) between the virulent VapA-pET23 proteins, which in this case served as the immobilized ligand molecule, against VapG-pET23, VapG-pET28, VapH-pET2 and VapE-pET23 which were applied as substrate binding partner molecules. Interestingly binding SPR method could provide Vap protein – Vap protein binding affinity and binding constant over a range of concentrations obtained from SPR binding responses (Besenicar *et al.*, 2006; Okoko, 2015). The VapA-pET23 was immobilised on a chip CM5 and other Vap proteins were flowed over it, allowing for measurement of binding responses like *in vivo* scenario. The SPR model adopted for this experiment was a custom affinity assay; this was used to monitors the interaction between the ligand (VapA-pET23) and substrates (VapG-pET23, VapG-pET28, VapH-pET2, and VapE-pET23).

5.12.1 VapA-pET23 Immobilization

The hypothesis, that when other Vap proteins are associated with VapA they form an assemblage that may attenuate VapA virulence (Coulson *et al.*, 2010). Recombinant C-terminal His -Vap A was chosen as the ligand and was immobilized via direct or covalent coupling, as presented in section 1.2.3a. The VapA-Pet23a was absorbed on a BIAcore CM5 sensor chip (via amine coupling as described in section 2.9.1.1). After 120 seconds injection of 650 $\mu\text{g}/\text{mL}$ VapA-pET23a at 5 $\mu\text{L}/\text{min}$, a signal of 1,911 response unit (RU) was obtained after the reference flow cell was subtracted (fc 2-1), see figure 5.13. This immobilization signal was found to be stable after the completion of blocking the non-reactive activated carboxymethylated dextran matrix to prevent the covalent capturing of the analyte.

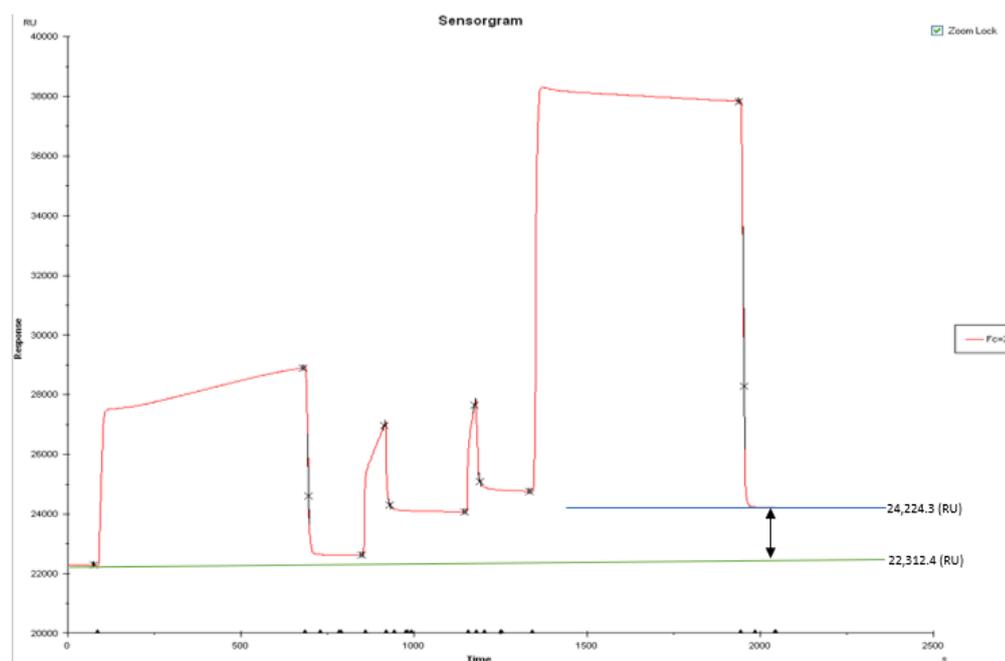


Figure 5.14: Surface Plasmon Resonance of *R. equi* C-terminal His-Vap immobilization

Sensorgram plot showing immobilization of recombinant *pET23a-VapA* to a CM5 chip. Four injections were captured, which resulted in an increased immobilization signal of 1,911 RU. The increase in RU from baseline is shown by the green and blue lines, respectively.

5.12.2 Binding Analysis of Recombinant *R. equi* Virulence Associated Proteins

The immobilization of VapA-pET23a was assessed by developing an SPR custom binding assay. The binding profile of synthesized immobilized VapA was characterized against other Vap proteins (VapG-pET23a, VapG-pET28a, VapH-pET28a, and VapE-pET23a). The binding of each synthesized recombinant Vap protein was then studied to evaluate the *in vitro* affinity and kinetic behaviour to VapA-pET23a. There is substantial evidence to suggest that virulent VapA- protein due to its unusual lipidation, is tethered to the surface of the cell wall (Okoko, 2014). For analytical runs using the custom assay, solutions of substrate Vap proteins were passed over the VapA-pET23a immobilized CM5 sensor chip for 90 seconds in an association phase at two concentrations (50 μ M and 100 μ M concentrations in HSB buffer; 10mM HEPES, 2M NaCl, 50 mM EDTA, 0.005% Tween-20). The analysis continued in a dissociation phase of 500 seconds per cycle, as buffer were injected over the chip. Results obtained from the protein-protein binding data were adjusted by applying reference subtraction (fc 2-1) of analyte signals passing concurrently through the reference flow-cell, Representative sensogram (adjusted images), as shown from figure 5.15 – 5.22. Consistent with the aim of this chapter, attention was given to understand whether recombinant Vap protein-protein interaction, are driven by the positioning (C and N- terminal) of the fusion tag on the recombinant proteins.

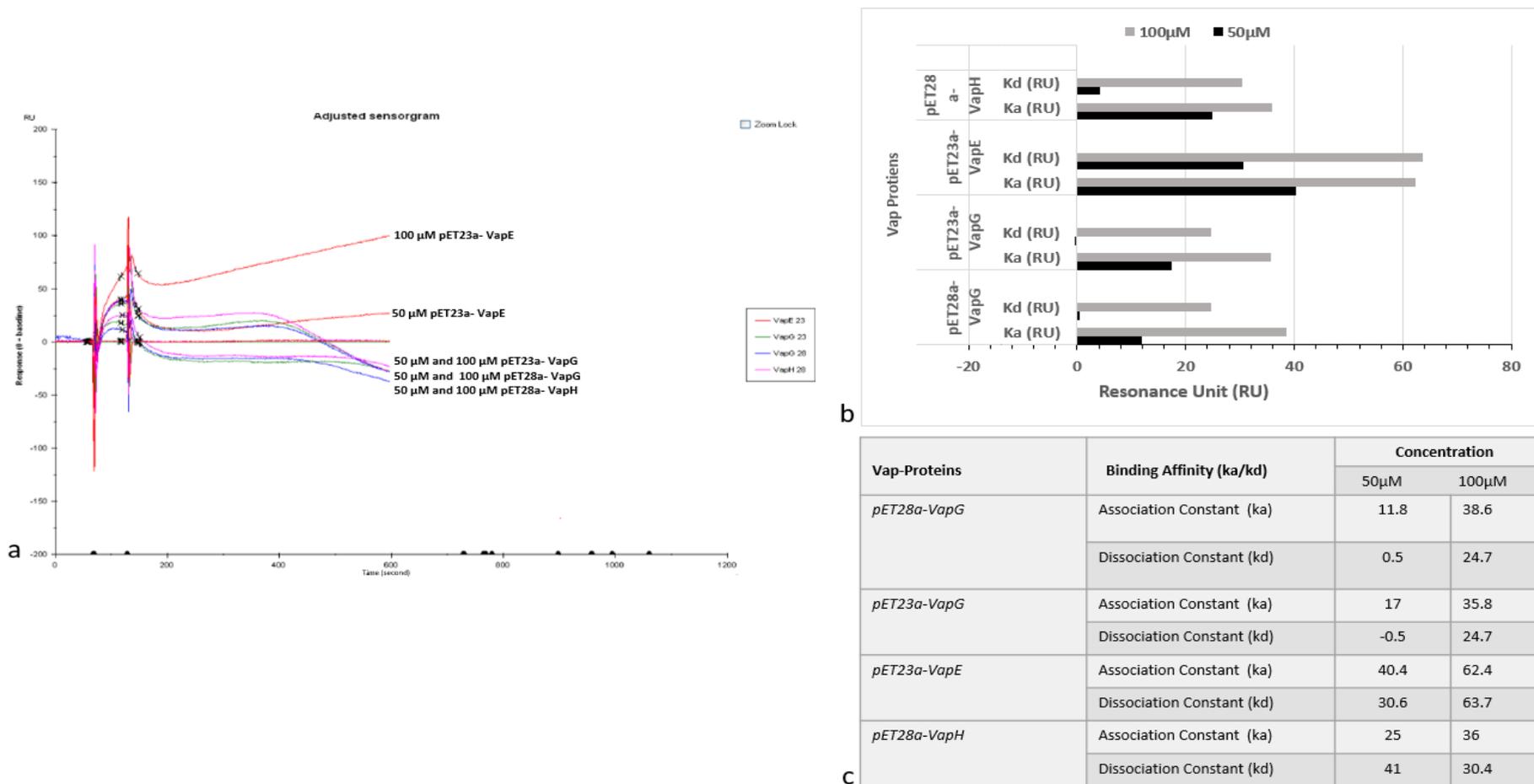


Figure 5.15 Interaction of *R. equi* VapG-pET23a, VapG-pET28a, VapH-pET28a and VapE-pET23a proteins with immobilized VapA-pET23a

(a) Sensorgram plot showing individual Vap-proteins flow over as substrate on VapA-pET23a modified CM5 chip at 50 μM and 100 μM (concentration-dependent) using a Biacore X100 **(b)** Bar chart showing association (ka) and dissociation (kd) constant values (RU) of all Vap – protein interactions against VapA-pET23a **(c)** Table showing ka/kd resonance unit (RU) for VapG-pET23a, VapG-pET28a, VapH-pET28a, and VapE-pET23a proteins binding interaction.

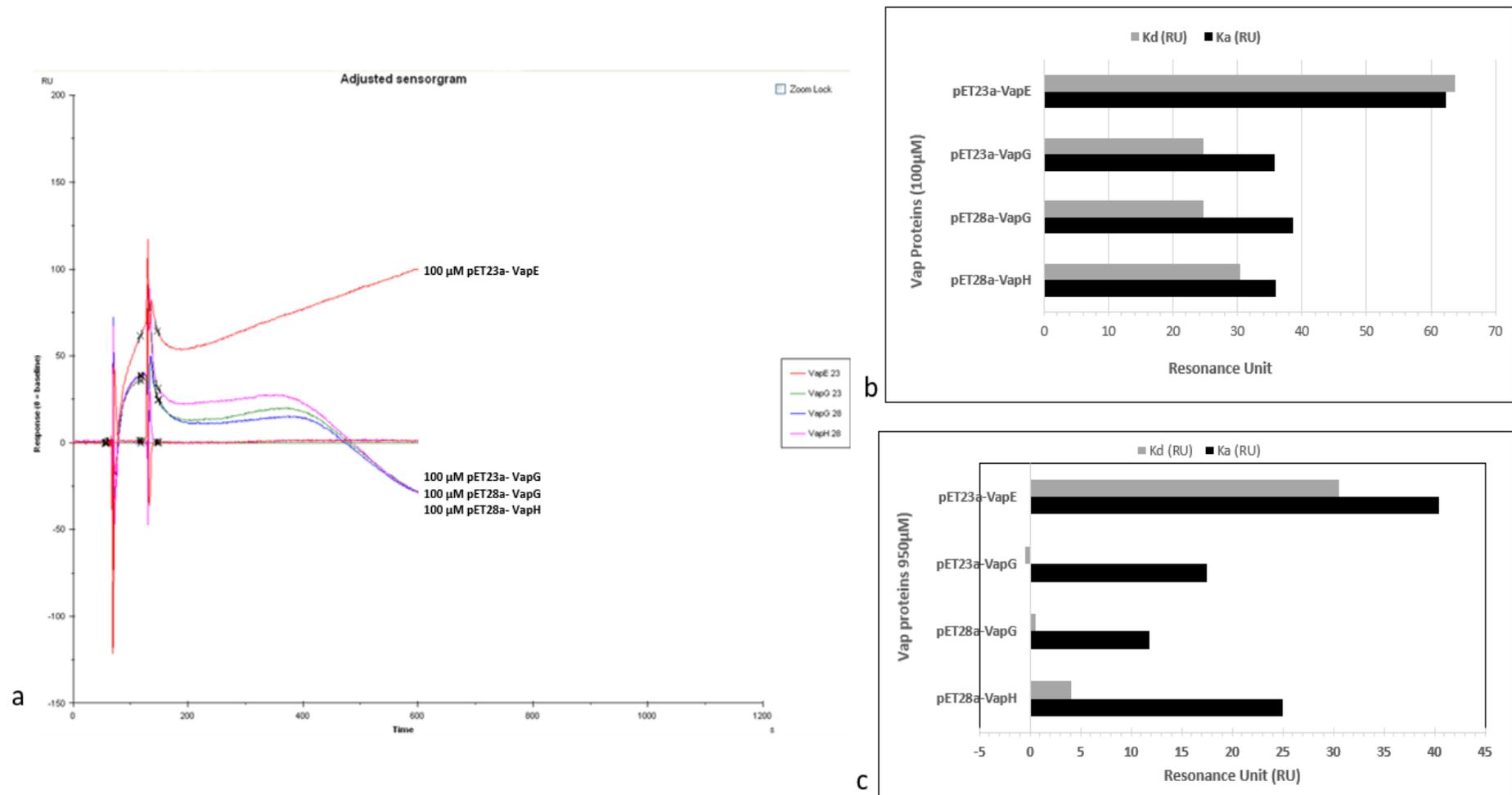


Figure 5.16 Interaction of *R. equi* Non-Concentration Dependent VapG-pET23a, VapG-pET28a, VapH-pET28a and VapE-pET23a proteins with immobilised VapA-pET23a

(a) Sensorgram plot showing a non-concentration dependent VapG-pET23a, VapG-pET28a, VapH-pET28a and VapE-pET23a proteins flowed over as substrate on VapA-pET23 modified CM5 chip at 100 μM using a Biacore X100 (b) Bar chart showing association (ka) and dissociation (kd) constant values (RU) of 100 μM VapG-pET23, VapG-pET28, VapH-pET28 and VapE-pET23 proteins interactions against VapA-pET23 (c) Bar chart showing association (ka) and dissociation (kd) constant values (RU) of 50 μM VapG-pET23, VapG-pET28, VapH-pET28 and VapE-pET23 proteins interactions against VapA-pET23a.

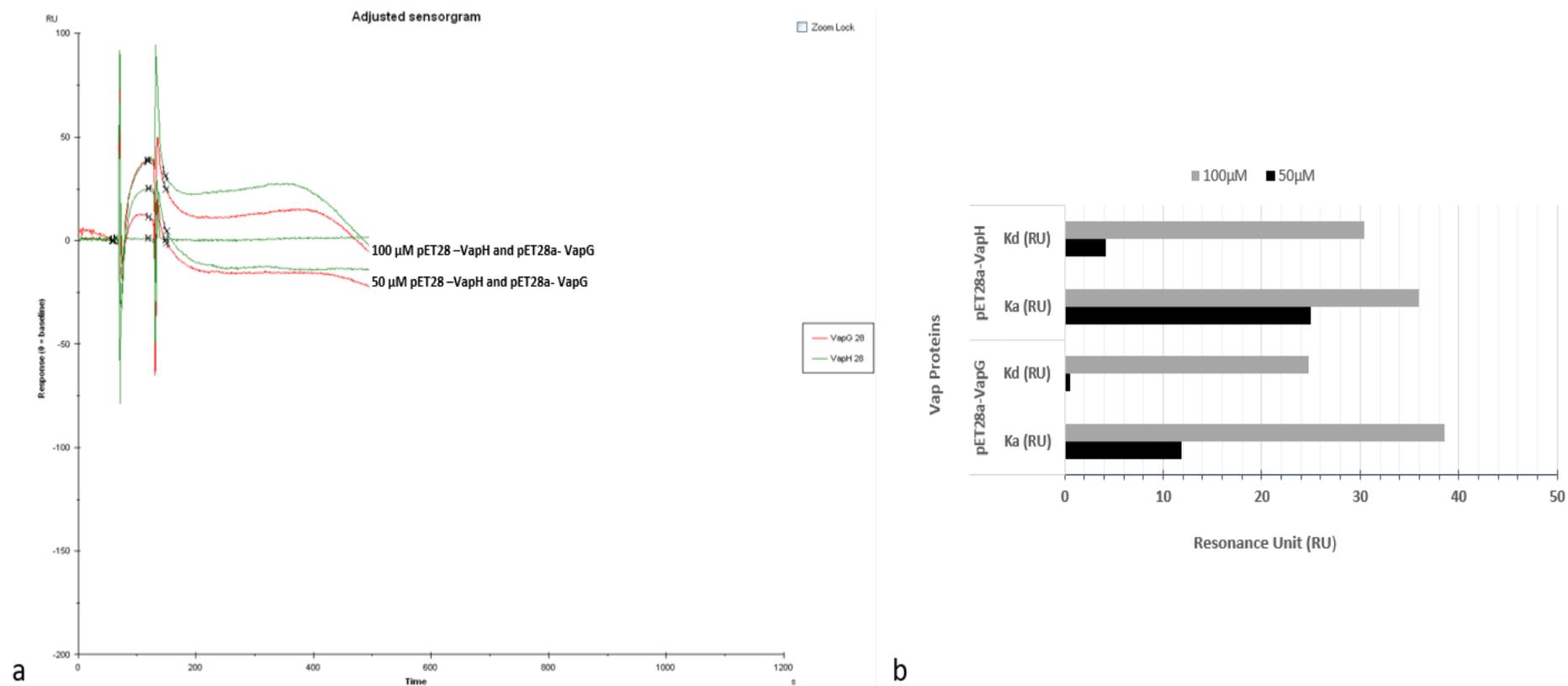


Figure 5.17 Interaction of Concentration Dependent of *R. equi* VapG-pET28a and VapH-pET28a proteins with immobilized VapA-pET23a

- (a) Sensorgram plot showing a concentration-dependent VapG-pET28a and VapH-pET28a proteins flown over as substrate on VapA-pET23 modified CM5 chip at 50μM and 100μM using a Biacore X100 (b) Bar chart showing association (ka) and dissociation (kd) constant values (RU) of both 50μM and 100μM concentration of VapG-pET28, and VapH-pET28 proteins interactions against VapA-pET23.

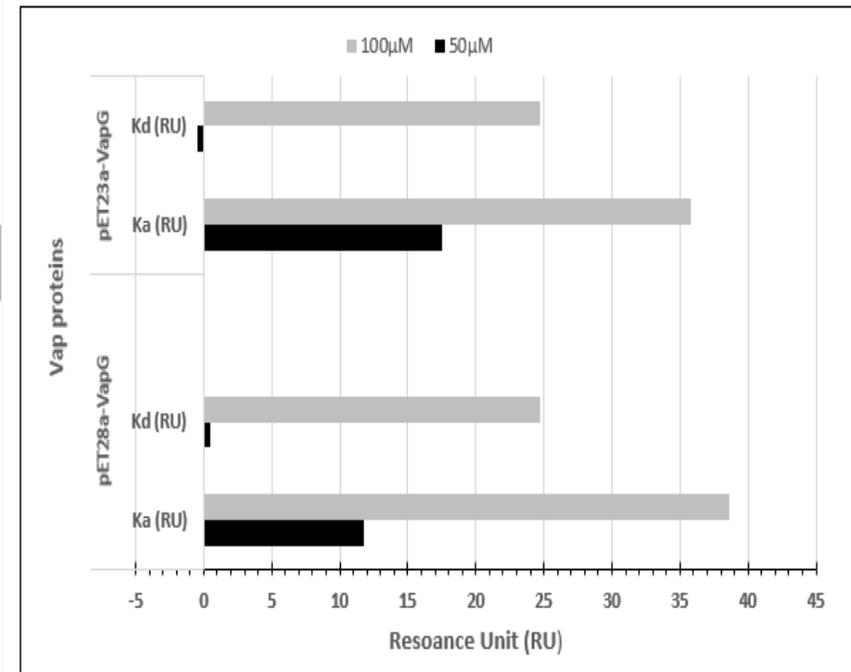
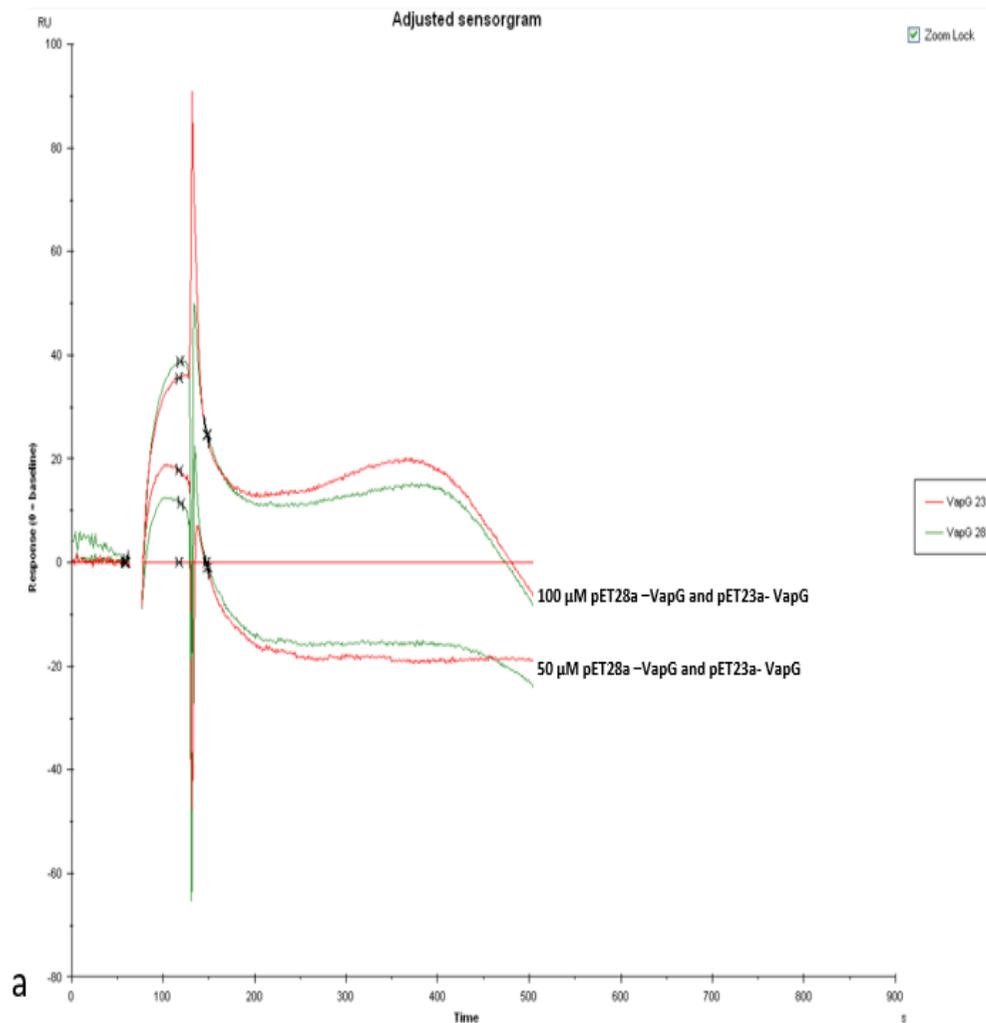


Figure 5.18 Interaction of Concentration Dependent of *R. equi* VapG-pET28a and VapG-pET23a proteins with immobilized VapA-pET23a

(a) Sensorgram plot showing a concentration-dependent VapG-pET28a and VapG-pET23a proteins flown over as substrate on VapA-pET23 modified CM5 chip at 50 μM and 100 μM using a Biacore X100 (b) Bar chart showing association (ka) and dissociation (kd) constant values (RU) of both 50 μM and 100 μM concentration of VapG-pET28a and VapG-pET23a proteins interactions against VapA-pET23.

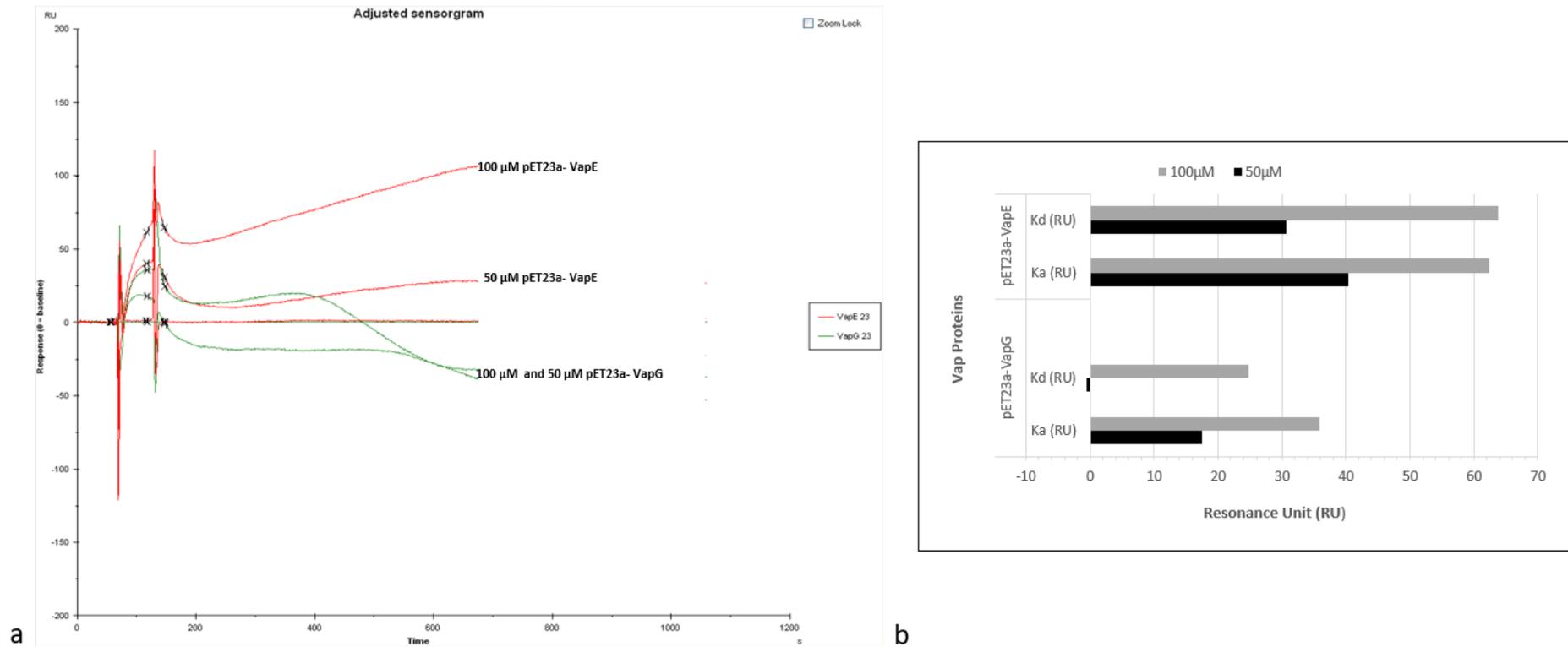


Figure 5.19 Interaction of Concentration Dependent of *R. equi* VapG-pET23a and VapE-pET23a proteins with immobilized VapA-pET23a

(a) Sensorgram plot showing a concentration-dependent VapG-pET23a and VapE-pET23a proteins flown over as substrate on VapA-pET23 modified CM5 chip at 50 μM and 100 μM using a Biacore X100 (b) Bar chart showing association (k_a) and dissociation (k_d) constant values (RU) of both 50 μM and 100 μM concentration of VapG-pET23a and VapE-pET23a proteins interactions against VapA-pET23.

A review of the 50 μ M and 100 μ M SPR binding profile of all Vap proteins under study against the immobilized recombinant pET23a-VapA indicated a concentration-dependent increased association constant (k_a) value in the order of highest to lowest k_a -value at 100 μ M sample concentration: VapE-pET23a (62.4 RU), VapG-pET28a (38.6 RU), VapG-pET23a (35.8 RU), and VapH-pET28a (30.4 RU) respectively as displayed in figure 5.15a 15.5b and 15.5c. The results also revealed that at 100 μ M sample concentration, of VapG-pET23a, VapG-pET28a and VapH-pET28a there was a gradual signal degradation of the immobilized recombinant VapA-pET23a and loss of binding just above 420 seconds of the dissociation phase. The order of association rate constant at 50 μ M sample concentration is VapE-pET23a (40.4 RU), VapH-pET28a (25 RU), VapG-pET23a (17 RU), and VapG-pET28a (11.8 RU) respectively as displayed in figure 5.15a 15.5b and 15.5c.

Similarly, at 50 μ M of VapG-pET23a, VapG-pET28a and VapH-pET28a also displayed rapid signal degradation of the ligand immobilized Vap protein just above 120 seconds of the dissociation. It was shown that during the dissociation phase, the experiment demonstrated that signal degradation at both concentrations deteriorated in different order patterns. At 50 μ M concentration, the signal degradation pattern was rapid in the order of VapH-pET28a (41 RU) > VapE-pET23a (30.6 RU) > VapG-pET28a (0.5 RU) > VapG-pET23a (-0.5 RU), while at 100 μ M concentration, the degradation pattern was gradual in the following order VapE-pET23a (63.7 RU) > VapH-pET28a (30.4 RU) > VapG-pET28a (24.7RU) > VapG-pET23a (24.7) see figure 5.16. Although the magnitude of apparent affinity showed to be largely concentration-dependent, the poor stability of recombinant VapA-pET23a protein may have caused a conformational change of the ligand on the

surface of the CM5 chip, hence the degradation. This degradation could be described as a biphasic dissociation, which causes dissociation to go beyond baseline instead of returning to baseline before regeneration, due to loss or change in characteristics of the immobilized VapA-pET23a protein. This biphasic dissociation behaviour was, however, not observed at 50 μ M and 100 μ M concentrations of recombinant VapE-pET23a protein sample during the dissociation phase of the experiment. Interestingly, the VapE-pET23a protein in a concentration-dependent manner showed the most significant association constant (k_a) /binding to the recombinant VapA-pET23a protein. The degree of significance for VapE-pET23a protein association binding constant (k_a), are summarized in Table 5.3 below, showing percentage comparison with other recombinant Vap proteins.

Table 5.3 Percentage comparison of pET23a-VapE k_a -value against Vap-proteins

Vap-Proteins	Percentage (%) k_a (RU) value of <i>pET23a-VapE</i> : other vap		Concentration	
	50 μ M	100 μ M	50 μ M	100 μ M
<i>pET28a-VapG</i>	70.7%	38.1%	11.8	38.6
<i>pET23a-VapG</i>	57.9%	42.6%	17	35.8
<i>pET28a-VapH</i>	38.1%	42.3%	25	36

The apparent affinity of Vap protein binding interactions were, evaluated, comparing Vap protein of different histidine fusion tag (N-terminal and C-terminal Vap G), of VapG-pET28a and VapG-pET23a, at 50 μ M and 100 μ M concentration (figure 5.18). The result showed that at 100 μ M concentration, the k_a value of VapG-pET28a and VapA-pET23a binding was 7.2% higher than the binding

association of VapG-pET23a and VapA-pET23a. At a lower concentration of 50 μ M, the binding association of VapG-pET23a to VapA-pET23a was 30.5% higher than the binding association between VapG-pET28a and VapA-pET23a.

The binding behavior of C-terminally tagged VapE and VapG was evaluated at 50 μ M and 100 μ M binding concentration (figure 5.19). The result showed binding association was consistently higher, as association rate constant for VapE-pET23a to VapA-pET23a binding interaction at 100 μ M concentration, had a 42% higher association rate more than the VapG-pET23a to VapA-pET23a binding. Similarly, at 50 μ M concentration, the association rate of VapE-pET23a to VapA-pET23a, was 44.4% higher than the measure of binding association between VapG-pET23a and VapA-pET23a protein interaction.

The VapE-pET23a to VapA-pET23a protein association demonstrated the highest and most significant binding association of 42% and 44.4% at 100 and 50 μ M concentration respectively, compared to all other vap proteins and concentration. Interestingly dissociation constant (k_d) in this SPR custom assay was thought not to be an accurate value or true reflection of the dissociation rate (K_{off}). A major observation in this binding interaction experiment was that the Vap protein-protein interaction was consistent in its demonstration of a fast K_{on} and K_{off} rate, as observed in figure 5.15 and 5.19.

The poor stability VapA-pET23a protein on the surface of the sensor chip over a long time suggested that although the binding association constant (k_a -values) rate, for all vap protein interaction could be retrieved and represents the apparent affinity of all Vap proteins under investigation, in this binding interaction studies. The behaviour of the dissociation fit (K_{off}), are indicative, that full kinetic assay and

data may prove difficult to be generated, and the fast off-rate are indicative to the inability of the BIAcore X100 machine to detect and characterize the kinetics data of this experimental *R.equi* vap-protein binding interaction.

5.13 Thermal Shift Assay

A review of the *R. equi* Vap protein interaction result showed that the control protein sample, successfully increased in thermal shift both in the absence and presence of ligand molecule respectively. At high temperature the protein begins to melt, to release its internal regions and unfold to become linearized. Being run in triplicates figure 5.20a shows the expected gradual increase of the panel, which starts from a region of low fluorescence, to attains its peak where proteins are unfolded and decrease proportionally in fluorescence and temperature gradually. Thermal shift cycle, here is indicated by a fluorescence across a temperature cycle ranging from 25°-95°C in 0.05-degree steps at 5 sec per step.

Application of *R.equi* Vap A in figure 5.20b showed the characteristic thermal shift for the control protein while VapA shows a very high initial fluorescence that slowly falls without any thermal shift observed. Similarly figure 5.20c shows a repeat of figure 5.20b but with the SYPRO ruby dye at different concentrations.

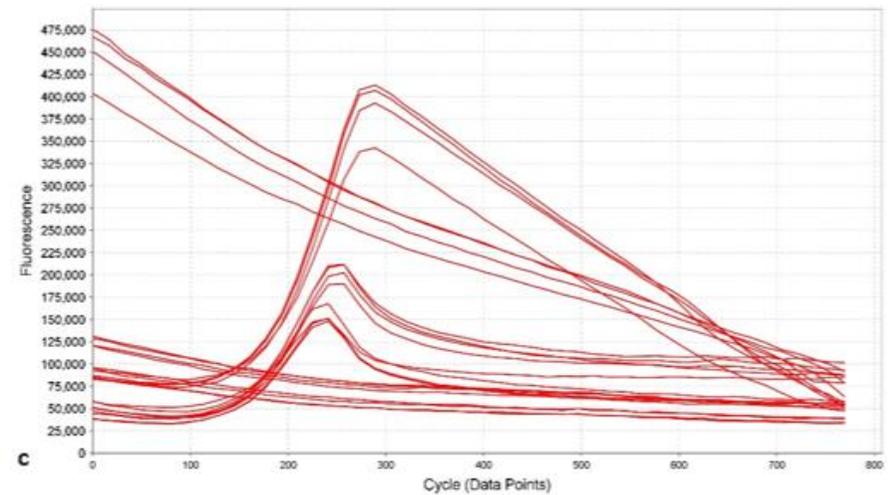
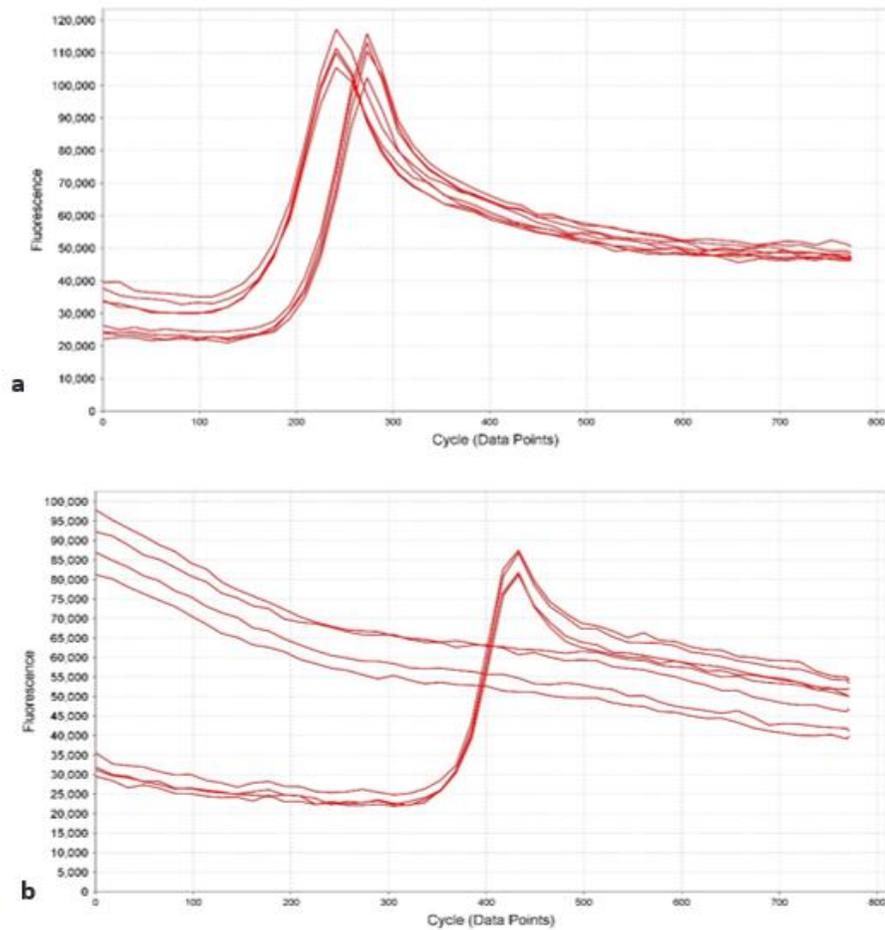


Figure 5.20: Initial thermal shift assay control and *R. equi* Vap A experiment.

The figure shows a range of experiments using the thermal shift assay kit (Thermo). Each figure shows fluorescence across a temperature cycle ranging from 25°-95°C in 0.05-degree steps at 5 sec per step each sample was run in triplicate. Panel A shows the kit control protein with and without ligand respectively with an increased thermal shift. Panel B Shows the characteristic thermal shift for the control protein but VapA shows a very high initial fluorescence that slowly falls. Panel C is a repeat of panel B but with the SYPRO ruby dye at different concentrations.

5.14 Discussion

5.14.1 Introduction

The *Rhodococcus equi* mediated pneumonia pathogenesis, remains a major cause of stock wastage, with significant cost implication for treatment and fatality in the equine industry (Dawson *et al.*, 2010; Thunes, 2018). Although a degree of protective immunity against the disease have been developed by some foals, around the world, endemic *Rhodococcus equi* contributes between 40% to 50% mortality rate of foals, and about 80% morbidity in affected foals (Giles *et al.*, 2016; Woodmansey, 2018).

Researchers in this field, have been targeting prevention of *Rhodococcus equi* through the development of a vaccine, (Lohmann *et al.*, 2013; Giles, *et al.*, 2015). The earlier lack of success in developing a vaccine was attributed to the ambiguity of the immunity to *R. equi* and the intracellular nature (complex cell envelope) of the organism (ability to survive harsh macrophage environment). The application of a wide range of antibiotics, for the management and the treatment of *R. equi* infections was not only a limited option, but demonstrated mixed results, when antibiotics combination therapy of rifampin and clarithromycin or azithromycin treatments were introduced (Giguere *et al.*, 2012; Venner *et al.*, 2013).

From the perception of our research, facts presently known, is that (i) the VapA has been described as a major virulence determinant since mutants that lack the VapA gene were found to be unable to cause *Rhodococcus equi* infection. (ii) Rofe and co-worker have successfully resolved the hypothesis that VapA disrupts lysosomal

function by allowing *R. equi* to survive intracellularly (Rofe *et al.*, 2017). To mitigate the concerns of macrophage disruption, study has shown VapA association with other vap proteins to be essentially useful towards attenuating VapA virulence (Jain *et al.*, 2003, Vazquez-Boland *et al.*, 2013). Although success have now been made with the discovery of the PNAG antibody vaccine, the many years of *R. equi* Vap study, the actual biology by which other vap proteins can successful self-associate with virulent plasmid (VapA), has become a focus of research and explored towards mitigating the *R. equi* disease in foals.

Okoko and co-workers first investigated the hypothesis of VapA and other vap proteins, self-association via biophysical protein- protein interaction using a CM5 carboxymethylated dextran chip, covalently immobilizing recombinant VapA on a surface plasmon resonance chip and assaying for binding affinity of other vap proteins that were run over the immobilised VapA (Okoko *et al.*, 2015). This current work builds on their recommendation and attempted to synthesis Vap proteins differently, exploiting the biology of affinity fusion tag proteins, that were cloned into N and C terminals of the vapG-pET23a, vapG-pET28a, vapH-pET28a, vapE-pET23a, proteins. This experiment validated that all vap plasmids investigated under this study were successfully cloned into pET23a and pET28a vectors, respectively, and retrieved for the purpose of synthesis of recombinant Vap proteins.

A fundamental question that we set out to address in this study, was whether the positioning of both N-terminal (pET-28) and C-terminal (pET-23a) fusion tag on the

R. equi Vap proteins, could inform or in fact enhance a range of Vap protein interaction with the virulent VapA protein, using the biophysical techniques.

5.14.2 Overexpression of recombinant Vap-His Tag fusion proteins

The experimental Vap proteins A, E, G and H were designed to be cloned using a plasmid vector pET23a and pET28a in this order for pET23a- *pVapA*, pET23a- *pVapE*, pET28a- *pVapG*, pET23a- *pVapG* and pET28a- *pVapH*. The pET28a and pET23a expresses an N and C-terminal fusion tagged hexa-histidine (HIS) protein respectively are transcribed by the T7 RNA polymerase gene in a *E. coli* expression system (Rosano and Ceccarelli, 2014). The final plasmids constructs were satisfactorily re-transformed into C43 (DE3) *E. coli* with exception of pET23a- *pVapD*, pET28a- *pVapD*, pET28a- *pVapE*, and pET23a- *pVapH* and overexpression of experimental vap proteins were induced in at 37°C for 2 hours. The pET expression system featuring the T7 promoter is by far widely used system for heterogeneous expression in *E. coli* (Graslund *et al.*, 2008).

The appreciable yield of recombinant soluble C-terminal His-tag *pVapA*, C-terminal His-tag-*pVapE*, N-terminal His-tag *pVapG*, C-terminal His-tag *pVapG* and N-terminal His-tag *pVapH* fusion protein recovery, meant that further optimisation of the insoluble fraction was discontinued. Previously a different model of *E. coli* expression strategy was adopted by Letek *et al.*, (2008) who utilised *E. coli* at 20°C for 24-48 hours to grow an *R. equi* vapB in the brain heart infusion medium to produce VapB *R. equi* proteins.

During a study of VapA disruption of endolysosome function, Rofo and co-workers, had successfully expressed *R. equi* VapA, using *E. coli* (B21 DE3 pLysS) strain at 37°C

were transformed plasmid tagged either to GST- or His₆-tagged proteins and were grown in 2TY broth media (Rofe *et al.*, 2016).

Osmotic shock was similarly mitigated during the overexpression of experimental vap proteins, by ensuring protein overexpression media (yeast extract media) were supplemented with osmotic shock buffers 0.5M sorbitol and 0.5 mM betaine which increases the solubility of recombinant proteins in bacterial expression (Hofmann, 2015). This may account for the recovery of soluble recombinant C-terminal His-tag *pVapA*, C-terminal His-tag-*pVapE*, N-terminal His-tag *pVapG*, C-terminal His-tag *pVapG* and N-terminal His-tag *pVapH* fusion protein under the described overexpression condition as described in section 2.5.2 and 2.5.3.

5.14.3 Recombinant Vap Protein Purification

The purity of soluble recombinant C-terminal His-tag *pVapA*, C-terminal His-tag-*pVapE*, N-terminal His-tag *pVapG*, C-terminal His-tag *pVapG* and N-terminal His-tag *pVapH* fusion protein, were satisfactorily resolved using immobilized metal ion affinity chromatography (IMAC) protein purification. The Ni-NTA sepharose, reacting with both the C and N-terminally tagged His - *R. equi* Vap proteins, achieved high level recombinant protein purification under both native and denatured condition. Purified recombinant His-tagged vaps proteins were easily detected due to the ability of histidine residues to bind several types of immobilized metal ions, including zinc, nickel, cobalt, and copper, under specific buffer conditions (Jansen, 2011).

Interestingly due to the small molecular size of His-tags, (~2.5 kDa), they were believed not have significantly interfere with the function and structure of

experimental *R. equi* vap proteins (Gräslund *et al.*, 2008; Booth *et al.*, 2018). Okoko and co-workers had similarly used IMAC purification method for protein purification of their *R. equi* vap protein, followed by an additional step of anion exchange chromatography (Okoko *et al.*, 2015)

5.14.4 SPR interaction of Vap proteins in free solution with immobilized C-terminal His-tag VapA

Evaluating C-terminal His-tag *pVapA* fusion protein binding interaction against C-terminal His-tag-*pVapE*, N-terminal His-tag *pVapG*, C-terminal His-tag *pVapG* and N-terminal His-tag *pVapH* fusion protein were modelled in vitro and in solution using a custom kinetics - 1:1 binding assay of the surface plasmon resonance, similarly, adopted by previous *R. equi* Vap protein interaction investigations (Okoko *et al.*, 2015; Drin, 2019). A C-terminal His-tag *pVapA* immobilisation level of 1,911RU affixed on the CM5 chip carboxymethyl dextran was satisfactory, aiming and achieving a high immobilisation RU in kinetic assay is thought increase binding hits. Considering that the limitations of mass transport effect is of no concern in this assay model. Okoko and co-workers had previously attained a 1,483RU immobilization level for VapA on a CM5-chip, during a binding interaction investigation with other non-virulent Vap proteins (Okoko *et al.*, 2015).

The SPR custom kinetics- 1:1 binding assay model applied, is referred as a custom protocol due to the contact time and dissociation time adjustment of 60 seconds and 500 seconds made respectively. In a conventional kinetics binding assay, the contact and dissociation times are 120 seconds and 600 seconds respectively. The rationale for developing a custom method, was to minimize the length of time utilised in running other Vap proteins over the immobilised C-terminal His-tag

pVapA fusion protein. Since the experiment was designed in a single kinetic run measured association (k_a) and dissociation (k_d) in a 1:1 binding ratio.

The parameters of interest here, are the k_a -value being the association rate constant ($M^{-1}s^{-1}$) and the k_d -value dissociation rate constant (S^{-1}). The results revealed that at two different concentrations of Vap protein SPR investigation the k_a -values change, meaning that the binding interaction between C-terminal His-tag *pVapA* and C-terminal His-tag-*pVapE*, N-terminal His-tag *pVapG*, C-terminal His-tag *pVapG* and N-terminal His-tag *pVapH* fusion proteins are dose-dependent.

The results suggest that, at $100\mu M$ sample concentration, C-terminal His-tag-*pVapE* has the highest association rate constant value of 62.4 RU, followed by VapG-pET28a (38.6 RU), VapG-pET23a (35.8 RU), and VapH-pET28a (30.4 RU). While at $50\mu M$ sample concentration, the order is not the same VapE-pET23a (40.4 RU), VapH-pET28a (25 RU), VapG-pET23a (17 RU), and VapG-pET28a (11.8 RU). This means that experimental *R. equi* Vap protein binding of VapA apart from being dose dependent, are also affected by the position of His tag affinity on either N or C terminal of the reacting Vap protein.

Kinetic- 1:1 binding model equation



Where the following parameters are fitted by the model:

A	Ligand
B	Substrate
AB	Ligand/Substrate Complex
K_a	Association rate constant ($M^{-1}s^{-1}$)
K_d	Dissociation rate constant (S^{-1})
Rmax	Analyte binding capacity of the surface (RU)
t_c	Flow rate-independent part of mass transfer constant
RI	Bulk refractive index contribution in sample

Although there were dissociation rate constant values generated for N-terminal His-tag *pVapG*, C-terminal His-tag *pVapG* and N-terminal His-tag *pVapH* in this experiment. It is thought that those values do not reflect the absolute dissociation behaviour as the immobilised C-terminal His-tag *pVapA* was observed to be unstable and gradually degrading on the surface of the biosensor chip. This was visibly observed in figure 5.15a as Rmax signal dropped below baseline and may be as a result of some conformational changes of the C-terminal His-tag *pVapA* on the CM5 chips surface. There is however an exception to the biphasic dissociation, with C-terminal His-tag-*pVapE*, whose dissociation rate behaviour showed to be consistent as expected in kinetic binding assay model, hence may not have been influenced by the ligand degradation.

To successfully generate kinetic data from this vap-protein - protein interaction experiment, the k_a and k_d values must demonstrate a 1:1 plot fitting ratio. Since that is not the case with experimental k_d -values obtained from this experiment, kinetic data for *R. equi* Vap protein-protein binding could not be generated.

The C-terminal His-tag *pVapA* degradation on the CM5 chip surface, prevented duplicate or triplicate binding interaction runs. These runs are essential because of the validity and robustness it confers on reproducible binding affinity and kinetics data generated. Similarly, the ligand degradation meant that kinetic data are unable to be generated due to false k_d -values.

The binding association rate constant (k_a -values) of N-terminal His-tag *pVapG*, C-terminal His-tag *pVapG* and N-terminal His-tag *pVapH* proteins in ratio to C-terminal His-tag-*pVapE* protein (table 5.3) helps to validate and conclude, that *R.*

equi vap protein – protein interaction is both non-concentration dependent and also non-dependent on the positioning of the His-tag on the substrate molecule either at the N or C-terminal based on data generated in this work.

5.14.5 Thermal Shift Assay interaction of *R.equi* Vap proteins

Thermal shift assay were attempted of *R.equi* VapA proteins, but did not work, because vapA protein appears to be hydrophobic and upon unfolding, were found to bind the SYPRO ruby dye. Although they both showed an initial high fluorescence in figure 5.20b and figure 5.20c that gradually fell, no thermal shift was observed. This suggests that *R. equi* vap protein are unstable under these experiment condition.

Chapter 6

General Discussion

6.0 Chapter Six

6.1 Final Thoughts and Reflections

6.1.1 Recombinant protein synthesis

The synthesis of recombinant proteins is essential aspect of biotechnology and biomedical sciences. For recombinant proteins to be beneficial for research studies, they usually are expressed and purified in high quantities. Protein synthesis via bacterial overexpression using *Escherichia coli* has been the focus of this thesis. This system is often a guaranteed source to produce high quantity and quality recombinant proteins required for research (Wagner *et al.*, 2008).

However, bacterial recombinant protein synthesis processes are not always as straight forward as the theory presents. For example, we found that although HMGR and AGTR1 genes successfully cloned into pGEX-6p-1, pMAL-c5x, pMAL-p5x and pET32a respectively in chapter 3 recombinant proteins could not be expressed. The limitations of bacterial protein overexpression is not surprising, often this can be overcome by determining the most stable fusion tag and expression strain for their protein synthesis work (Booth *et al.*, 2018).

The application of different fusion tags (with strong promoter, RBS, and transcription terminator), codon-optimization, low overexpression temperature and optimised protease-deficient host strains were troubleshooting approaches employed to facilitate the expression of the hydrophobic trans-membrane and highly toxic CHD proteins we attempted to overexpress in chapter 3. But to no avail and certain classes of proteins are always going to be difficult or potentially impossible to express in bacteria.

The application of araBAD operon based pBAD vector system, could be applied in the future towards the expression of proteins such as HMGCR and AGTR1 proteins. Széliová and co, had this system may have advantages suitable for the expression of toxic and hydrophobic proteins. These advantages include tight regulation of transcription, moderately high expression levels, and induction by an inexpensive and non-toxic monosaccharide L-arabinose (Széliová *et al.*, 2016), but this does not directly address the issue of hydrophobicity.

However, proteins under investigation in chapter 4 and 5 confirmed that bacteria (*E. coli*) were viable, efficient, and effective host system, for the production of high quantity recombinant heterologous soluble proteins that can be utilised for function studies. We know from previous work (Wisdom Okoye's Masters Project, unpublished) that CCL5 is almost insoluble as a GST-fusion protein, but in this work expressing CCL5 using MBP fusion-tag, resolved the solubility concerns.

6.1.2 Protein Purification Method Validation

Two different protein purification methods, the affinity purification on amylose resin and immobilized metal ion affinity chromatography (IMAC) purification were applied through this work in chapter 4 and 5 respectively. The results showed that the pure recombinant CCL5/MBP proteins recovered after batch purification in chapter 4, had nucleic acid impurities or contamination. This was evident in figure 4.25 where pure CCL5/MBP proteins showed to attain peak absorbance at 260nm, rather than at 280nm under UV-spectrum light. On the other hand, purified *R. equi* Vap proteins recovered from IMAC purification in chapter 5, had no nucleic acid

impurities as demonstrated in figure 5.11. As IMAC purified *R. equi* Vap proteins showed to attain peak absorbance at 280nm.

This suggests that protein purification methods impact the extent of genomic DNA contamination present in a recombinant pure protein sample post-purification. The result demonstrates also that chromatographic based protein purification methods (used for Vap proteins) show to be more efficient, high-throughput and robust than non-chromatographic based protein purification methods (used for CCL5 proteins). However, this adds time and cost to the process, which has to be considered when comparing it to the additional steps of nuclease digestion and the dialysis required by batch purification method. Ultimately, this depends on the workflow required.

6.1.3 Sample preparation and Optimization

The nucleic acid contaminated pure recombinant CCL5/MBP proteins were treated by DNase and RNase and membrane dialysis as described in section 2.8.0. Optimised pure recombinant CCL5/MBP proteins without nucleic acid contamination are shown thereafter in figure 4.26 to attain peak absorbance at 280nm.

During preliminary SPR immobilization of biotinylated heparin (Heparin-biotin sodium salt acquired from Sigma Aldrich) on a streptavidin (SA) biosensor chip, the results, revealed presence and interfering activities of free flowing and excess biotin in stock sample. The interference created an unregulated, highly specific, and irreversible biotin and streptavidin binding complex (Liu *et al.*, 2016). Consequently, the immobilization of Sigma Aldrich purchased biotinylated heparin

sodium salt could not be achieved. We resolved and optimised the Heparin-biotin sodium salt product by dialysis using a Slide-A-Lyzer 2K (2,000 MWC) dialysis cassette (Thermo Scientific, USA) as described in section 2.8.1. Excess non-conjugated biotin, a vitamin of 0,244 kDa molecular mass, was found to be eliminated through the permeable 2,000 MWC membrane of the Slide-A-Lyzer dialysis cassette, leaving with the cassette an optimized 15kDa biotinylated heparin free of excess biotin contamination. Further application of the dialyzed Sigma Aldrich Heparin-biotin sodium salt showed a successful immobilization profile of the surface of SA biosensor chip figure 4.21.

6.1.4 Biophysical functional Assay

Biophysical functional assay is important in to biological, biomolecular and drug discovery research and development. The demand to analyse a wide range of samples (proteins, nucleic acids, lipids and membrane-associated molecules, carbohydrates, low molecular weight compounds (> 200 Da), viruses/bacteria and whole cells) for biomolecular interactions are currently on the rise (Robin, 2015). Biophysical assays are useful in determining binding site mapping, thermodynamic profiles, specificity, kinetics, and affinity (Mboge *et al.*, 2018). Each of these biophysical techniques has its advantages and disadvantages, and they are best used in combination.

In the context of this study, although the characteristic and science of the various proteins investigated in chapter 3, 4 and 5 differ in biology, the concept of investigating protein – substrate (as in chapter 3 and 4) or protein – protein (as in

chapter 5) binding interactions through the application of biophysical technique remained consistent.

The SPR as widely applied in this work, is certainly one of the most appropriate high-throughput screening methods (Genick *et al.*, 2014). It is a very sensitive technique with its limit of detection increasing proportionally with the model of the machine. The SPR requires only micrograms of sample and can detect binding events at physiologically relevant nanomolar concentrations. A vital technical consideration of for the application of SPR biophysical assay in this work, is because the proteins under investigation do not need to be radio-labelled or modified as SPR is a label-free technique. Although the SPR experiments ran with few limitations and future run will require further modifications and sample concentration optimization. The application of this SPR technique successfully demonstrated the two types of coupling method in the SPR instrumentation which are the indirect and direct coupling as in chapter 4 and 5 respectively. Relatively true binding data were also generated, using apparent affinity data via steady-state assay. Also, affinity constant information's were obtained for CCL5/MBP – GAG and *R. equi* Vap protein - protein binding interactions in chapter 4 and 5 respectively.

This techniques and results obtained from this work, demonstrates that a combination of biological, functional, and biophysical tools can generate some useful scientific data. These data can be useful in understanding the biology of complex protein - protein and protein – ligand interactions, hence enhancing research and development.

It may be useful to further SPR investigation of these binding interaction using a more recent model of SPR machine with a higher limit of detection. With the

availability of funding, already synthesized recombinant wild type and mutagenic CCL5 alongside *pET23a- pVapD*, *pET28a- pVapD*, *pET28a- pVapE*, and *pET23a- pVapH*, *pET23a- pVapE*, *pET28a- pVapG*, *pET23a- pVapG* and *pET28a- pVapH* could have been purchased and applied to SPR binding interaction assay, with biotinylated heparin as the ligand for the CCL5/GAG assay and *pET23a- pVapA* and *pET28a- pVapA* as ligands for *R. equi* Vap protein-protein interaction studies.

6.2 Conclusion

It is the view of this work that *E. coli* except for few toxic gene continues to demonstrate high efficiency in the production of heterologous recombinant proteins. That this work reflects a series of troubleshooting, method validation, and method optimization, in biological to biophysical assays in a producible and systematic manner. Although we successfully cloned the wildtype proteins under investigation in chapter 3, as well as generated crucial SNP data of mutants with clinical annotation, this work was unable to address the aim and objective of the project in chapter 3. This was solely, due to lack of progress solving toxicity issues with overexpression of AGTR1 and the full length HMGCR proteins.

Chemokine – GAG binding interaction data showed that CCL5 binding interaction with heparin was concentration dependent and that the recombinant WT-CCL5-MBP protein, demonstrated the highest apparent affinity for the GAG Heparin in this experiment, followed by CCL5-MutLow-MBP and CCL5-MutHigh-MBP which is consistent with literature.

The SRP functional assay data of *R. equi* vap protein – protein interaction, concluded that VapA binding interaction with other experimental vap proteins were both non concentration dependent and also not dependent on the positioning of the His-tag on the substrate molecule either at the N or C-terminal.

6.3 Recommendations for Future Work

1. To explore means through which wild type and mutagenic AGTR1 genes can be overexpressed, using *E. coli* overexpression system where the host strain is able to mitigate AGTR1 toxicity. This can be achieved by applying lower copy number plasmids containing tightly regulated promoters (like the *araP_{BAD}* promoter)
2. To explore means through which full-length wild type HMGCR genes can be overexpressed, using *E. coli* overexpression system thereby mitigating the cascade of unexplained events that leads to truncation of HMGCR gene.
3. Further evaluate potential method of separating maltose binding protein (MBP) tag, from the experimental wild type and mutagenic recombinant CCL5/MBP fusion complex.
4. To validate SPR steady-state apparent affinity interaction of recombinant wild type and mutagenic CCL5 against GAG-Heparin using 2 units of higher concentrations and running binding assay in triplicates.
5. Obtaining additional data on the interactions of the recombinant wild type and mutagenic CCL5 proteins using transendothelium flow assay to enhance the initial preliminary data.
6. To clone, overexpress and purify of *pET23a- pVapD*, *pET28a- pVapD*, *pET28a- pVapE*, and *pET23a- pVapH* and apply them towards SPR binding affinity and

kinetics interaction studies alongside *pET23a- pVapE*, *pET28a- pVapG*, *pET23a- pVapG* and *pET28a- pVapH*. Experiment to have *pET28a- pVapA* as the immobilized ligand.

7. To desalt the Biotinylated Heparin, using Zeba Spin desalting column and valid the sample using mass spectrometer, before biophysical analysis

8. To explore means of mitigating *pET23a- pVapA* instability and degradation on the surface of CM5 biosensor chip, by applying a running buffer with less NaCl concentration. Since high salt concentration interferes with the electrostatic interaction activity on the surface of the CM5 biosensor chip.

9. To subject all recombinant protein products to Mass spectrometry prior to binding affinity assay

References

- Adams, J. (2008) The complexity of gene expression, protein interaction, and cell differentiation. *Nature Education* 1(1):110
- Adnan, M., (2020). *Problem In Re-Binding Of Maltose Binding Protein To Amylose Resin*. [online] ResearchGate. Available at:<https://www.researchgate.net/post/Problem_in_rebinding_of_maltose_binding_protein_to_amylose_resin> [Accessed 12 March 2020].
- Akiba, H., Tamura, H., Kiyoshi, M., Yanaka, S., Sugase, K., Caaveiro, J. and Tsumoto, K. (2019). Structural and thermodynamic basis for the recognition of the substrate-binding cleft on hen egg lysozyme by a single-domain antibody. *Scientific Reports*, 9(1). Kastritis, P. and Bonvin, A. (2013). On the binding affinity of macromolecular interactions: daring to ask why proteins interact. *Journal of The Royal Society Interface*, 10 (79), p.20120835.
- Allen, S., Crown, S. and Handel, T. (2007). Chemokine:Receptor Structure, Interactions, and Antagonism. *Annual Review of Immunology*, 25(1), pp.787-820.
- Almulki, L., Noda, K., Amini, R., Schering, A., Garland, R., Nakao, S., Nakazawa, T., Hisatomi, T., Thomas, K., Masli, S. and Hafezi-Moghadam, A., 2008. Surprising up-regulation of P-selectin glycoprotein ligand-1 (PSGL-1) in endotoxin-induced uveitis. *The FASEB Journal*, 23(3), pp.929-939.
- Alvarez, H. M., Silva, R. A., Herrero, H., Hernandez, M. A. and Villalba, M. S. (2013) Metabolism of triacylglycerols in *Rhodococcus equi* species: insights from physiology and molecular genetics. *Journal of Molecular Biochemistry*2(1): 69-78.

- An, G., Wu, F., Huang, S., Feng, L., Bai, J., Gu, S. and Zhao, X., 2019. Effects of CCL5 on the biological behavior of breast cancer and the mechanisms of its interaction with tumor-associated macrophages. *Oncology Reports*,
- Anthis, N. and Clore, G. (2013). Sequence-specific determination of protein and peptide concentrations by absorbance at 205 nm. *Protein Science*, 22(6), pp.851-858.
- Apweiler, R., Hermjakob, H. and Sharon, N. (1999). On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1473(1), pp.4-8.
- Archana, V.K., Sampath, K.K, Ann-Kathrin, U. and Martin, G. (2015). Heparan sulphate as a regulator of leukocyte recruitment in inflammation. *Curr Protein Pept Sci*. 16(1): 77–86
- Arimont, A., Sun, S., Smit, M.J., Leurs, R., De Esch, I.J. and De Graaf, C. (2017). Structural Analysis of Chemokine Receptor-Ligand Interactions. *J Med Chem*. 60(12): 4735–4779
- Ateeq, B., Tomlins, S. and Chinnaiyan, A., (2009). AGTR1 as a therapeutic target in ER-positive and ERBB-negative breast cancer cases. *Cell Cycle*, 8(23), pp.3794-3795.
- Bacon, K., Baggiolini, M., Broxmeyer, H., Horuk R, R., Lindley, I., Mantovani, A., Maysushima, K., Murphy, P., Nomiyama, H., Oppenheim, J., Rot, A., Schall, T., Tsang, M., Thorpe, R., Van Damme, J., Wadhwa, M., Yoshie, O., Zlotnik, A. and Zoon, K. (2002). Chemokine/Chemokine Receptor Nomenclature. *Journal of Interferon & Cytokine Research*, 22(10), pp.1067-1068.

- Baker, R. T. (1996). Protein expression using ubiquitin fusion and cleavage. *Curr. Opin. Biotechnol.* 7, 541–546.
- Baron, O. L. and Pauron, D. (2014). Protein-lipid Interaction Analysis by Surface Plasmon Resonance (SPR). *Bio-protocol* 4(18): e1237.
- Barth S., Huhn M., Matthey B., Klimka A., Galinski E. & Engert A. (2000). Compatible solute-supported periplasmic expression of functional recombinant proteins under stress conditions. *Appl. Env. Microbiol.* 66, 1572-1579.
- Bauer, P., Welbourne, T., Shigematsu, T., Russell, J., Granger, DN. (2000). Endothelial expression of selectins during endotoxin preconditioning. *Am J Physiol Reg Integr Comp Physiol.* 279: R2015–21.
- Baumgarten, T., Ytterberg, A., Zubarev, R. and de Gier, J. (2018). Optimizing Recombinant Protein Production in the Escherichia coli Periplasm Alleviates Stress. *Applied and Environmental Microbiology*, 84(12), pp. e00270-18.
- Beck, A. and Liu, H. (2019). Macro- and Micro-Heterogeneity of Natural and Recombinant IgG Antibodies. *Antibodies*, 8(1), p.18.
- Benjamin, E.J., Muntner, P., Alonso, A., Bittencourt, M.S., Callaway, C.W. and Carson, A.P, *et al.*, (2019). Heart disease and stroke statistics—update: a report from the American Heart Association. *Circulation.* 2019;139(10): e56–528.
- Benoit, S., Benachour, A., Taouji, S., Auffray, Y. and Hartke, A. (2002) H₂O₂, which causes macrophage-related stress, triggers induction of expression of virulence-associated plasmid determinants in *Rhodococcus equi*. *Infection and Immunity* 70(7): 3768-3776.

- Benoit, S., Taouji, S., Benachour, A. and Hartke, A. (2000) Resistance of *Rhodococcus equi* to acid pH. *International Journal of Food Microbiology* 55(1-3): 295-298.
- Bernfield, M., Götte, M., Park, P., Reizes, O., Fitzgerald, M., Lincecum, J. and Zako, M. (1999). Functions of Cell Surface Heparan Sulfate Proteoglycans. *Annual Review of Biochemistry*, 68(1), pp.729-777.
- Bird, L. E. (2011). High throughput construction and small scale expression screening of multi-tag vectors in *Escherichia coli*. *Methods* 55, 29–37.
- Bischoff, M.F., and Rodwell, V.W. (1992); *Biochem. Med. Metab. Biology*. 48, 149.
- Blanpain, C., Buser, R., Power, A., et al (2001). A chimeric MIP-1 α /RANTES protein demonstrates the use of different regions of the RANTES protein to bind and activate its receptors. *J Leukoc Biol*; 69 pp977-85
- Bolten, S., Rinas, U. and Scheper, T., (2018). Heparin: role in protein purification and substitution with animal-component free material. *Applied Microbiology and Biotechnology*, 102(20), pp.8647-8660.
- Booth, W., Schlachter, C., Pote, S., Ussin, N., Mank, N., Klapper, V., Offermann, L., Tang, C., Hurlburt, B. and Chruszcz, M., 2018. Impact of an N-terminal Polyhistidine Tag on Protein Thermal Stability. *ACS Omega*, 3(1), pp.760-768.
- Bornhorst, J. A., and Falke, J. J. (2000). Purification of proteins using polyhistidine affinity tags. *Methods Enzymol.* 326, 245–254
- Bronowska A.K. (2011), Thermodynamics of ligand-protein interactions: Implications for molecular design. In: Moreno-Piraján J.C., editor.

- Thermodynamics—Interaction Studies—Solids, Liquids and Gases. InTech; Rijeka, Croatia: pp. 1–48.
- Bryne, G. A., Russel, D. A., Chen, X. and Meijer, W. G. (2007) Transcriptional regulation of the *virR* operon of the intracellular pathogen *Rhodococcus equi*. *Journal of Bacteriology* 189(14): 5082–5089.
- Bryne, G. A., Russel, D. A., Chen, X. and Meijer, W. G. (2007) Transcriptional regulation of the *virR* operon of the intracellular pathogen *Rhodococcus equi*. *Journal of Bacteriology* 189(14): 5082–5089
- Burton, A. J., Giguere, S., Sturgill, T. L., Berghaus, L. J., Slovis, N. M., Whitman, J. L., Levering, C., Kuskie, K. R. and Cohen, N. D. (2013) Macrolide- and rifampin-resistant *Rhodococcus equi* on a horse breeding farm, Kentucky, USA. *Emerging Infectious Diseases* 19(2): 282–285.
- Butt, T. R., Edavettal, S. C., Hall, J. P., and Mattern, M. R. (2005). SUMO fusion technology for difficult-to-express proteins. *Protein Expr. Purif.* 43, 1–9.
- Byrne, G.A., Boland, C.A., O’Connell, E.P. and Meijer, W.G. (2008) Differential mRNA stability of *vapAICD* operon of the facultative intracellular pathogen *Rhodococcus equi*. *FEMS Microbiology Letters* 280(1): 89–94.
- Cambien, B., Richard-Fiardo, P., Karimjee, B., Martini, V., Ferrua, B., Pitard, B., Schmid-Antomarchi, H. and Schmid-Alliana, A. (2011). CCL5 Neutralization Restricts Cancer Growth and Potentiates the Targeting of PDGFR β in Colorectal Carcinoma. *PLoS ONE*, 6(12), p.e28842.
- Cardona, A., Sasse, M., Liu, L., Cardona, S., Mizutani, M., Savarin, C., Hu, T. and Ransohoff, R. (2008). Scavenging roles of chemokine receptors: chemokine

receptor deficiency is associated with increased levels of ligand in circulation and tissues. *Blood*, 112(2), pp.256-263.

Cardoso, S. A., Oliveira, A. F., Ruas, L. P., Trevisani, M. M., De Oliveira, L. L., Hanna, E. S., Roque-Barreira, M. C. and Soares, S. G. (2013) Nasal vaccination with attenuated Salmonella expression VapA: TLR2 activation is not essential for protection against *R. equi* infection. *Vaccine*31(41): 4528-4535.

Carter, N., Ali, S. and Kirby, J., 2003. Endothelial inflammation: the role of differential expression of N-deacetylase/N-sulphotransferase enzymes in alteration of the immunological properties of heparan sulphate. *Journal of Cell Science*, 116(17), pp.3591-3600.

Caufield, J., Wimble, C., Shary, S., Wuchty, S. and Uetz, P. (2017). Bacterial protein meta-interactomes predict cross-species interactions and protein function. *BMC Bioinformatics*, 18(1).

Centers for Disease Control and Prevention. (2020). *Heart Attack* | *cdc.gov*. [online] Available at: https://www.cdc.gov/heartdisease/heart_attack.htm [Accessed 22 Feb. 2020].

Centers for Disease Control and Prevention. 2020. *Heart Attack* | *Cdc.Gov*. [online] Available at: https://www.cdc.gov/heartdisease/heart_attack.htm [Accessed 22 February 2020].

Centre for Disease Control (2017) *High Blood Pressure Fact Sheet*, Washington DC: CDC.

Chaires J.B. Calorimetry and thermodynamics in drug design. *Annu. Rev. Biophys.* 2008; 37:135–151. doi: 10.1146/annurev.biophys.36.040306.132812.

- Chasman D. I., Posada D., Subrahmanyam L., Cook N. R., Stanton V. P., Ridker P. M. (2004) 'Pharmacogenetic Study Of Statin Therapy And Cholesterol Reduction', *JAMA*, 291, pp. 2821-2827.
- Chatterjee, K., Nostramo, R., Wan, Y. and Hopper, A. (2018). tRNA dynamics between the nucleus, cytoplasm and mitochondrial surface: Location, location, location. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, 1861(4), pp.373-386.
- Chen, J., Almo, S. and Wu, Y. (2017). General principles of binding between cell surface receptors and multi-specific ligands: A computational study. *PLOS Computational Biology*, 13(10), p.e1005805.
- Chen, K., Bao, Z., Tang, P., Gong, W., Yoshimura, T. and Wang, J. (2018). Chemokines in homeostasis and diseases. *Cellular & Molecular Immunology*, 15(4), pp.324-334.
- Chen, Y., Chen, S., Li, J., Chen, P., Lee, Y., Li, K., Zarivach, R., Sun, Y. and Sue, S., 2019. Integrative model to coordinate the oligomerization and aggregation mechanisms of CCL5.
- Chen, Y., Song, J., Sui, S.F. and Wang, D.N. (2003) DNaK and DNaJ facilitated the folding process and reduced inclusion body formation of magnesium transporter CorA overexpressed in Escherichia coli. **Protein Expression Purif.** Vol 32 pp: 221–231.
- Chen, Y., Taton, A., Go, M., London, R., Pieper, L., Golden, S. and Golden, J. (2016). Self-replicating shuttle vectors based on pANS, a small endogenous plasmid of the unicellular cyanobacterium *Synechococcus elongatus* PCC 7942. *Microbiology*, 162(12), pp.2029-2041.

- Chen, Y.C., Chen, Y.D., Li, X., Post, W., Herrington, D., Polak, J.F., Rotter, J.I. and Taylor, K.D. (2009) The HMGCoA reductase gene and lipid and lipoprotein levels: the multi-ethnic study of atherosclerosis. **Lipids**; 44(8): pp733-43.
- Chi-Yen, L., Zhen, H., Wei, W., Andrew, W., Congzhou, Wang. and Li, Niu. (2015) Enhancing Protein Expression in HEK-293 Cells by Lowering Culture Temperature. *Journal.pone. PLoS ONE* 10(4): e0123562
- Cho, H.J; Lee, Y; Chang, R.S; Hahm, M.S, and Kim, M.K, et al. (2008). Maltose binding protein facilitates high-level expression and functional purification of the chemokines RANTES and SDF-1alpha from *Escherichia coli*. *Protein Expr Purif* 60: 37–45
- Choi, J. H., Keum, K. C., and Lee, S. Y. (2006). Production of recombinant proteins by high cell density culture of *Escherichia coli*. *Chem. Eng. Sci.* 61, 9.
- Cisek, A., Rzewuska, M., Witkowski, L. and Binek, M. (2014). Antimicrobial resistance in *Rhodococcus equi*. *Acta Biochimica Polonica*, 61(4).
- Clausen, T., Sandoval, D., Spliid, C., Pihl, J., Perrett, H., Painter, C., Narayanan, A., Majowicz, S., Kwong, E., McVicar, R., Thacker, B., Glass, C., Yang, Z., Torres, J., Golden, G., Bartels, P., Porell, R., Garretson, A., Laubach, L., Feldman, J., Yin, X., Pu, Y., Hauser, B., Caradonna, T., Kellman, B., Martino, C., Gordts, P., Chanda, S., Schmidt, A., Godula, K., Leibel, S., Jose, J., Corbett, K., Ward, A., Carlin, A. and Esko, J., 2020. SARS-CoV-2 Infection Depends on Cellular Heparan Sulfate and ACE2. *Cell*, 183(4), pp.1043-1057.e15.
- Cohen, N. (2015). *Rhodococcus equi* Foal Pneumonia. *Veterinary Clinics of North America: Equine Practice*, 30(3), pp.609-622.

- Cohen, N. D., O'Connor, M. S., Chaffin, M. K. and Martens, R. J. (2005) Farm characteristics and management processes associated with development of *Rhodococcus equi* pneumonia in foals. *Journal of American Veterinary Medical Association* 226(3): 404-413.
- Cooper-DeHoff Rhonda M, Johnson Julie A (2015). Hypertension pharmacogenomics: in search of personalized treatment approaches., *Nature reviews. Nephrology*
- Corsini, A., Maggi, F.M., Catapano, A.L (1995); *Pharmacol. Res.* 31, 9
- Costa, S. J., Almeida, A., Castro, A., Domingues, L., and Besir, H. (2013). The novel Fh8 and H fusion partners for soluble protein expression in *Escherichia coli*: a comparison with the traditional gene fusion technology. *Appl. Microbiol. Biotechnol.* 97, 6779–6791.
- Coulson, G. B., Agarwal, S. and Hondalus, M. K. (2010) Characterization of the role of the pathogenicity island and vapG in the virulence of the intracellular actinomycete pathogen *Rhodococcus equi*. *Infection and Immunity* 78(8): 3323-3334.
- Cox, J., Dean, R., Roberts, C. and Overall, C., 2020. Matrix Metalloproteinase Processing of CXCL11/I-TAC Results In Loss Of Chemoattractant Activity And Altered Glycosaminoglycan Binding.
- Cronin, C. (2019). High-volume shake flask cultures as an alternative to cellbag technology for recombinant protein production in the baculovirus expression vector system (BEVS). *Protein Expression and Purification*, p.105496.

- Cywes-Bentley, C., Rocha, J., Bordin, A., Vinacur, M., Rehman, S., Zaidi, T., Meyer, M., Anthony, S., Lambert, M., Vlock, D., Giguère, S., Cohen, N. and Pier, G., 2018. Antibody to Poly-N-acetyl glucosamine provides protection against intracellular pathogens: Mechanism of action and validation in horse foals challenged with *Rhodococcus equi*. *PLOS Pathogens*, 14(7), p.e1007160.
- Darrah, P. A., Hondalus, M. K., Chen, Q., Ischiropoulos, H. and Mosser, D. M. (2000) Cooperation between reactive oxygen and nitrogen intermediates in killing of *Rhodococcus equi* by activated macrophages. *Infection and Immunity* 68(6): 3587-3593.
- Dawson, T., Horohov, D., Meijer, W. and Muscatello, G., 2010. Current understanding of the equine immune response to *Rhodococcus equi*. An immunological review of *R. equi* pneumonia. *Veterinary Immunology and Immunopathology*, 135(1-2), pp.1-11.
- Day, A. and Prestwich, G. (2001). Hyaluronan-binding Proteins: Tying Up the Giant. *Journal of Biological Chemistry*, 277(7), pp.4585-4588.
- De Boer, H. A., Comstock, L. J., and Vasser, M. (1983). The tac promoter: a functional hybrid derived from the trp and lac promoters. *Proc. Natl. Acad. Sci. U.S.A.* 80, 21–25.
- De Marco, A. (2009). Strategies for successful recombinant expression of disulfide bond-dependent proteins in *Escherichia coli*. *Microbial Cell Factories*, 8(1), p.26.
- De Marco, A. (2015). Recombinant antibody production evolves into multiple options aimed at yielding reagents suitable for application-specific needs. *Microbial Cell Factories*, 14(1).

- Deshauer, C., Morgan, A., Ryan, E., Handel, T., Prestegard, J. and Wang, X., 2015. Interactions of the Chemokine CCL5/RANTES with Medium-Sized Chondroitin Sulfate Ligands. *Structure*, 23(6), pp.1066-1077.
- Dessen, A. and Xu, W. (2010). Structural biology of protein functional regulation. *Current Opinion in Structural Biology*, 20(6), pp.711-713.
- Dilworth, M., Piel, M., Bettaney, K., Ma, P., Luo, J., Sharples, D., Poyner, D., Gross, S., Moncoq, K., Henderson, P., Miroux, B. and Bill, R., 2018. Microbial expression systems for membrane proteins. *Methods*, 147, pp.3-39.
- Doolittle RF. (1983). Angiotensinogen is related to the antitrypsin-antithrombin-ovalbumin family. *Science*.;222:417-419.
- Drin, G., 2019. Intracellular Lipid Transport. *Methods in Molecular Biology*, 1.
- Du, X., Li, Y., Xia, Y., Ai, S., Liang, J., Sang, P., Ji, X. and Liu, S. (2016). Insights into Protein–Ligand Interactions: Mechanisms, Models, and Methods. *International Journal of Molecular Sciences*, 17(2), p.144.
- Duan, G. and Walther, D. (2015). The Roles of Post-translational Modifications in the Context of Protein Interaction Networks. *PLOS Computational Biology*, 11(2), p.e1004049.
- Dudley, Q., Karim, A. and Jewett, M. (2014). Cell-free metabolic engineering: Biomanufacturing beyond the cell. *Biotechnology Journal*, 10(1), pp.69-82.
- Duff, M.R. Jr. and Howell, E.E. Thermodynamics and solvent linkage of macromolecule-ligand interactions. *Methods*. 2014; 76:51–60.
- Dumon-Seignovert, L., Cariot, G. and Vuillard, L., 2004. The toxicity of recombinant proteins in *Escherichia coli*: a comparison of overexpression in BL21(DE3),

- C41(DE3), and C43(DE3). *Protein Expression and Purification*, 37(1), pp.203-206.
- Dyer, D., Migliorini, E., Salanga, C., Thakar, D., Handel, T. and Richter, R. (2017). Differential structural remodelling of heparan sulfate by chemokines: the role of chemokine oligomerization. *Open Biology*, 7(1), p.160286.
- Dyer, D., Salanga, C., Volkman, B., Kawamura, T. and Handel, T. (2016). The dependence of chemokine–glycosaminoglycan interactions on chemokine oligomerization. *Glycobiology*, 26(3), pp.312-326 *Protein Science*, 80(1).
- Elissalde, G. S., Renshaw, H. W. and Walberg, J. A. (1980) *Corynebacterium equi*: an interhost review with emphasis on the foal. *Comparative Immunology, Microbiology and Infectious Diseases* 3(4): 433–445.
- Embl.de. (2020). *Protein Purification - Extraction and Clarification - Removal of DNA* - EMBL. [online] Available at: https://www.embl.de/pepcore/pepcore_services/protein_purification/extraction_clarification/removal_dna/ [Accessed 9 Feb. 2020].
- Endo, A., Kuroda, M. and Tsujita, Y. (1976) ML-236A, ML-236B, and ML-236C, new inhibitors of cholesterol synthesis produced by *Penicillium citrinum*. *Journal of Antibiotics*. 29 (12): 1346–8.
- Esko, J. and Lindahl, U. (2001). Molecular diversity of heparan sulfate. *Journal of Clinical Investigation*, 108(2), pp.169-173.
- Eswarakumar, V., Lax, I. and Schlessinger, J. (2005). Cellular signaling by fibroblast growth factor receptors. *Cytokine & Growth Factor Reviews*, 16(2), pp.139-149.

- Fan, H., Conn, A., Williams, P., Diggs, S., Hahm, J., Gamper, H., Hou, Y., O'Leary, S., Wang, Y. and Blaha, G. (2017). Transcription–translation coupling: direct interactions of RNA polymerase with ribosomes and ribosomal subunits. *Nucleic Acids Research*, 45(19), pp.11043-11055.
- Farrugia, B., Lord, M., Melrose, J. and Whitelock, J., 2018. The Role of Heparan Sulfate in Inflammation, and the Development of Biomimetics as Anti-Inflammatory Strategies. *Journal of Histochemistry & Cytochemistry*, 66(4), pp.321-336.
- Fee, C.J, Fredericks-Short F, Billikanti JM, Damodaran VB (22–26 Jun 2008). Measurement of Electrostatic Interactions of PEGylated Proteins Using a Novel Multi-Channel Surface Plasmon Resonance Technique. Recovery of Biological Products XIII. Quebec City, Quebec. Archived from the original on 2010-11-04.
- Fermas, S., Gonnet, F., Sutton, A., Charnaux, N., Mulloy, B., Du, Y., Baleux, F., Daniel, R. (2008) Sulfated oligosaccharides (heparin and fucoidan) binding and dimerization of stromal cell-derived factor-1(SDF-1/CXCL 12) are coupled as evidenced by affinity CE-MS analysis. *Glycobiology* 18,1054–1064.75.
- Fernandez-Mora, E., Polidori, M., Luhrmann, A., Schaible, U. E. and Haas, A. (2005) Maturation of *Rhodococcus equi*-containing vacuoles is arrested after completion of the early endosome stage. *Traffic* 6(8): 635-653.
- Fox, J., Tyler, R., Peterson, F., Dyer, D., Zhang, F., Linhardt, R., Handel, T. and Volkman, B., 2016. Examination of Glycosaminoglycan Binding Sites on the XCL1 Dimer. *Biochemistry*, 55(8), pp.1214-1225.

- Frangioni, J.V. and Neel, B.G. (1993) Solubilisation and purification of enzymatically active glutathione s-transferase (pGEX) fusion proteins. **Anal Biochem** 210:179-87.
- Frauschuh, A. *et al.* (2007) "Molecular cloning and characterization of a highly selective chemokine-binding protein from the tick *Rhipicephalus sanguineus*," *Journal of Biological Chemistry*, 282(37), pp. 27250–27258. Available at: <https://doi.org/10.1074/jbc.m704706200>.
- Friesen, J. F and Rodwell, V. W. (2004) 'The 3-Hydroxy-3-Methylglutaryl Coenzyme-A (HMG-Coa) eductases ', *Genome Biology*, 5, pp. 248.
- Gallagher, J.T., Walker, A. (1985). Molecular distinctions between heparan sulphate and heparin. Analysis of sulphation patterns indicates that heparan sulphate and heparin are separate families of *N*-sulphated polysaccharides. *Biochem. J.* 230, 665–674
- Garratt R.C., Valadares N.F., Bacheга J.F.R. (2013) Oligomeric Proteins. In: Roberts G.C.K. (eds) *Encyclopedia of Biophysics*. Springer, Berlin, Heidelberg
- Gavins, F., Yilmaz, G. and Granger, D. (2007). The Evolving Paradigm for Blood Cell-Endothelial Cell Interactions in the Cerebral Microcirculation. *Microcirculation*, 14(7), pp.667-681.
- Ge, B., Li, J., Wei, Z. *et al.* Functional expression of CCL8 and its interaction with chemokine receptor CCR3. *BMC Immunol* **18**, 54 (2017). <https://doi.org/10.1186/s12865-017-0237-5>
- Geddes, C. (2016). *REVIEWS IN PLASMONICS 2015*. New York: SPRINGER INTERNATIONAL PU, p.316.

- Genick, C., Barlier, D., Monna, D., Brunner, R., Bé, C., Scheufler, C. and Ottl, J., 2014. Applications of Biophysics in High-Throughput Screening Hit Validation. *Journal of Biomolecular Screening*, 19(5), pp.707-714.
- Giguère, S., Cohen, N., Keith Chaffin, M., Hines, S., Hondalus, M., Prescott, J. and Slovis, N. (2012). Rhodococcus equi: Clinical Manifestations, Virulence, and Immunity. *Journal of Veterinary Internal Medicine*, 25(6), pp.1221-1230.
- Giles, C., Ndi, O., Barton, M. and Vanniasinkam, T., 2016. An Adenoviral Vector Based Vaccine for Rhodococcus equi. *PLOS ONE*, 11(3), p.e0152149.
- Giles, C., Vanniasinkam, T., Ndi, S. and Barton, M., 2015. Rhodococcus equi (Prescottella equi) vaccines; the future of vaccine development. *Equine Veterinary Journal*, 47(5), pp.510-518.
- Goenka, S. and Rao, C.M. (2001) Expression of recombinant zeta-crystallin in Escherichia coli with the help of GroEL/ES and its purification. **Protein Expression Purif.**; Vol 21: pp 260–267.
- Golemis E (2002) *Protein-protein interactions: A molecular cloning manual*. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press. p ix, 682.
- Golub, B. (1967). Lung Abscess Due to Corynebacterium equi. *Annals of Internal Medicine*, 66(6), p.1174.
- Gotte, M. Syndecans in inflammation. *FASEB J.* 2003, 17, 575–591 Theoharides, T.C.; Alysandratos, K.-D.; Angelidou, A.; Delivanis, D.-A.; Sismanopoulos, N.; Zhang, B.; Asadi, S.; Vasiadi, M.; Weng, Z.; Miniati, A.; et al. Mast cells and inflammation. *Biochim. Biophys. Acta* 2012, 1822, 21–33.

- Graf, C., Stankiewicz, M., Kramer, G. and Mayer, M. (2009). Spatially and kinetically resolved changes in the conformational dynamics of the Hsp90 chaperone machine. *The EMBO Journal*, 28(5), pp.602-613.
- Grand, E., Kovensky, J., Pourceau, G., Toumieux, S. and Wadouachi, A. (2014). ChemInform Abstract: Anionic Oligosaccharides: Synthesis and Applications. *ChemInform*, 45(38), p.195-235
- Granger, D. and Kubes, P. (1994). The microcirculation and inflammation: modulation of leukocyte-endothelial cell adhesion. *Journal of Leukocyte Biology*, 55(5), pp.662-675.
- Graslund, S. et al., (2008). Protein production and purification. **Nat. Methods** 5, 135–146.
- Gräslund, S.; Nordlund, P.; Weigelt, J.; Hallberg, B. M.; Bray, J.; Gileadi, O.; Knapp, S.; Oppermann, U.; Arrowsmith, C.; Hui, R.; Ming, J.; dhe-Paganon, S.; Park, H. W.; Savchenko, A.; Yee, A.; Edwards, A.; Vincentelli, R.; Cambillau, C.; Kim, R.; Kim, S. H.; Rao, Z.; Shi, Y.; Terwilliger, T. C.; Kim, C. Y.; Hung, L. W.; Waldo, G. S.; Peleg, Y.; Albeck, S.; Unger, T.; Dym, O.; Prilusky, J.; Sussman, J. L.; Stevens, R. C.; Lesley, S. A.; Wilson, I. A.; Joachimiak, A.; Collart, F.; Dementieva, I.; Donnelly, M. I.; Eschenfeldt, W. H.; Kim, Y.; Stols, L.; Wu, R.; Zhou, M.; Burley, S. K.; Emtage, J. S.; Sauder, J. M.; Thompson, D.; Bain, K.; Luz, J.; Gheyi, T.; Zhang, F.; Atwell, S.; Almo, S. C.; Bonanno, J. B.; Fiser, A.; Swaminathan, S.; Studier, F. W.; Chance, M. R.; Sali, A.; Acton, T. B.; Xiao, R.; Zhao, L.; Ma, L. C.; Hunt, J. F.; Tong, L.; Cunningham, K.; Inouye, M.; Anderson, S.; Janjua, H.; Shastry, R.; Ho, C. K.; Wang, D.; Wang, H.; Jiang, M.; Montelione, G. T.; Stuart, D. I.; Owens, R. J.; Daenke, S.; Schutz, A.;

- Heinemann, U.; Yokoyama, S.; Bussow, K.; Gunsalus, K. C. Protein production and purification *Nat. Methods* **2008**, 5, 135–146
- Grodberg, J., Dunn J.J. (1988) ompT encodes the *Escherichia coli* outer membrane protease that cleaves T7 RNA polymerase during purification. *J.Bacteriol.* 170:1245-1253
- Guan, C. et al. (1987) *Gene* 67, 21–30.
- Gundelly, P. (2015). *Rhodococcus equi* in Immunocompromised and Immunocompetent Individuals and Antibiotic Susceptibility of Clinical Isolates: A Case Series of 12 Patients with Literature Review. *Open Forum Infectious Diseases*, 2(suppl_1).
- Gupta, S. K., Dangi, A. K., Smita, M., and Dwivedi, S. (2019). “Effectual bioprocess development for protein production,” in *Applied Microbiology and Bioengineering*, ed P. Shukla (London: Academic Press), 203–227.
- Gurel, V., Lambert, K., Page, A. E., Loynachan, A. T., Huges, K., Timoney, J. F., Fettinger, M., Horohov, D. W. and McMichael, J. (2013) Streptolysin-O/antibiotics adjunct therapy modulates site-specific expression of extracellular matrix and inflammatory genes in lungs of *Rhodococcus equi* infected foals. *Veterinary Research Communications* 37(2): 145-154.
- Hammarstrom, M., Hellgren, N., Van Den Berg, S., Berglund, H., and Hard, T. (2002). Rapid screening for improved solubility of small human proteins produced as fusion proteins in *Escherichia coli*. *Protein Sci.* 11, 313–321.
- Hammarstrom, M., Woestenenk, E. A., Hellgren, N., Hard, T., and Berglund, H. (2006). Effect of N-terminal solubility enhancing fusion proteins on yield of purified target protein. *J. Struct. Funct. Genomics* 7, 1–14.

- Handel TM, Johnson Z, Crown SE, Lau EK, Sweeney M, Proudfoot AE. Regulation of protein function by glycosaminoglycans—as exemplified by chemokines. *Annu Rev Biochem.* 2005; 74:385–410.
- Harrow, J., Nagy, A., Reymond, A. *et al.* Identifying protein-coding genes in genomic sequences. *Genome Biol* **10**, 201 (2009). <https://doi.org/10.1186/gb-2009-10-1-201>
- Hay, E. (1991). *Cell biology of extracellular matrix*. 2nd ed. New York: Springer US.
- Hayashi K, Morooka N, Yamamoto Y, Fujita K, Isono K, Choi S, Ohtsubo E, Baba T, Wanner BL, Mori H, and Horiuchi T. Highly accurate genome sequences of *Escherichia coli* K-12 strains MG1655 and W3110. *Mol Syst Biol.* 2006; 2:2006.0007.
- Hebert, L., Bidaud, P., Goux, D., Benachour, A., Laugier, C. and Petry, S. (2014) Study of lysozyme resistance in *Rhodococcus equi*. *Current Microbiology* 68(3): 352-357.
- Hebron, H. R., Yang, Y., & Hang, J. (2009). Purification of genomic DNA with minimal contamination of proteins. *Journal of biomolecular techniques: JBT*, 20(5), 278–281.
- Hickey, M.J; Kanwar, S; McCafferty, D.M; Granger, D.N, Eppihimer, M.J, and Kubes, P. (1999) Varying Roles of E-selectin and P-selectin in different microvascular beds in response to antigen. *J Immunol.*; 162: 1137–43.
- Hines, M. T., Paasch, K. M., Alperin, D. C., Palmer, G. H., Westhoff, N. C. and Hines, S. A. (2001) Immunity to *Rhodococcus equi*: antigen-specific recall responses in the lungs of adult horses. *Veterinary Immunology and Immunopathology* 79(1-2): 101–113.

- Hines, S. A., Stone, D. M., Hines, M. T., Alperin, D. C., Knowles, D. P., Norton, L. K., Hamilton, M. J., Davis, W. C. And McGuire, T. C. (2003) Clearance of virulent but not avirulent *Rhodococcus equi* from the lungs of adult horses is associated with intracytoplasmic gamma interferon production by CD4+ and CD8+T Lymphocytes. *Clinical and Diagnostic Laboratory Immunology* 10(2): 208-215.
- Hirayama, D., Iida, T. and Nakase, H., 2017. The Phagocytic Function of Macrophage-Enforcing Innate Immunity and Tissue Homeostasis. *International Journal of Molecular Sciences*, 19(1), p.92.
- Hofmann, A., 2015. *Methods In Structural Chemistry*. 2nd ed. Australia: The Printing Hub, 3/1 Colemans Road, Carrum Downs, Vic 3201, p.43.
- Hondalus, M. K. (1997) Pathogenesis and virulence of *Rhodococcus equi*. *Veterinary Microbiology* 56(3-4): 257-268.
- Hondalus, M. K. and Mosser, D. M. (1994) Survival and replication of *Rhodococcus equi* in Macrophages. *Infection and Immunity* 62(10): 4167-4175.
- Hoogewerf, A.J. Kuschert, G.S. Proudfoot, A.E. *et al.* (1997), Glycosaminoglycans mediate cell surface oligomerization of chemokines *Biochemistry*, 36 pp. 13570-13578.
- Hopp, T. P., Prickett, K. S., Price, V. L., Libby, R. T., March, C. J., Pat Cerretti, D., *et al.* (1988). A short polypeptide marker sequence useful for recombinant protein identification and purification. *Nat. Biotechnol.* 6, 1204–1210.
- Horohov, D. W., Loynachan, A. T., Page, A. E., Hughes, K., Timoney, J. F., Fettingner, M., Hatch, T., Spaulding, J. G. and McMichael J. (2011) The use of streptolysin

- O (SLO) as an adjunct therapy for *Rhodococcus equipneumonia* in foals. *Veterinary Microbiology* 154(1-2): 156-162.
- Ignatov, A., Robert, J., Gregory-Evans, C. and Schaller, H. (2006). RANTES stimulates Ca²⁺ mobilization and inositol trisphosphate (IP₃) formation in cells transfected with G protein-coupled receptor 75. *British Journal of Pharmacology*, 149(5), pp.490-497.
- Iizasa, E. and Nagano, Y. (2018). Highly efficient yeast-based in vivo DNA cloning of multiple DNA fragments and the simultaneous construction of yeast/*Escherichia coli* shuttle vectors. *BioTechniques*, 40(1), pp.79-83.
- Imberty A, Lortat-Jacob H, Pérez S. Structural view of glycosaminoglycan-protein interactions. *Carbohydr Res.* 2007; 342:430–439.
- Istvan, E. S, Deisenhofer, J. (2001) 'Structural Mechanism for Statin Inhibition of HMG-Coa Reductase. *Science*', 292, pp. 1160-1164.
- Istvan, E. S. (2001) Bacterial and Mammalian HMG-CoA Reductases: Related Enzymes with Distinct Architectures. *Curr. Opin. Struc. Biol.*11, 746–751.
- Istvan, E.S. and Deisenhofer, J. (2003) 'Structural Mechanism for Statin Inhibition of HMG-Coa Reductase. *Science*', 292, pp. 1160-1164.
- Istvan, E.S., Palnitkar, M., Buchanan, S.K. and Deisenhofer, J. (2000) Crystal structure of the catalytic portion of human HMG-CoA reductase: insights into regulation of activity and catalysis: **EMBO J.** 2000 Mar 1; 19(5): 819–830.
- Jacks, S., Giguere, S. and Nguyen, A. (2003). In Vitro Susceptibilities of *Rhodococcus equi* and Other Common Equine Pathogens to Azithromycin,

- Clarithromycin, and 20 Other Antimicrobials. *Antimicrobial Agents and Chemotherapy*, 47(5), pp.1742-1745.
- Jacks, S., Giguere, S., Crawford, P. and Castleman, W. (2007). Experimental Infection of Neonatal Foals with *Rhodococcus equi* Triggers Adult-Like Gamma Interferon Induction. *Clinical and Vaccine Immunology*, 14(6), pp.669-677.
- Jackson, R., Busch, S., Cardin, A. (1991) Glycosaminoglycans: Molecular Properties, Protein Interactions, and Role in Physiological Processes *Physiol. Rev.*, 71, pp.481– 539
- Jahanshahi, P., Zalnezhad, E., Sekaran, S. *et al.* Rapid Immunoglobulin M-Based Dengue Diagnostic Test Using Surface Plasmon Resonance Biosensor. *Sci Rep* 4, 3851 (2014). <https://doi.org/10.1038/srep03851>
- Jain, S., Bloom, B. R. and Hondalus, M. K. (2003) Deletion of *vapA* encoding virulence associated protein-A attenuates the intracellular actinomycete *Rhodococcus equi*. *Molecular Microbiology* 50(1): 115-128.
- Jansen, J-C. (Editor). (2011). Protein Purification: Principles, High Resolution Methods, and Applications. 3rd edition. Volume 151 of Methods of Biochemical Analysis. John Wiley & Sons, Hoboken, NJ
- Jayanthi, S., Gundampati, R. and Kumar, T., (2018). Simple and Efficient Purification of Recombinant Proteins Using the Heparin-Binding Affinity Tag. *Current Protocols in Protein Science*, 90(1).
- Kalinowski, M., Grądziński, Z., Jarosz, Ł., Kato, K., Hieda, Y., Kakuda, T. and Takai, S. (2016). Plasmid Profiles of Virulent *Rhodococcus equi* Strains Isolated from Infected Foals in Poland. *PLOS ONE*, 11(4), p.e0152887.

- Kanaly, S. T., Hines, S. A. and Palmer, G. H. (1995) Cytokine modulation alters pulmonary clearance of *Rhodococcus equi* and development of granulomatous pneumonia. *Infection and Immunity* 63(8): 3037-3041.
- Kapust, R. B., and Waugh, D. S. (1999). *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci.* 8, 1668–1674.
- Karav, S., Talak, E., Tuncer, M. and Ozleyen, A., 2017. The Effect of Fusion Tags on Enzyme Specificity and Protein Purification Efficiency. *International Journal of Agriculture Innovations and Research*, 6(3), pp.462-463.
- Kastritis, P. and Bonvin, A. (2013). On the binding affinity of macromolecular interactions: daring to ask why proteins interact. *Journal of The Royal Society Interface*, 10(79), p.20120835.
- Khan, F., Legler, P. M., Mease, R. M., Duncan, E. H., Bergmann-Leitner, E. S., and Angov, E. (2012). Histidine affinity tags affect MSP1(42) structural stability and immunodominance in mice. *Biotechnol. J.* 7, 133–147.
- Kilcoyne, I., Nieto, J. and Vaughan, B. (2013). Tibial osteomyelitis caused by *Rhodococcus equi* in a mature horse. *Equine Veterinary Education*, 26(6), pp.283-286.
- Kim, S. and Misra, A., 2007. SNP Genotyping: Technologies and Biomedical Applications. *Annual Review of Biomedical Engineering*, 9(1), pp.289-320.
- Klug, W., Cummings, M., Spencer, C. and Palladino, M. (2016). *Essentials of genetics*. 9th ed. Edinburgh, England: Pearson, pp.296-297.
- Koike, G., Horiuchi, M., Yamada, T., Szpirer, C., Jacob, H. J. and Dzau, V. J. (1994) Human type 2 angiotensin II receptor gene: cloned, mapped to the X

chromosome, and its mRNA is expressed in the human lung. *Biochem. Biophys. Res. Commun.* 203: 1842-1850.

Koopmann, W. Ediriwickrema, C., Krangel M.S. (1999), Structure and function of the glycosaminoglycan binding site of chemokine macrophage-inflammatory protein-1 beta *J. Immunol.*, 163 pp. 2120-2127

Korpimäki, T., Kurittu, J., and Karp, M. (2003). Surprisingly fast disappearance of beta-lactam selection pressure in cultivation as detected with novel biosensing approaches. *J. Microbiol. Methods* 53, 37–42.

Köwitsch, A., Zhou, G. and Groth, T. (2018). Medical application of glycosaminoglycans: a review. *Journal of Tissue Engineering and Regenerative Medicine*, 12(1), pp. e23-e41.

Krauss, R.M., Mangravite, L.M., Smith, J.D., Medina, M.W., Wang, D., Guo, X., Rieder, M.J., Simon, J.A., Hulley, S.B., Waters, D., Saad, M., Williams, P.T., Taylor, K.D., Yang, H., Nickerson, D.A. and Rotter, J.I. (2008) Variation in the 3-hydroxyl-3-methylglutaryl coenzyme a reductase gene is associated with racial differences in low-density lipoprotein cholesterol response to simvastatin treatment. *Circulation*: 117(12):1537-44.

Krieger, E., Geretti, E., Brandner, B., Goger, B., Wells, T. N., Kungl, A. J. (2004) A structural and dynamic model for the interaction of interleukin-8 and glycosaminoglycans: support from isothermal fluorescence titrations. *Proteins* 54, 768–775.76.

Kumar (2019). Macrophages: The Potent Immunoregulatory Innate Immune Cells, *Macrophage Activation - Biology and Disease*, Khalid Hussain Bhat, IntechOpen, DOI: 10.5772/intechopen.88013. Available from:

- <https://www.intechopen.com/books/macrophage-activation-biology-and-disease/macrophages-the-potent-immunoregulatory-innate-immune-cells>
- Kunji, E., Ruprecht, J., King, M., Zögg, T., Aleksandrova, A., Pardon, E., Crichton, P. and Steyaert, J., 2018. The structural mechanism of transport by the mitochondrial ADP/ATP carrier. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 1859, pp. e31-e32.
- Ladbury J.E. Calorimetry as a tool for understanding biomolecular interactions and an aid to drug design. *Methods Mol. Biol.* 2015; 1278:223–232.
- Laguri, C., Arenzana-Seisdedos, F. and Lortat-Jacob, H., (2008). Relationships between glycosaminoglycan and receptor binding sites in chemokines-the CXCL12 example *Carbohydr. Res.*, 343, pp. 2018-2023
- Laguri, C., Sadir, R. and Rueda, P. *et al.* (2007). The novel CXCL12gamma isoform encodes an unstructured cationic domain which regulates bioactivity and interaction with both glycosaminoglycans and CXCR4 *PLoS ONE*, 2, p. e1110.
- Lawrence, M. and Springer, T. (1991). Leukocytes roll on a selectin at physiologic flow rates: Distinction from and prerequisite for adhesion through integrins. *Cell*, 65(5), pp.859-873.
- Lazard, D., Briend-Sutren, M. M., Villageois, P., Mattei, M.-G., Strosberg, A. D. and Nahmias, C. (1994) Molecular characterization and chromosome localization of a human angiotensin II AT2 receptor gene highly expressed in fetal tissues. *Receptors Channels* 2: 271-280.
- Letek, M., González, P., MacArthur, I., Rodríguez, H., Freeman, T. C., Valero-Rello, A., Blanco, M., Buckley, T., Cherevach, I., Fahey, R., Hapeshi, A., Holdstock, J., Leadon, D., Navas, J., Ocampo, A., Quail, M. A., Sanders, M., Scotti, M.

- M., Prescott, J.F., Fogarty, U., Meijer, W. G., Parkhill, J., Bentley, D. and Vázquez-Boland, J. A. (2010) The genome of a Pathogenic *Rhodococcus*: Cooptive virulence underpinned by key gene acquisitions. *Plos Genetics* 6(9): e1001145.
- Letek, M., Ocampo-Sosa, A., Sanders, M., Fogarty, U., Buckley, T., Leadon, D., Gonzalez, P., Scotti, M., Meijer, W., Parkhill, J., Bentley, S. and Vazquez-Boland, J. (2008). Evolution of the *Rhodococcus equi* vap Pathogenicity Island Seen through Comparison of Host-Associated vapA and vapB Virulence Plasmids. *Journal of Bacteriology*, 190(17), pp.5797-5805.
- Letek, M., Ocampo-Sosa, A., Sanders, M., Fogarty, U., Buckley, T., Leadon, D., Gonzalez, P., Scotti, M., Meijer, W., Parkhill, J., Bentley, S. and Vazquez-Boland, J., 2008. Evolution of the *Rhodococcus equi* vap Pathogenicity Island Seen through Comparison of Host-Associated vapA and vapB Virulence Plasmids. *Journal of Bacteriology*, 190(17), pp.5797-5805.
- Leung, K., Dacks, J. and Field, M., 2008. Evolution of the Multivesicular Body ESCRT Machinery; Retention Across the Eukaryotic Lineage. *Traffic*, 9(10), pp.1698-1716.
- Levy, J. (2009). The Unexpected Pleiotropic Activities of RANTES. *The Journal of Immunology*, 182(7), pp.3945-3946.
- Li, Jin-Ye, Yi-Chen Chen, Yi-Zong Lee, Chun-Hsiang Huang, and Shih-Che Sue. 2020. "N-terminal Backbone Pairing Shifts in CCL5-¹²AAA¹⁴ Dimer Interface: Structural Significance of the FAY Sequence" *International Journal of Molecular Sciences* 21, no. 5: 1689. <https://doi.org/10.3390/ijms21051689>

- Liang, W.G., Triandafillou, C.G., Huang, T.Y., Zulueta, M.M., Banerjee, S., Dinner, A.R., Hung, S.C., Tang, W.J. (2016), Crystal structure of CC chemokine 5 (CCL5). *Proc. Natl. Acad. Sci. USA* 113, 5000–5005
- Lindahl, U., Lidholt, K., Spillmann, D. and Kjellén, L. (1994). More to “heparin” than anticoagulation. *Thrombosis Research*, 75(1), pp.1-32.
- Linder, R. and Bernheimer, A. W. (1997) Oxidation of macrophage membrane cholesterol by intracellular *Rhodococcus equi*. *Veterinary Microbiology* 56(3-4): 269-276. Linder, R. and Bernheimer, A. W. (1997) Oxidation of macrophage membrane cholesterol by intracellular *Rhodococcus equi*. *Veterinary Microbiology* 56(3-4): 269-276.
- Lindman, S., Xue, W., Szczepankiewicz, O., Bauer, M., Nilsson, H. and Linse, S., 2006. Salting the Charged Surface: pH and Salt Dependence of Protein G B1 Stability. *Biophysical Journal*, 90(8), pp.2911-2921.
- Liu, F., Zhang, J. and Mei, Y., 2016. The origin of the cooperativity in the streptavidin-biotin system: A computational investigation through molecular dynamics simulations. *Scientific Reports*, 6(1).
- Liu, M., Wang, B., Wang, F., Yang, Z., Gao, D., Zhang, C., et al. (2019). Soluble expression of single-chain variable fragment (scFv) in *Escherichia coli* using superfolder green fluorescent protein as fusion partner. *Appl. Microbiol. Biotechnol.* 103, 6071–6079.
- Liu, Y., Zhao, T.J., Yan, Y.B. and Zhou, H.M. (2005) Increase of soluble expression in *Escherichia coli* cytoplasm by a protein disulfide isomerase gene fusion system. *Protein Expression Purif.*; 44:155–161.

- MacArthur, I., Parreira, V. R., Lepp, D., Mutharia, L. M., Vazquez-Boland, J. A. and Prescott, J. F. (2011) The sensor kinase MprB is required for *Rhodococcus equi* virulence. *Veterinary Microbiology* 147(1-2): 133-141.
- Magnusson, H. (1923). Spezifische infektiöse Pneumonie beim Fohlen. Ein neuer Eiterreger beim Pferd *Archiv Wissenschaft Praktit-Tierheilkunde* 50: 22-38.
- Maillard, L., Saito, N., Hlawaty, H., Friand, V., Suffee, N., Chmilewsky, F., Haddad, O., Laguillier, C., Guyot, E., Ueyama, T., Oudar, O., Sutton, A. and Charnaux, N., 2014. RANTES/CCL5 mediated-biological effects depend on the syndecan-4/PKC signaling pathway. *Biology Open*, 3(10), pp.995-1004.
- Makoff, A. J., and Oxer, M. D. (1991). High level heterologous expression in *E. coli* using mutant forms of the lac promoter. *Nucleic Acids Res.* 19, 2417–2421.
- Malekian, R., Sima, S., Jahanian-Najafabadi, A., Moazen, F., and Akbari, V. (2019). Improvement of soluble expression of GM-CSF in the cytoplasm of *Escherichia coli* using chemical and molecular chaperones. *Protein Expr. Purif.* 160, 66–72.
- Mango, R., Biocca, S., Del Vecchio, F., Clementi, F., Sangiuolo, F., Amati, F., Filareto, A., Grelli, S., Spitalieri, P., Filesi, I., Favalli, C., Lauro, R., Mehta, J. L, Romeo, F., Novelli, G. (2005) 'In Vivo And In Vitro Studies Support', *That A New Splicing Isoform of OLR1 Gene Is Protective Against Acute Myocardial Infarction*, 97, pp. 152-158.
- Marisa, W. M., and Ronald, Krauss, M. (2009) 'The Role of HMGR Alternative Splicing in Statin Efficacy Trends', *Cardiovasc Med.*, 19(5), pp. 173-177.
- Maron, B.J., Gardin, J.M., Flack, J.M., Gidding, S.S., Kurosaki, T.T. and Bild, D.E. (1995); Prevalence of hypertrophic cardiomyopathy in a general population

of young adults: echocardiographic analysis of 4111 subjects in the CARDIA study. *Circulation*.92:785–789

Martin, L., Blanpain, C., Garnier, P., Wittamer, V., Parmentier, M. and Vita, C. (2001). Structural and Functional Analysis of the RANTES-Glycosaminoglycans Interactions. *Biochemistry*, 40(21), pp.6303-6318.

Martin, S. and Clements, J. (2013). Correlating Structure and Energetics in Protein-Ligand Interactions: Paradigms and Paradoxes. *Annual Review of Biochemistry*, 82(1), pp.267-293.

Martínez-Burgo, B., Cobb, S., Pohl, E., Kashanin, D., Paul, T., Kirby, J., Sheerin, N. and Ali, S. (2019). A C-terminal CXCL 8 peptide based on chemokine–glycosaminoglycan interactions reduce neutrophil adhesion and migration during inflammation. *Immunology*, 157(2), pp.173-184.

Martínez-Solís, M., Herrero, S. and Targovnik, A. (2018). Engineering of the baculovirus expression system for optimized protein production. *Applied Microbiology and Biotechnology*, 103(1), pp.113-123.

Mboge, M., Kota, A., McKenna, R. and Frost, S., 2018. Biophysical, Biochemical, and Cell Based Approaches Used to Decipher the Role of Carbonic Anhydrases in Cancer and to Evaluate the Potency of Targeted Inhibitors. *International Journal of Medicinal Chemistry*, 2018, pp.1-18.

McCullough, P.A., Chan, C.T., Weinhandl, E.D., Burkart, J.M. and Bakris, G.L. (2016); Intensive Hemodialysis, Left Ventricular Hypertrophy, and Cardiovascular Disease. *Am J Kidney Dis*. 68(5S1):S5-S14.

McDermott, M.M. (2007); The international pandemic of chronic cardiovascular disease.

- McNaughton, E., Eustace, A., King, S., Sessions, R., Kay, A., Farris, M., Broadbridge, R., Kehoe, O., Kungl, A. and Middleton, J., 2018. Novel Anti-Inflammatory Peptides Based on Chemokine–Glycosaminoglycan Interactions Reduce Leukocyte Migration and Disease Severity in a Model of Rheumatoid Arthritis. *The Journal of Immunology*, 200(9), pp.3201-3217.
- Medina, M.W., Gao, F., Ruan, W., Rotter, J.I, Krauss, R.M., (2008) 'Alternative Splicing Of 3-Hydroxy-3-Methylglutaryl Coenzyme A Reductase Is Associated with Plasma Low- Density Lipoprotein Cholesterol Response To Simvastatin.', *Circulation*, 118, pp. 355-362.
- Mehta, P.K., Wei, J. and Wenger, N.K (2014). "Ischemic heart disease in women: A focus on risk factors.". *Trends in Cardiovascular Medicine*. **25** (2): 140–151.
- Meijer, W. and Prescott, J. (2004). *Rhodococcus equi*. *Veterinary Research*, 35(4), pp.383-396.
- Mendis, S., Puska, P., Norrving, B. (2011) *Global Atlas on Cardiovascular Disease Prevention and Control*, Geneva: World Health Organization Organization in collaboration with the World Heart Federation and the World Stroke Organization. pp. 3–18.
- Mensah, G., Roth, G. and Fuster, V. (2019). The Global Burden of Cardiovascular Diseases and Risk Factors. *Journal of the American College of Cardiology*, 74(20), pp.2529-2532.
- Metzemaekers, M., Vanheule, V., Janssens, R., Struyf, S. and Proost, P. (2018). Overview of the Mechanisms that May Contribute to the Non-Redundant Activities of Interferon-Inducible CXC Chemokine Receptor 3 Ligands. *Frontiers in Immunology*, 8.

- Mikic, D. (2006). Rhodococcus equi infection. *Vojnosanitetski pregled*, 63(11), pp.957-962.
- Miroux, B. and Walker, J., 1996. Over-production of Proteins in Escherichia coli: Mutant Hosts that Allow Synthesis of some Membrane Proteins and Globular Proteins at High Levels. *Journal of Molecular Biology*, 260(3), pp.289-298.
- Mizumoto, S., Fongmoon, D., Sugahara, K. (2013) Interaction of chondroitin sulfate and dermatan sulfate from various biological sources with heparin-binding growth factors and cytokines. *Glycoconj. J.* 30, 619–632. 77.
- Moffatt, B. A., and Studier, F. W. (1987). T7 lysozyme inhibits transcription by T7 RNA polymerase. *Cell* 49, 221–227.
- Moghadasian M. (1999); Clinical pharmacology of 3-hydroxy-3-methylglutaryl coenzyme A reductase inhibitors. *Life Sci* 65: 1329-37
- Mohanty, S., Jagannathan, L., Ganguli, G., Padhi, A., Roy, D., Alaridah, N., Saha, P., Nongthomba, U., Godaly, G., Gopal, R., Banerjee, S. and Sonawane, A. (2015). A Mycobacterial Phosphoribosyltransferase Promotes Bacillary Survival by Inhibiting Oxidative Stress and Autophagy Pathways in Macrophages and Zebrafish. *Journal of Biological Chemistry*, 290(21), pp.13321-13343.
- Moore, K.L.; Patel, K.D.; Bruehl, R.E.; Fugang, L.; Johnson, D.A.; Lichenstein, H.S.; Cummings, R.D.; Bainton, D.F.; McEver, R.P. P-selectin glycoprotein ligand-1 mediates rolling of human neutrophils on P-selectin. *J. Cell Biol.* **1995**, 128, 661–671.

- Morla, S. (2019). Glycosaminoglycans and Glycosaminoglycan Mimetics in Cancer and Inflammation. *International Journal of Molecular Sciences*, 20(8), p.1963.
- Mosser, D. and Hondalus, M. (1996). Rhodococcus equi: An emerging opportunistic pathogen. *Trends in Microbiology*, 4(1), pp.29-33.
- Murata, K., Ochiai, Y., Akashio, K. (1985). Polydispersity of Acidic Glycosaminoglycan Components in Human-Liver and the Changes at Different Stages in Liver-Cirrhosis. *Gastroenterology*, 89 pp 1248– 1257
- Murphy PM., Baggiolini M, Charo IF, Hébert CA, Horuk R, Matsushima K, Miller LH, Oppenheim JJ, Power CA (2000). "International union of pharmacology. XXII. Nomenclature for chemokine receptors" (abstract page). *Pharmacol. Rev.* 52 (1): 145–76.
- Muscatello, G. (2012) Rhodococcus equi pneumonia in the foal-Part 1: Pathogenesis and epidemiology. *The Veterinary Journal* 192(1): 20-26.
- Nakagawa, R., Moki, H., Hayashi, K., Oniwa, K., Tokuyama, K., Kakuda, T., Yoshioka, K. And Takai, S. (2018). A case report on disseminated *Rhodococcus equi* infection in a Japanese black heifer. *Journal of Veterinary Medical Science*, 80(5), pp.819-822.
- Navratilova, I., Dioszegi, M. and Myszka, D. (2006). Analyzing ligand and small molecule binding activity of solubilized GPCRs using biosensor technology. *Analytical Biochemistry*, 355(1), pp.132-139.
- Nhlbi.nih.gov. (2020). *How the Heart Works | National Heart, Lung, and Blood Institute (NHLBI)*. [online] Available at: <https://www.nhlbi.nih.gov/health-topics/how-heart-works> [Accessed 22 Feb. 2020].

- Nie, Y., Viola, C., Bieniossek, C., Trowitzsch, S., Vijayachandran, L., Chaillet, M., Garzoni, F. and Berger, I. (2009). Getting a Grip on Complexes. *Current Genomics*, 10(8), pp.558-572.
- Nirathanan S., Awal W., Niranjana N.R. (2016) Snake α -Neurotoxins and the Nicotinic Acetylcholine Receptor. In: Gopalakrishnakone P., Inagaki H., Mukherjee A., Rahmy T., Vogel CW. (eds) Snake Venoms. Toxinology. Springer, Dordrecht
- Norouzi, R., Hojati, Z. and Badr, Z. (2016). Overview of the recombinant proteins purification by affinity tags and tags exploit systems. *Journal of Fundamental and Applied Sciences*, 8(3), p.90.
- Nowbar, A., Gitto, M., Howard, J., Francis, D. and Al-Lamee, R. (2019). Mortality from Ischemic Heart Disease. *Circulation: Cardiovascular Quality and Outcomes*, 12(6).
- Okoko, T., Blagova, E., Whittingham, J., Dover, L. and Wilkinson, A., 2015. Structural characterisation of the virulence-associated protein VapG from the horse pathogen *Rhodococcus equi*. *Veterinary Microbiology*, 179(1-2), pp.42-52.
- Olczyk, P., Mencner, Ł. and Komosińska-Vassev, K. (2015). Diverse Roles of Heparan Sulfate and Heparin in Wound Repair. *BioMed Research International*, 2015, pp.1-7.
- Oldfield, C., Bonella, H., Renwick, L., Dodson, H. I., Alderson, G. and Goodfellow, M. (2004) Rapid determination of vapA/vapB genotype in *Rhodococcus equi* using a differential polymerase chain reaction method. *Antonie van Leeuwenhoek* 85(4): 317-326.
- Ono, C., Okamoto, T., Abe, T. and Matsuura, Y. (2018). Baculovirus as a Tool for Gene Delivery and Gene Therapy. *Viruses*, 10(9), p.510.

- Ori, A., Wilkinson, M. and Fernig, D., 2011. A Systems Biology Approach for the Investigation of the Heparin/Heparan Sulfate Interactome. *Journal of Biological Chemistry*, 286(22), pp.19892-19904.
- Osborne, T.F. and Espenshade, P.J. (2009) 'Evolutionary Conservation and Adaptation in the Mechanism that Regulates SREBP Action: What A Long, Strange Trip It's Been', *PPAR Res*, 10
- Oschatz, C., Maas, C., Lecher, B., Jansen, T., Björkqvist, J., Tradler, T., Sedlmeier, R., Burfeind, P., Cichon, S., Hammerschmidt, S., Müller-Esterl, W., Wuillemin, W., Nilsson, G. and Renné, T., 2011. Mast Cells Increase Vascular Permeability by Heparin-Initiated Bradykinin Formation In Vivo. *Immunity*, 34(2), pp.258-268.
- Osterloh, K., Ewert, U. and Pries, A. (2002). Interaction of albumin with the endothelial cell surface. *American Journal of Physiology-Heart and Circulatory Physiology*, 283(1), pp.H398-H405.
- Oohira, A.; Wight, T. N.; Bornstein, P. Sulfated Proteoglycans Synthesized by Vascular Endothelial Cells in Culture *J. Biol. Chem.* 1983, 258, 2014– 2021.
- Owczarek, B., Gerszberg, A., and Hnatuszko-Konka, K. (2019). A brief reminder of systems of production and chromatography-based recovery of recombinant protein biopharmaceuticals. *Biomed Res. Int.* 2019, 1–13.
- Øynebråten, I., Barois, N., Bergeland, T., Küchler, A., Bakke, O. and Haraldsen, G., 2015. Oligomerized, filamentous surface presentation of RANTES/CCL5 on vascular endothelial cells. *Scientific Reports*, 5(1).
- Page, F. (1967). Re-Definition of the Genus *Acanthamoeba* with Descriptions of Three Species. *The Journal of Protozoology*, 14(4), pp.709-724.

- Paraskevopoulou, V., and Falcone, F. (2018). Polyionic tags as enhancers of protein solubility in recombinant protein expression. *Microorganisms* 6:47.
- Parish, C., 2006. The role of heparan sulphate in inflammation. *Nature Reviews Immunology*, 6(9), pp.633-643.
- Patching S.G. Surface plasmon resonance spectroscopy for characterisation of membrane protein-ligand interactions and its potential for drug discovery. *Biochim. Biophys. Acta*. 2014; 1838:43–55.
- Patton, K. M., McGuire, T. C., Hines, M. T., Mealey, R. H. and Hines, S. A. (2005) Rhodococcus equi-specific cytotoxic T-lymphocytes in immune horses and development in asymptomatic foals. *Infection and Immunity* 73(4): 2083–2093.
- Perozzo R., Folkers G., Scapozza L. Thermodynamics of protein-ligand interactions: History, presence, and future aspects. *J. Recept. Signal. Transduct. Res.* 2004; 24:1–52. doi: 10.1081/RRS-120037896.
- Petry, S., Sévin, C., Fleury, M., Duquesne, F., Foucher, N., Laugier, C., Henry-Amar, M. and Tapprest, J. (2017). Differential distribution of vprA-positive *Rhodococcus equi* in affected and unaffected horse-breeding farms. *Veterinary Record*, 181(6), pp.145.1-145.
- Phizicky EM, Fields S (1995) Protein-protein interactions: Methods for detection and analysis. *Microbiol Rev* 59:94–123.
- Piccicchi, A., Šiaučiūnaiteė-Gaubard, L., Petit-Hartlein, I., Sadir, R., Revilloud, J., Caro, L., Vivaudou, M., Fieschi, F., Moreau, C. and Vivès, C., 2014. C-Terminal Engineering of CXCL12 and CCL5 Chemokines: Functional

- Characterization by Electrophysiological Recordings. *PLoS ONE*, 9(1), p.e87394.
- Pomin, V. H. (2014) Heparin-Binding Proteins (Chemokines and Defensins) and their complexes with glycosaminoglycans from the solution NMR perspective. *Curr. Protein Pept. Sci.*15,738–744.80.
- Poolkhet, C., Chumsing, S., Wajjwalku, W., Minato, C., Otsu, Y. and Takai, S. (2010) Plasmid profiles and prevalence of intermediately virulent *Rhodococcus equi* from pigs in Nakhonpathom province, Thailand: Identification of a new variant of the 70-kb virulent plasmid, type 18. *Veterinary Medicine International*. DOI 4061/2010/491624.
- Porowińska, D., Wujak, M., Roszek, K. and Komoszyński, M. (2013). Prokaryotic expression systems. *Postępy Higieny i Medycyny Doświadczalnej*, 67, pp.119-129.
- Porto, A. C. R. C., Petrello, M. S., Hoge, A. Y. A., Ambrozio, G. R., Massoco, C. O. and Oliveira, C. A. (2014) Equine colostrum: more than IgG. *Journal of Equine Veterinary Science* 34(1): 239.
- Prescott, J. (1991). *Rhodococcus equi*: an animal and human pathogen. *Clinical Microbiology Reviews*, 4(1), pp.20-34.
- Prescott, J. F. (1987) Epidemiology of *Rhodococcus equi* infection in horses. *Veterinary Microbiology* 14(3): 211-214.
- Proudfoot A.E.I., Borlat F. (2000) Purification of Recombinant Chemokines from *E. coli*. In: Proudfoot A.E.I., Wells T.N.C., Power C.A. (eds) *Chemokine Protocols. Methods in Molecular Biology*, vol 138. Humana Press

- Proudfoot, A. and Ugucioni, M., 2016. Modulation of Chemokine Responses: Synergy and Cooperativity. *Frontiers in Immunology*, 7.
- Proudfoot, A., Fritchley, S., Borlat, F., Shaw, J., Vilbois, F., Zwahlen, C., Trkola, A., Marchant, D., Clapham, P. and Wells, T. (2001). The BBXB Motif of RANTES Is the Principal Site for Heparin Binding and Controls Receptor Selectivity. *Journal of Biological Chemistry*, 276(14), pp.10620-10626.
- Proudfoot, A., Handel, T., Johnson, Z., Lau, E., LiWang, P., Clark-Lewis, I., Borlat, F., Wells, T. and Kosco-Vilbois, M. (2003). Glycosaminoglycan binding and oligomerization are essential for the in vivo activity of certain chemokines. *Proceedings of the National Academy of Sciences*, 100(4), pp.1885-1890
- Proudfoot, A., Johnson, Z., Bonvin, P. and Handel, T. (2017). Glycosaminoglycan Interactions with Chemokines Add Complexity to a Complex System. *Pharmaceuticals*, 10(4), p.70.
- Quinlan, R.A. et al (1989). *Cell Sci.* 82, 7252-7255.
- Ramírez-Clavijo, S. and Montoya-Ortíz, G. (2013). *Autoimmunity*. 1st ed. Bogota (Colombia): El Rosario University Press, p.Chapter 1 Gene expression and regulation.
- Raran-Kurussi, S., and Waugh, D. S. (2012). The ability to enhance the solubility of its fusion partners is an intrinsic property of maltose-binding protein but their folding is either spontaneous or chaperone-mediated. *PLoS ONE* 7:e49589. doi: 10.1371/journal.pone.0049589
- Rathore, J. and Gautam, L., 2014. Expression, Purification, and Functional Analysis of Novel RelE Operon from *X. nematophila*. *The Scientific World Journal*, 2014, pp.1-7.

Repeke C.E., Garlet T.P., Vieira A.E., Broll D., Cunha F.Q., Garlet G.P. (2018) CCL5.

In: Choi S. (eds) Encyclopedia of Signaling Molecules. Springer, Cham.

ResearchGate. 2013. *Researchgate*. [online] Available at:

<[https://www.researchgate.net/post/Problem_in_re-](https://www.researchgate.net/post/Problem_in_re-binding_of_maltose_binding_protein_to_amylose_resin)

[binding_of_maltose_binding_protein_to_amylose_resin](https://www.researchgate.net/post/Problem_in_re-binding_of_maltose_binding_protein_to_amylose_resin)> [Accessed 13

March 2020].

Reuss, S. and Cohen, N. (2015). Update on Bacterial Pneumonia in the Foal and Weanling. *Veterinary Clinics of North America: Equine Practice*, 31(1), pp.121-135.

Reuss, S. M., Chaffin, M. K. and Cohen, N. D. (2009) Extrapulmonary disorders associated with *Rhodococcus equi* infections in foals: 150 cases (1987-2007). *Journal of the American Veterinary Medical Association* 235(7): 855-863.

Reuten, R., Nikodemus, D., Oliveira, M., Patel, T., Brachvogel, B., Breloy, I., Stetefeld, J. and Koch, M., 2016. Maltose-Binding Protein (MBP), a Secretion-Enhancing Tag for Mammalian Protein Expression Systems. *PLOS ONE*, 11(3), p.e0152386.

Rhodes, D.R., Ateeq, B., Cao, Q., Tomlins, S.A., Mehra, R., Laxman, B., Kalyana-Sundaram, S., Lonigro, R.J., Helgeson, B.E., Bhojani, M.S., Rehemtulla, A., Kleer, C.G., Hayes, D.F., Lucas, P.C., Varambally, S. and Chinnaiyan, A.M. (2009) AGTR1 overexpression defines a subset of breast cancer and confers sensitivity to losartan, an AGTR1 antagonist. *Proc Natl Acad Sci U S A*; 106(25):10284-9

- Rhodococcus Equi*: An Emerging Pathogen. (2002). *Infectious Diseases in Clinical Practice*, 11(3), p.170.
- Ribero, M. G., Takai, S., Guazzelli, A., Lara, G. H. B., da Silva, A. V., Fernandes, M. C., Condas, L. A. Z., Siqueira, A. K. and Salerno, T. (2011) Virulence genes and plasmid profiles in *Rhodococcus equi* isolates from pigs and wild boars (*Sus scrofa*) in Brazil. *Research in Veterinary Science* 91(3): 478-481.
- Rinas, U., Garcia-Fruitós, E., Corchero, J., Vázquez, E., Seras-Franzoso, J. and Villaverde, A. (2017). Bacterial Inclusion Bodies: Discovering Their Better Half. *Trends in Biochemical Sciences*, 42(9), pp.726-737.
- Robinson, P., 2015. Enzymes: principles and biotechnological applications. *Essays in Biochemistry*, 59, pp.1-41.
- Rofe, A., Davis, L., Whittingham, J., Latimer-Bowman, E., Wilkinson, A. and Pryor, P. (2016). The *Rhodococcus equi* virulence protein VapA disrupts endolysosome function and stimulates lysosome biogenesis. *MicrobiologyOpen*, 6(2), p.e00416.
- Roohvand, F., Shokri, M., Abdollahpour-Alitappeh, M. and Ehsani, P. (2017). Biomedical applications of yeast- a patent view, part one: yeasts as workhorses for the production of therapeutics and vaccines. *Expert Opinion on Therapeutic Patents*, 27(8), pp.929-951.
- Rosano, G. and Ceccarelli, E. (2014). Recombinant protein expression in *Escherichia coli*: advances and challenges. *Frontiers in Microbiology*, 5(172).
- Rosano, G. L., and Ceccarelli, E. A. (2014). Recombinant protein expression in *Escherichia coli*: advances and challenges. *Front. Microbiol.* 5:172.

- Rosano, G. L., Morales, E. S., and Ceccarelli, E. A. (2019). New tools for recombinant protein production in *Escherichia coli*: a 5-year update. *Protein Sci.* 28, 1412–1422.
- Rot, A. (1993). Neutrophil attractant/activation protein-1 (interleukin-8) induces in vitro neutrophil migration by haptotactic mechanism. *European Journal of Immunology*, 23(1), pp.303-306.
- Rubio, N., Fish, K., Trimmer, B. and Kaplan, D. (2019). Possibilities for Engineered Insect Tissue as a Food Source. *Frontiers in Sustainable Food Systems*, 3.
- Rzewuska, M., Wilkowski, L., Cisek, A. A., Stefanska, I., Chrobak, D., Stefaniuk, E., Kizerwetter-Swida, M. and Takai, S. (2014) Characterization of *Rhodococcus equi* isolates from submaxillary lymph nodes of wild boars (*Sus scrofa*), red deer (*Cervus elaphus*) and roe deer (*Capreolus capreolus*). *Veterinary Microbiology* 172(1-2): 272-278.
- Saesen, E., Sarrazin, S., Laguri, C., Sadir, R., Maurin, D., Thomas, A., Imberty, A. and Lortat-Jacob, H., 2013. Insights into the Mechanism by Which Interferon- γ Basic Amino Acid Clusters Mediate Protein Binding to Heparan Sulfate. *Journal of the American Chemical Society*, 135(25), pp.9384-9390.
- Sahdev, S., Khattar, S. and Saini, K. (2007). Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies. *Molecular and Cellular Biochemistry*, 307(1-2), pp.249-264.
- Saïda, F., Uzan, M., Odaert, B. and Bontems, F. (2006) Expression of highly toxic genes in *E. coli*: special strategies and genetic tools. **Curr Protein Pept Sci.** 7 (1):47-56.

- Salanga, C., Dyer, D., Kiselar, J., Gupta, S., Chance, M. and Handel, T., 2014. Multiple Glycosaminoglycan-binding Epitopes of Monocyte Chemoattractant Protein-3/CCL7 Enable It to Function as a Non-oligomerizing Chemokine. *Journal of Biological Chemistry*, 289(21), pp.14896-14912.
- Santillán, M. and Mackey, M. (2008). Quantitative approaches to the study of bistability in the lac operon of *Escherichia coli*. *Journal of The Royal Society Interface*, 5(suppl_1), pp. S29–S39.
- Sarvaiya, P., Guo, D., Ulasov, I., Gabikian, P. and Lesniak, M. (2013). Chemokines in tumour progression and metastasis. *Oncotarget*, 4(12).
- Satyanarayana, T. and Kunze, G. (2017). *Yeast diversity in human welfare*. 1st ed. Singapore: Springer, pp.215-250.
- Scanlon, P.J, Faxon, D.P., Audet, A.M., Carabello, B., Dehmer, G.J. and Eagle, K.A et al. (1999) 'Guidelines for Coronary Angiography.A Report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee On Coronary Angiography)', *J Am Coll Cardiol* , 33, pp. 1756-824.
- Schall ,T., Jongstra, J., Dyer, BJ., et al. (1988) A human T cell-specific molecule is a member of a new gene *family*. *J Immunol* ;141 pp1018-25
- Schall, T., Bacon, K., Toy, K. and Goeddel, D. (1990). Selective attraction of monocytes and T lymphocytes of the memory phenotype by cytokine RANTES. *Nature*, 347(6294), pp.669-671. Murphy PM International union of pharmacology. Update on chemokine receptor nomenclature. *Pharmacol Rev* 2002; 54:227-9

- Schimmel, L., Heemskerk, N. and van Buul, J. (2017). Leukocyte transendothelial migration: A local affair. *Small GTPases*, 8(1), pp.1-15.
- Scott, M., Graham, B., Verrall, R., Dixon, R., Schaffner, W. and Tham, K. (1995). *Rhodococcus equi*—An Increasingly Recognized Opportunistic Pathogen: Report of 12 Cases and Review of 65 Cases in the Literature. *American Journal of Clinical Pathology*, 103(5), pp.649-655.
- Segerer, S., Djafarzadeh, R., Gröne, H., Weingart, C., Kerjaschki, D., Weber, C., Kungl, A., Regele, H., Proudfoot, A. and Nelson, P. (2009). Selective Binding and Presentation of CCL5 by Discrete Tissue Microenvironments during Renal Inflammation. *Journal of the American Society of Nephrology*, 18(6), pp.1835-1844
- Sekar Kathiresan, M.D., Olle Melander, M.D., Ph.D., Dragi Anevski, Ph.D., Candace Guiducci, B.S., Noel P. Burt, B.S., Charlotta Roos, M.Sc., Joel N. Hirschhorn, M.D., Ph.D., Goran Berglund, M.D., Ph.D., Bo Hedblad, M.D., Ph.D., Leif Groop, M.D., Ph.D., David M. Altshuler, M.D., Ph.D., Christopher Newton-Cheh, M.D., M.P.H., and Marju Orho-Melander, Ph.D. (2008). Polymorphisms Associated with Cholesterol and Risk of Cardiovascular Events *The new england journal of medicine new england journal med* 358;12
- Seo, S., Girard, M., de la Cruz, M. and Mirkin, C., 2019. The Importance of Salt-Enhanced Electrostatic Repulsion in Colloidal Crystal Engineering with DNA. *ACS Central Science*, 5(1), pp.186-191.
- Seo, Y., Andaya, A., Blei holder, C., Leary, J. A. (2013) Differentiation of CC vs CXC chemokine dimers with GAG octa saccharide binding partners: an ion

mobility mass spectrometry approach. *J. Am. Chem. Soc.* 135,4325–4332.79.

Sepuru, K., Nagarajan, B., Desai, U. and Rajarathnam, K., 2018. Structural basis, stoichiometry, and thermodynamics of binding of the chemokines KC and MIP2 to the glycosaminoglycan heparin. *Journal of Biological Chemistry*, 293(46), pp.17817-17828.

Sharon, D. and Kamen, A. (2017). Advancements in the design and scalable production of viral gene transfer vectors. *Biotechnology and Bioengineering*, 115(1), pp.25-40.

Shintani, M., Sanchez, Z. and Kimbara, K. (2015). Genomics of microbial plasmids: classification and identification based on replication and transfer systems and host taxonomy. *Frontiers in Microbiology*, 6.

Šiekštelė, R., Veteikytė, A., Tvaska, B. and Matijošytė, I. (2015). Yeast *Kluyveromyces lactis* as host for expression of the bacterial lipase: cloning and adaptation of the new lipase gene from *Serratia* sp. *Journal of Industrial Microbiology & Biotechnology*, 42(10), pp.1309-1317.

Singh, A., Upadhyay, V., Upadhyay, A., Singh, S. and Panda, A. (2015). Protein recovery from inclusion bodies of *Escherichia coli* using mild solubilization process. *Microbial Cell Factories*, 14(1).

Singh, S., Bani Baker, Q. and Singh, D., 2022. Molecular docking and molecular dynamics simulation. *Bioinformatics*, pp.291-304.

Smith, D. B., and Johnson, K. S. (1988). Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione *S*-transferase. *Gene* 67, 31–40.

- Sokol, C. and Luster, A. (2015). The Chemokine System in Innate Immunity. *Cold Spring Harbor Perspectives in Biology*, 7(5), p.a016303.
- Solá, R. and Griebenow, K. (2010). Glycosylation of Therapeutic Proteins. *BioDrugs*, 24(1), pp.9-21.
- Song, G., Shao, X., Wu, Q., Xu, Z., Liu, Y. and Guo, Z. (2016). Novel Bioluminescent Binding Assays for Ligand–Receptor Interaction Studies of the Fibroblast Growth Factor Family. *PLOS ONE*, 11(7), p.e0159263.
- Spada, S., Yamazaki, T. and Vanpouille-Box, C., 2019. Detection and quantification of cytosolic DNA. *Methods in Enzymology*, pp.17-33.
- Spillmann, D., Witt, D. and Lindahl, U., 1998. Defining the Interleukin-8-binding Domain of Heparan Sulfate. *Journal of Biological Chemistry*, 273(25), pp.15487-15493.
- Springer, T. (1994). Traffic signals for lymphocyte recirculation and leukocyte emigration: The multistep paradigm. *Cell*, 76(2), pp.301-314.
- Stano, N. M., and Patel, S. S. (2004). T7 lysozyme represses T7 RNA polymerase transcription by destabilizing the open complex during initiation. *J. Biol. Chem.* 279, 16136–16143.
- Steele, K., Stone, B., Franklin, K., Fath-Goodin, A., Zhang, X., Jiang, H., Webb, B. and Geisler, C. (2017). Improving the baculovirus expression vector system with vankyrin-enhanced technology. *Biotechnology Progress*, 33(6), pp.1496-1507.
- Steinbrecher T., Labahn A. Towards accurate free energy calculations in ligand protein-binding studies. *Curr. Med. Chem.* 2010; 17:767–785.

- Studier, F. W. (2014). Stable expression clones and auto-induction for protein production in *E. coli*. *Methods Mol. Biol.* 1091, 17–32.
- Suganuma, T., Ino, K., Shibata, K., Kajiyama, H., Nagasaka, T., Mizutani, S. and Kikkawa, F. (2005) Functional expression of the angiotensin II type 1 receptor in human ovarian carcinoma cells and its blockade therapy resulting in suppression of tumor invasion, angiogenesis, and peritoneal dissemination. **Clin Cancer Res.**; pp11:2686–94.
- Sultan, I., Rahman, S., Jan, A., Siddiqui, M., Mondal, A. and Haq, Q. (2018). Antibiotics, Resistome and Resistance Mechanisms: A Bacterial Perspective. *Frontiers in Microbiology*, 9.
- Swanson, K. M. and Hohl, R. J. (2006) 'Anti -Cancer Therapy: Targeting the Mevalonate Pathway', *Curr Cancer Drug Targets*, (6), pp. 15-37.
- Széliová, D., Krahulec, J., Šafránek, M., Lišková, V. and Turňa, J., 2016. Modulation of heterologous expression from P BAD promoter in Escherichia coli production strains. *Journal of Biotechnology*, 236, pp.1-9.
- Takai, S. (1997). Epidemiology of Rhodococcus equi infections: A review. *Veterinary Microbiology*, 56(3-4), pp.167-176.
- Takai, S., Hines, S. A., Sekizaki, T., Nicholson, V. M., Alperin, D. A., Osaki, M., Takamatsu, D., Nakamura, M., Suzuki, K., Ogino, N., Kakuda, T., Dan, H. and Prescott, J. F. (2003) DNA Sequence and comparison of virulence plasmids from Rhodococcus equi ATCC 33701 and 103. *Infection and Immunity* 68(12): 6840-6847.
- Tan, S. and Yiap, B. (2009). DNA, RNA, and Protein Extraction: The Past and The Present. *Journal of Biomedicine and Biotechnology*, 2009, pp.1-10.

- Terpe, K. (2003). Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. *Appl. Microbiol. Biotechnol.* 60, 523–533.
- ThermoFisher.com. (2020). *Overview of Cell Lysis and Protein Extraction | Thermo Fisher Scientific - UK*. [online] Available at: <https://www.thermofisher.com/uk/en/home/life-science/protein-biology/protein-biology-learning-center/protein-biology-resource-library/pierce-protein-methods/overview-cell-lysis-and-protein-extraction.html> [Accessed 9 Feb. 2020].
- Thomas, J.L., Braus, P.A. (1988); Coronary artery disease in women: a historical perspective, *Arch Intern Med*, vol. 158 (pg. 333-337)
- Thompson, S., Martínez-Burgo, B., Sepuru, K., Rajarathnam, K., Kirby, J., Sheerin, N. and Ali, S. (2017). Regulation of Chemokine Function: The Roles of GAG-Binding and Post-Translational Nitration. *International Journal of Molecular Sciences*, 18(8), p.1692.
- Thu, H., Le, Hyung-Suk, K., Andrew, M., Allen, R.F., Spurney, O.S. and Thomas, M. C. (2003) Physiological Impact of Increased Expression of the AT1 Angiotensin Receptor. **Hypertension**. 42:507-514.
- Thunes, C., 2018. Breeding and Reproduction, Diseases and Conditions, Foal Care and Problems, Horse Care, Mare Care and Problems, Pneumonia (Rhodococcus), Respiratory Problems, Vaccinations. In: *Researchers Study Potential R. equi Vaccine*. [online] Lexington, KY: The Horse. Available at: <https://thehorse.com/157057/researchers-study-potential-r-equi-vaccine/> [Accessed 2 April 2020].

- Tolstorukov, M., Virnik, K., Zhurkin, V. and Adhya, S. (2016). Organization of DNA in a bacterial nucleoid. *BMC Microbiology*, 16(1).
- Tomlins, S.A, Rhodes, D.R, Perner, S, Dhanasekaran, S.M, Mehra, R, Sun XW, Varambally S, Cao X, Tchinda J, Kuefer R, Lee C, Montie JE, Shah RB, Pienta KJ, Rubin MA, Chinnaiyan AM. (2005) Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. **Science**. 310:644–8.
- Toyooka, K., Takai, S. and Kirikae, T. (2005) *Rhodococcus equi* survive a phagolysosomal environment in macrophages by suppressing acidification of the phagolysosome. *Journal of Medical Microbiology* 54(11): 1007-1015.
- Tummino, P. and Copeland, R. (2008). Residence Time of Receptor–Ligand Complexes and Its Effect on Biological Function. *Biochemistry*, 47(20), pp.5481-5492.
- Van der Geize, R., Grommen, A. F. W., Hessels, G. I., Jacobs, A. A. C. and Dijkhuizen, L. (2011) The steroid catabolic pathway of the intracellular pathogen *Rhodococcus equi* is important for pathogenesis and a target for vaccine development. *Plos Pathogens* 7(8): e1002181.
- Vanheule, V., Janssens, R., Boff, D., Kitic, N., Berghmans, N., Ronsse, I., Kungl, A., Amaral, F., Teixeira, M., Van Damme, J., Proost, P. and Mortier, A., 2015. The Positively Charged COOH-terminal Glycosaminoglycan-binding CXCL9 (74–103) Peptide Inhibits CXCL8-induced Neutrophil Extravasation and Monosodium Urate Crystal-induced Gout in Mice. *Journal of Biological Chemistry*, 290(35), pp.21292-21304.
- Vazquez-Boland, J. A., Giguere, S., Hapeshi, A., MacArthur, I., Anastasi, E. and Valero-Rello, A. (2013) *Rhodococcus equi* : The many facets of a pathogenic

actinomycete. *Veterinary Microbiology*. *Veterinary Microbiology* 167 (1-2): 9- 33.

Velazquez Benito, A., Juste Tejero, C., Perez Lazaro, C. and Santos Lasaosa, S. (2013) Cerebral abscess due to *Rhodococcus equi* with pseudotumour immunocompetent patient. *Neurologia* 28(8): 522-524.

Venner, M., Astheimer, K., Lammer, M. and Giguere, S. (2013) Efficacy of mass antimicrobial treatment of foals with subclinical pulmonary abscesses associated with *Rhodococcus equi*. *Journal of Veterinary Internal Medicine* 27(1):171–176

Vera-Cabrera, L., Ortiz-Lopez, R., Elizondo-Gonzalez, R. and Ocampo-Candiani, J. (2013) Complete genome sequence analysis of *Nocardia brasiliensis* HUJEG-1 reveals a saprobic lifestyle and the genes needed for human pathogenesis. *Plos One* 8(6): e65425.

Victor, X., Nguyen, T., Ethirajan, M., Tran, V., Nguyen, K. and Kuberan, B. (2009). Investigating the Elusive Mechanism of Glycosaminoglycan Biosynthesis. *Journal of Biological Chemistry*, 284(38), pp.25842-25853.

Villalba, M. S., Hernandez, M. A., Silva, R. A. and Alvarez, H. M. (2013) Genome sequences of triacylglycerol metabolism in *Rhodococcus* as a platform for comparative genomics. *Journal of Molecular Biochemistry* 2: 94-105.

Viswanathan, R., Fajardo, E., Steinberg, G., Haller, M. and Fiser, A. (2019). Protein—protein binding supersites. *PLOS Computational Biology*, 15(1), p.e1006704.

- Vivès, R. R., Sadir, R., Imbert, A., Rencurosi, A., Lortat-Jacob, H. (2002) A kinetics and modeling study of RANTES (9-68) binding to heparin reveals a mechanism of cooperative oligomerization. *Biochemistry* 41,14779–14789.
- Von Bargen, K. and Haas, A. (2009) Molecular and infection biology of the horse pathogen *Rhodococcus equi*. *FEMS Microbiology Reviews* 33(5): 870-891.
- Von Bargen, K., Polidori, M., Becken, U., Huth, G., Prescott, J. and Haas, A., 2009. *Rhodococcus equi* Virulence-Associated Protein A Is Required for Diversion of Phagosome Biogenesis but Not for Cytotoxicity. *Infection and Immunity*, 77(12), pp.5676-5681.
- Von Hundelshausen, P., Agten, S., Eckardt, V., Blanchet, X., Schmitt, M., Ippel, H., Neideck, C., Bidzhekov, K., Leberzammer, J., Wichapong, K., Faussner, A., Drechsler, M., Grommes, J., van Geffen, J., Li, H., Ortega-Gomez, A., Megens, R., Naumann, R., Dijkgraaf, I., Nicolaes, G., Döring, Y., Soehnlein, O., Lutgens, E., Heemskerk, J., Koenen, R., Mayo, K., Hackeng, T. and Weber, C., 2017. Chemokine interactome mapping enables tailored intervention in acute and chronic inflammation. *Science Translational Medicine*, 9(384), p.eeah6650.
- Vu, T.T.T., Koo, B., Song, J. *et al.* Soluble overexpression and purification of bioactive human CCL2 in *E. coli* by maltose-binding protein. *Mol Biol Rep* **42**, 651–663 (2015). <https://doi.org/10.1007/s11033-014-3812-3>
- Wagner, S. *et al.* (2008) Tuning *Escherichia coli* for membrane protein overexpression. **Proc. Natl Acad. Sci. USA** **105**, 14 371–14 376
- Walker JM (2009) *The Protein Protocols Handbook. Third Edition*. New York (NY): Springer-Verlag New York, LLC.

- Walscheid, K., Hennig, M., Heinz, C., Wasmuth, S., Busch, M., Bauer, D., Dietzel, M., Deeg, C. and Heiligenhaus, A., 2014. Correlation Between Disease Severity and Presence of Ocular Autoantibodies in Juvenile Idiopathic Arthritis-Associated Uveitis. *Investigative Ophthalmology & Visual Science*, 55(6), p.3447.
- Wang, S., Liu, S., Sun, H., Huang, T., Chan, C., Yang, C., Yeh, H., Huang, Y., Chou, W., Lin, Y. and Tang, C., 2015. CCL5/CCR5 axis induces vascular endothelial growth factor-mediated tumor angiogenesis in human osteosarcoma microenvironment. *Carcinogenesis*, 36(1), pp.104-114.
- Wang, X., Coulson, G. B., Miranda-CasoLuengo, A. A., Miranda-CasoLuengo, R., Hondalus, M. K. and Meijer, W. G. (2014) IcgA is a virulence factor of *Rhodococcus equi* that modulates intracellular growth. *Infection and Immunity* 82(5): 1793-1800.
- Waugh, D., 2016. Crystal structures of MBP fusion proteins. *Protein Science*, 25(3), pp.559-571.
- Webb, L.M.; Ehrenguber, M.U.; Clark-Lewis, I.; Baggiolini, M.; Rot, A. Binding to heparan sulfate or heparin enhances neutrophil responses to interleukin 8. *Proc. Natl. Acad. Sci. USA* **1993**, 90, 7158–7162.
- Weickert, M. J., Doherty, D. H., Best, E. A. and Olins, P. O. (1996) Optimisation of heterologous protein production in *Escherichia coli*. *Current Opinion in Biotechnology* 7 (5): 494-499.
- Weiss, R., Esko, J. and Tor, Y. (2017). Targeting heparin and heparan sulfate protein interactions. *Organic & Biomolecular Chemistry*, 15(27), pp.5656-5668.
- Weiss, S. (2000). *Nature Structural Biology*, 7(9), pp.724-729.

- White, G., Iqbal, A. and Greaves, D. (2013). CC Chemokine Receptors and Chronic Inflammation—Therapeutic Opportunities and Pharmacological Challenges. *Pharmacological Reviews*, 65(1), pp.47-89.
- Whitelock, J.M., Lozzo R, V. (2005) Heparan sulfate: a complex polymer charged with biological activity. *Chem Rev* 10 5 pp: 2745-2764.
- Willen, M., Sorrell, J., Lekan, C., Davis, B. and Caplan, A. (1990). Patterns of Glycosaminoglycan/Proteoglycan Immunostaining in Human Skin During Aging. *Journal of Investigative Dermatology*, 96(6), pp.968-974.
- Williams, I. (1998). Fibroblast: Encyclopedia of Immunology. 2nd ed. London: Academic Press.
- Willingham-Lane, J., Berghaus, L., Giguère, S. and Hondalus, M. (2016). Influence of Plasmid Type on the Replication of *Rhodococcus equi* in Host Macrophages. *mSphere*, 1(5).
- Willingham-Lane, J., Coulson, G. and Hondalus, M., 2018. Identification of a VapA virulence factor functional homolog in *Rhodococcus equi* isolates housing the pVAPB plasmid. *PLOS ONE*, 13(10), p.e0204475.
- Wingfield, P. (2015). Overview of the Purification of Recombinant Proteins. *Current Protocols in Protein Science*, 80(1).
- Wingfield, P. (2017). Overview of the Purification of Recombinant Proteins. *Current Protocols in Protein Science*
- Woodmansey, D., 2018. Scientists make leap forward in R equi vaccine to save foals. *Veternary Times*, [online] (12). Available at: <<https://www.vettimes.co.uk/news/scientists-make-leap-forward-in-r-equi-vaccine-to-save-foals/>> [Accessed 2 April 2020].

- Xu, D. and Noyes, M. (2014). Understanding DNA-binding specificity by bacteria hybrid selection. *Briefings in Functional Genomics*, 14(1), pp.3-16.
- Xu,D., Esko, J.D. (2014). Demystifying heparan sulfate-protein interactions. *Annu Rev Biochem* 129-157.
- Yager, J. A. (1987) The pathogenesis of *Rhodococcus equi* pneumonia in foals. *Veterinary Microbiology* 14(3): 225-232.
- Yeung, B., Chong, P. and Petillo, P. (2002). SYNTHESIS OF GLYCOSAMINOGLYCANS. *Journal of Carbohydrate Chemistry*, 21(7-9), pp.799-865.
- Ze-hui Liu, Di Huang, Xiang-jing Fu, Peng Cheng & En-qi Du (2018) Comparison of three commonly used fusion tags for the expression of nanobodies in the cytoplasm of *Escherichia coli*, *Biotechnology & Biotechnological Equipment*, 32:2, 462-469,
- Zhang W.Y., Zhang H.P. Genomics of Lactic Acid Bacteria. In: Zhang H.P., Cai Y.M., editors. *Lactic Acid Bacteria-Fundamentals and Practice*. 1st ed. Springer Publishing Inc.; New York, NY, USA: 2014. pp. 235–238 (15).
- Andersen, D. and Krummen, L. (2002). Recombinant protein expression for therapeutic applications. *Current Opinion in Biotechnology*, 13(2), pp.117-123.
- Zhang, Z., Li, Z.H., Wang, F., Fang, M., Yin, C.C., Zhou, Z.Y., Lin, Q. and Huang, H.L. (2002) Overexpression of DsbC and DsbG markedly improves soluble and functional expression of single-chain fv antibodies in *Escherichia coli*. *Protein Express Purif.*; Vol: 26: pp 218–228.
- Zhao X, Li G, Liang S. 2013. Several affinity tags commonly used in chromatographic purification. *J. Anal. Methods Chem.* 2013, 581093.

- Zheng, G., Zhou, H. and Li, C. (2014). Serum-Free Culture of the Suspension Cell Line QB-Tn9-4s of the Cabbage Looper, *Trichoplusia*, is Highly Productive for Virus Replication and Recombinant Protein Expression. *Journal of Insect Science*, 14(24), pp.1-14.
- Zhou, G., Niepel, M., Saretia, S. and Groth, T., 2015. Reducing the inflammatory responses of biomaterials by surface modification with glycosaminoglycan multilayers. *Journal of Biomedical Materials Research Part A*, 104(2), pp.493-502.
- Zhou, H., Liao, J., Aloor, J., Nie, H., Wilson, B., Fessler, M., Gao, H. and Hong, J., 2012. CD11b/CD18 (Mac-1) Is a Novel Surface Receptor for Extracellular Double-Stranded RNA To Mediate Cellular Inflammatory Responses. *The Journal of Immunology*, 190(1), pp.115-125.
- Zhu, J., Tao, P., Mahalingam, M., Sha, J., Kilgore, P., Chopra, A. and Rao, V., 2019. A prokaryotic-eukaryotic hybrid viral vector for delivery of large cargos of genes and proteins into human cells. *Science Advances*, 5(8), p. eaax0064.
- Zink, M. C., Yager, J. A., Prescott, J. F. and Fernando, M. A. (1987) Electron microscopic investigation of intracellular events after ingestion of *Rhodococcus equi* by foal alveolar macrophages. *Veterinary Microbiology* 14(3): 295-305.

Appendix

Number 1

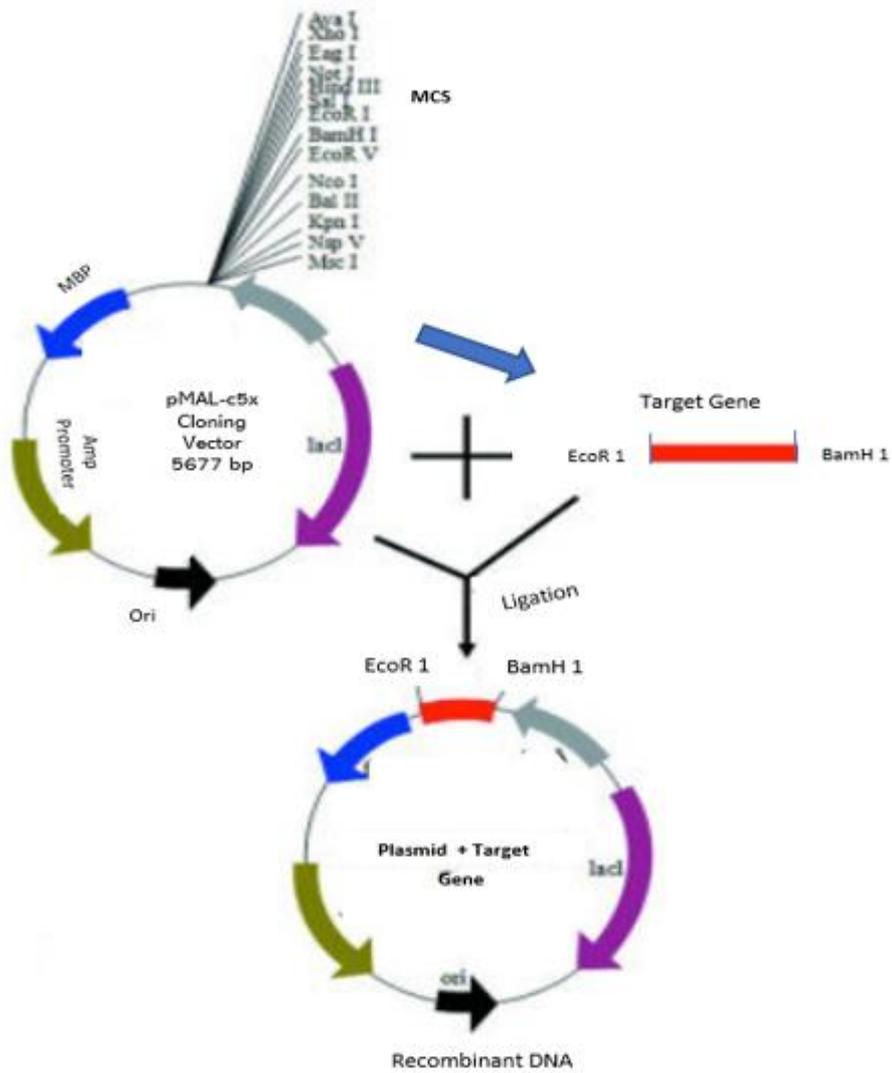


Illustration of Gene Cloning Process

Showing an *EcoR 1* and *BamH 1* digested target gene cloned into the pMAL-c5x vector expression plasmid, towards the production of recombinant genes.

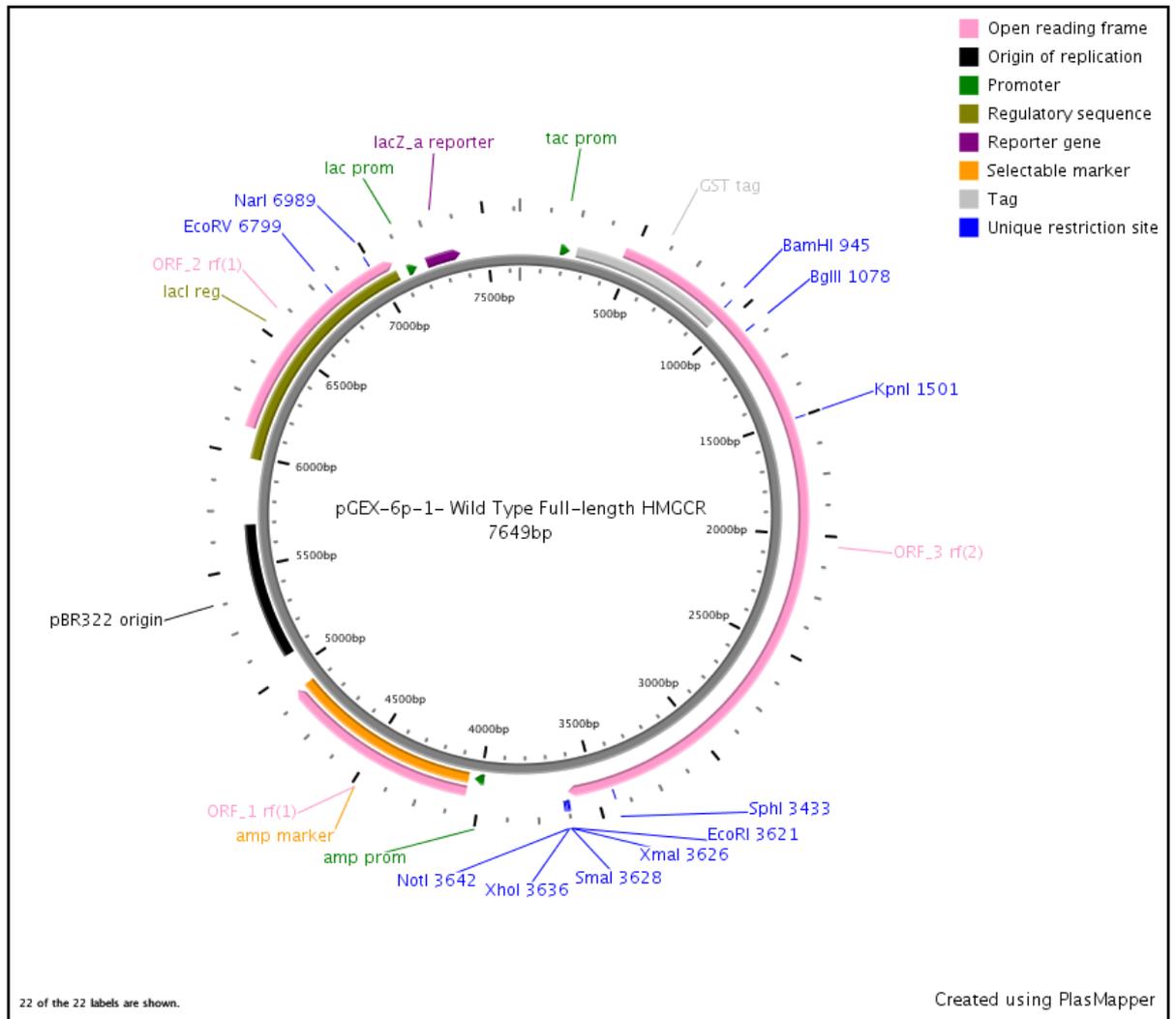
Number 2

Cloning vector pGEX-6p-1BamHI-WT Full-Length HMGR-EcoRI
complete sequence

ACGTTATCGACTGCACGGTGCACCAATGCTTCTGGCGTCAGGCAGCCATCGGAAGCTGTG
GTATGGCTGTGCAGGTCGTAAATCACTGCATAATTCGTGTCGCTCAAGGCGCACTCCCGT
TCTGGATAATGTTTTTTCGCGCCGACATCATAACGGTCTGGCAAATATTCTGAAATGAGC
TGTTGACAATTAATCATCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTCA
CACAGGAAACAGTATTCATGTCCCCTATACTAGGTTATTGGAAAATTAAGGGCCTTGTGC
AACCCACTCGACTTCTTTTGGAAATATCTTGAAGAAAATATGAAGAGCATTTGTATGAGC
GCGATGAAGGTGATAATGGCGAAACAAAAGTTTGAATTGGGTTTGGAGTTTCCCAATCT
TCCTTATTATATTGATGGTGATGTTAAATTAACACAGTCTATGGCCATCATACTTATAT
AGCTGACAAGCACAAACATGTTGGGTGGTTGTCCAAAAGAGCGTGCAGAGATTTCAATGCT
TGAAGGAGCGGTTTTGGATATTAGATACGGTGTTCGAGAATTGCATATAGTAAAGACTT
TGAAACTCTCAAAGTTGATTTTTCTTAGCAAGCTACCTGAAATGCTGAAAATGTTTGAAGA
TCGTTTTATGTCATAAAACATATTTAAATGGTGATCATGTAACCCATCCTGACTTCATGTT
GTATGACGCTCTTGATGTTGTTTTATACATGGACCCAATGTGCCTGGATGCGTTCCCAA
ATTAGTTTGTTTTAAAAACGTATTGAAGCTATCCACAAAATTGATAAGTACTTGAAATC
CAGCAAGTATATAGCATGGCCTTTGCAGGGCTGGCAAGCCACGTTTGGTGGTGGCGACCA
TCCTCCAAAATCGGATCTGGAAGTTCTGTTCCAGGGGCCCTGGGATCCCCGATGTTGTC
AAGACTTTTTTCGAATGCATGGCCTCTTTGTGGCCTCCCATCCCTGGGAAGTCATAGTGGG
GACAGTGACACTGACCATCTGCATGATGTCCATGAACATGTTTACTGGTAACAATAAGAT
CTGTGGTTGGAATTATGAATGTCCAAAGTTTGAAGAGGATGTTTTGAGCAGTGACATTAT
AATTCTGACAATAACACGATGCATAGCCATCCTGTATATTTACTTCCAGTTCAGAAATTT
ACGTCAACTTGGATCAAATATATTTTGGGTATTGCTGGCCTTTTCACAATTTTCTCAAG
TTTTGTATTACAGTACAGTTGTCATTCACTTCTTAGACAAAAGAAATTGACAGGCTTGAATGA
AGCTTTGCCCTTTTTCTACTTTTTGATTGACCTTTCCAGAGCAAGCACATTAGCAAAGTT
TGCCCTCAGTTCCAACTCACAGGATGAAGTAAGGGAAAATATTGCTCGTGGAATGGCAAT
TTTAGGTCCTACGTTTACCCTCGATGCTCTTGTGAATGTCTTGTGATTGGAGTTGGTAC
CATGTCAGGGGTACGTCAGCTTGAAATTATGTGCTGCTTTGGCTGCATGTCAGTTCTTGC
CAACTACTTCGTGTTTCATGACTTTCTTCCCAGCTTGTGTGTCCTTGGTATTAGAGCTTTC
TCGGGAAAGCCGCGAGGGTCGTCCAATTTGGCAGCTCAGCCATTTTGCCCGAGTTTTAGA
AGAAGAAGAAAATAAGCCGAATCCTGTAACCTCAGAGGGTCAAGATGATTATGTCTCTAGG
CTTGGTTCTTGTTCATGCTCACAGTCGCTGGATAGCTGATCCTTCTCCTCAAACAGTAC
AGCAGATACTTCTAAGGTTTTATTAGGACTGGATGAAAATGTGTCCAAGAGAATTGAACC
AAGTGTTCCTCTGGCAGTTTTATCTCTCTAAAATGATCAGCATGGATATTGAACAAGT
TATTACCCTAAGTTTAGCTCTCCTTCTGGCTGTCAAGTACATCTTCTTTGAACAAACAGA
GACAGAATCTACACTCTCATTAAAAAACCTATCACATCTCCTGTAGTGACACAAAAGAA
AGTCCCAGACAATTGTTGTAGACGTGAACCTATGCTGGTCAGAAATAACCAGAAATGTGA
TTCAGTAGAGGAAGAGACAGGGATAAACCAGAGAAAGAAAGTTGAGGTTATAAAACCTT
AGTGGCTGAAACAGATACCCCAACAGAGCTACATTTGTGGTTGGTAACTCCTCCTTACT
CGATACTTCATCAGTACTGGTGACACAGGAACCTGAAATTGAACTTCCAGGGAACCTCG
GCCTAATGAAGAATGTCTACAGATACTTGGGAATGCAGAGAAAGGTGCAAAATTCCTTAG
TGATGCTGAGATCATCCAGTTAGTCAATGCTAAGCATATCCCAGCCTACAAGTTGGAAAC
TCTGATGGAACTCATGAGCGTGGTGTATCTATTCGCCGACAGTTACTTTCCAAGAAGCT
TTCAGAACCTTCTTCTCTCCAGTACCTACCTTACAGGGATTATAAATACTCCTTGGTGAT
GGGAGCTTGTGTGAGAATGTTATTGGATATATGCCCATCCCTGTTGGAGTGGCAGGACC
CCTTTGCTTAGATGAAAAGAATTTCAAGTTCCAATGGCAACAACAGAAGGTTGTCTTGT

GGCCAGCACCAATAGAGGCTGCAGAGCAATAGGTCTTGGTGGAGGTGCCAGCAGCCGAGT
CCTTGCAGATGGGATGACTCGTGGCCAGTTGTGCGTCTTCCACGTGCTTGTGACTCTGC
AGAAGTGAAAGCCTGGCTCGAAACATCTGAAGGGTTCGCAGTGATAAAGGAGGCATTTGA
CAGCACTAGCAGATTTGCACGTCTACAGAACTTCATACAAGTATAGCTGGACGCAACCT
TTATATCCGTTTCCAGTCCAGGTCCAGGGATGCCATGGGGATGAACATGATTTCAAAGG
TACAGAGAAAGCACTTTCAAACCTTCACGAGTATTTCCCTGAAATGCAGATTCTAGCCGT
TAGTGGTAAC TATTGTACTGACAAGAAACCTGCTGCTATAAATTGGATAGAGGGAAGAGG
AAAATCTGTTGTTTGTGAAGCTGTCATTCCAGCCAAGGTTGTCAGAGAAGTATTAAGAC
TACCACAGAGGCTATGATTGAGGTCAACATTAACAAGAATTTAGTGGGCTCTGCCATGGC
TGGGAGCATAGGAGGCTACAACGCCCATGCAGCAAACATTGTCACCGCCATCTACATTGC
CTGTGGACAGGATGCAGCACAGAATGTTGGTAGTTCAAACCTGTATTACTTTAATGGAAGC
AAGTGGTCCCACAAATGAAGATTTATATATCAGCTGCACCATGCCATCTATAGAGATAGG
AACGGTGGGTGGTGGGACCAACCTACTACCTCAGCAAGCCTGTTTGCAGATGCTAGGTGT
TCAAGGAGCATGCAAAGATAATCCTGGGGAAAATGCCCGGCAGCTTGCCCGAATTGTGTG
TGGGACCGTAATGGCTGGGGAATTGTCACCTATGGCAGCATTGGCAGCAGGACATCTTGT
CAAAGTCACATGATTCACAACAGGTTCGAAGATCAATTTACAAGACCTCCAAGGAGCTTG
CACCAAGAAGACAGCCTGAGAATTCGCGGGTCGACTCGAGCGGCCGCATCGTGACTGACT
GACGATCTGCCTCGCGCGTTCGCGGTGATGACGGTGAAAACCTGACACATGCAGCTCCCG
GAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCG
TCAGCGGGTGTGGCGGGTGTGCGGGCGCAGCCATGACCCAGTCACGTAGCGATAGCGGA
GTGTATAATTCTTGAAGACGAAAGGGCCTCGTGATACGCCATTTTTTATAGGTTAATGTC
ATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACC
CCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCC
TGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTC
GCCCTTATTCCTTTTTTGCGGCATTTTGCCTTCTGTTTTTGTCTACCCAGAAACGCTG
GTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGAT
CTCAACAGCGGTAAGATCCTTGAGAGTTTTTCGCCCCGAAGAACGTTTTTCCAATGATGAGC
ACTTTTAAAGTTCTGCTATGTGGCGCGGTATATCCCGTGTGACGCCGGGCAAGAGCAA
CTCGGTGCGCCGATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAA
AAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGT
GATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCT
TTTTTGCACAACATGGGGGATCATGTAACCTCGCCTTGATCGTTGGGAACCGGAGCTGAAT
GAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGCAGCAATGGCAACAACGTTG
CGCAAAC TATTA ACTGGCGAACTACTTACTCTAGCTTCCC GGCAACAATTAATAGACTGG
ATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTT
ATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGG
CCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATG
GATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAAC TG
TCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACCTTCATTTTTAATTTAAA
AGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAATCCCTTAACGTGAGTTT
TCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTT
TTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAACCACCGCTACCAGCGGTGGTTTGT
TTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACCTGGCTTCAGCAGAGCGCAG
ATACCAAATACTGTCTTCTAGTG TAGCCGTAGTTAGGCCACCACTTCAAGA ACTCTGTA
GCACCGCCTACATACTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGAT
AAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTTCG
GGCTGAACGGGGGTTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTG
AGATACTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGAC
AGGTATCCGGTAAGCGGCAGGGTTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGA
AACGCCTGGTATCTTTATAGTCCTGTGCGGTTTTCGCCACCTCTGACTTGAGCGTCGATTT
TTGTGATGCTCGTCAGGGGGGCGGAGCCTATGAAAAACGCCAGCAACGCGGCCTTTTTTA

CGGTTCTGGCCTTTTGGCTGGCCTTTTGGCTCACATGTTCTTTCTGCGTTATCCCCTGAT
TCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACG
ACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTTCTC
CTTACGCATCTGTGCGGTATTTACACCCGCATAAATTCGACACCATCGAATGGTGCAAA
ACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGAGTCAATTCAGGGTGGTGAATGTG
AAACCAGTAACGTTATACGATGTCGCAGAGTATGCCGGTGTCTTATCAGACCGTTTCC
CGCGTGGTGAACCAGGCCAGCCACGTTTCTGCGAAAACGCGGGAAAAAGTGGAAAGCGGCG
ATGGCGGAGCTGAATTACATTCCCAACCGCGTGGCACAACAACCTGGCGGGCAAACAGTCG
TTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGCACGCGCCGTCGCAAATTGTCGCG
GCGATTAATCTCGCGCCGATCAACTGGGTGCCAGCGTGGTGGTGTGATGGTAGAACGA
AGCGGCGTCGAAGCCTGTAAAGCGGCGGTGCACAATCTTCTCGCGCAACGCGTCAGTGGG
CTGATCATTAACTATCCGCTGGATGACCAGGATGCCATTGCTGTGGAAGCTGCCTGCACT
AATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGACACCCATCAACAGTATTATTTTC
TCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATCTGGTCGCATTGGGTACCAGCAA
ATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGGCGCGTCTGCGTCTGGCTGGCTGG
CATAAATATCTCACTCGCAATCAAATTCAGCCGATAGCGGAACGGGAAGGCGACTGGAGT
GCCATGTCCGGTTTTCAACAAACCATGCAAATGCTGAATGAGGGCATCGTTCCTCACTGCG
ATGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGG
CTGCGCGTTGGTGCGGATATCTCGGTAGTGGGATACGACGATAACCGAAGACAGCTCATGT
TATATCCCGCCGTCAACCACCATCAAACAGGATTTTCGCTGCTGGGGCAAACCAGCGTG
GACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGAAGGGCAATCAGCTGTTGCCCGTC
TCACTGGTGAAAAGAAAAACCACCCTGGCGCCCAATACGCAAACCGCCTCTCCCCGCGCG
TTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGA
GCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCAGGCTTTACACTTTAT
GCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTACACAGGAAACAG
CTATGACCATGATTACGGATTCAGTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACC
CTGGCGTTACCCAACCTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATA
GCGAAGAGGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGC
GCTTTGCCTGGTTTTCCGGCACAGAAAGCGGTGCCGGAAAGCTGGCTGGAGTGCATCTTC
CTGAGGCCGATACTGTCGTCTCCCTCAAACCTGGCAGATGCACGGTTACGATGCGCCCA
TCTACACCAACGTAACCTATCCCATTACGGTCAATCCGCCGTTTGTTCACCGGAGAATC
CGACGGGTTGTTACTCGCTCACATTTAATGTTGATGAAAGCTGGCTACAGGAAGGCCAGA
CGCGAATTATTTTTGATGGCGTTGGAATT



Map template of pGEX-6p-1- Wild Type Full-length HMCCR

Schematics of HMCCR presented as being cloned into the BamHI and EcoRI cloning site of pGEX-6p-1. Map created using PlasMapper (Dong *et al.*, 2004).

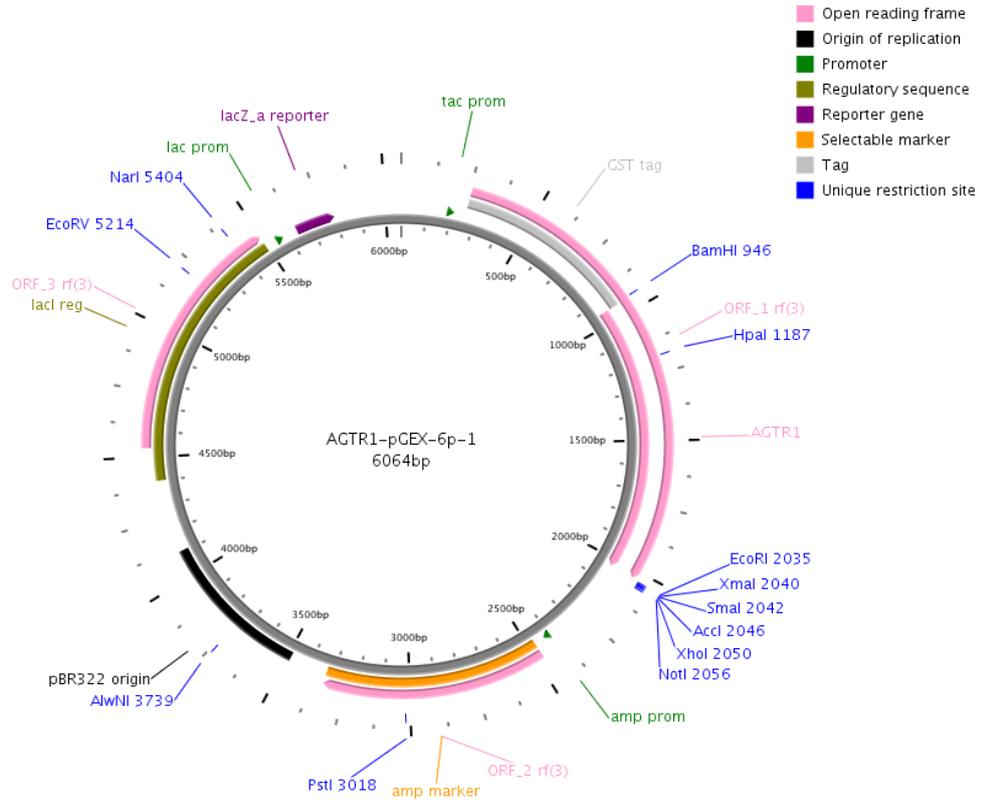
Number 3

Cloning vector pGEX-6p-1BamHI-AGTR1-EcoRI complete sequence

ACGTTATCGACTGCACGGTGCACCAATGCTTCTGGCGTCAGGCAGCCATCGGAAGCTGTG
GTATGGCTGTGCAGGTTCGTAAATCACTGCATAATTCGTGTCGCTCAAGGCGCACTCCCGT
TCTGGATAATGTTTTTTCGCGCCGACATCATAACGGTTCTGGCAAATATTCTGAAATGAGC
TGTTGACAATTAATCATCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTCA
CACAGGAAACAGTATTCATGTCCCCTATACTAGGTTATTGGAAAATTAAGGGCCTTGTGC
AACCCACTCGACTTCTTTTGGAAATATCTTGAAGAAAAATATGAAGAGCATTTGTATGAGC
GCGATGAAGGTGATAATGGCGAAACAAAAAGTTTGAATTGGGTTTGGAGTTTCCCAATCT
TCCTTATTATATTGATGGTGATGTTAAATTAACACAGTCTATGGCCATCATACGTTATAT
AGCTGACAAGCACAAACATGTTGGGTGGTTGTCCAAAAGAGCGTGCAGAGATTTCAATGCT
TGAAGGAGCGGTTTTGGATATTAGATACGGTGTTCGAGAATTGCATATAGTAAAGACTT
TGAAACTCTCAAAGTTGATTTTTCTTAGCAAGCTACCTGAAATGCTGAAAATGTTTCGAAGA
TCGTTTTATGTCATAAAACATATTTAAATGGTGATCATGTAACCCATCCTGACTTCATGTT
GTATGACGCTCTTGATGTTGTTTTATACATGGACCCAATGTGCCTGGATGCGTTCCCAA
ATTAGTTTGTTTTAAAAACGTATTGAAGCTATCCACAAAATTGATAAGTACTTGAAATC
CAGCAAGTATATAGCATGGCCTTTGCAGGGCTGGCAAGCCACGTTTGGTGGTGGCGACCA
TCCTCCAAAATCGGATCTGGAAGTTCTGTTCCAGGGGCCCTGGGATCCCCGATGATTCT
CAACTCTTCTACTGAAGATGGTATTTAAAAGAATCCAAGATGATTGTCCCAAAGCTGGAAG
GCATAATTACATATTTGTCATGATTCCCTACTTTATACAGTATCATCTTTGTGGTGGGAAT
ATTTGGAACAGCTTGGTGGTGATAGTCATTTACTTTTATATGAAGCTGAAGACTGTGGC
CAGTGTTTTTCTTTTGAATTTAGCACTGGCTGACTTATGCTTTTTTACTGACTTTGCCACT
ATGGGCTGTCTACACAGCTATGGAATACCGCTGGCCCTTTGGCAATTACCTATGTAAGAT
TGCTTCAGCCAGCGTCAGTTTCAACCTGTACGCTAGTGTGTTTCTACTCACGTGTCTCAG
CATTGATCGATACCTGGCTATTGTTACCCCAATGAAGTCCCGCCTTCGACGCACAATGCT
TGTAGCCAAAGTCACCTGCATCATATTTGGCTGCTGGCAGGCTTGGCCAGTTTGCCAGC
TATAATCCATCGAAATGTATTTTTTCAATTGAGAACACCAATATTACAGTTTGTGCTTTCCA
TTATGAGTCCCAAATTTCAACCCTCCCGATAGGGCTGGGCCTGACCAAAAATATACTGGG
TTTCTGTTTTCTTTTTCTGATCATTTCTTACAAGTTATACTCTTATTTGGAAGGCCCTAAA
GAAGGCTTATGAAATTCAGAAGAACAACCAAGAAATGATGATATTTTTTAAGATAATTAT
GGCAATTGTGCTTTTTCTTTTTCTTTTCTGGATTCCCACCAAATATTTCACTTTTCTGGA
TGATTTGATTCAACTAGGCATCATACTGACTGTAGAATTGCAGATATTGTGGACACGGC
CATGCCTATCACCATTTGTATAGCTTATTTTAAACAATTGCCTGAATCCTCTTTTTTTATGG
CTTTCTGGGGAAAAATTTAAAAGATATTTTCTCCAGCTTCTAAAATATATTTCCCCAAA
AGCCAAATCCCACTCAAACCTTTCAACAAAAATGAGCACGCTTTTCTACCGCCCCCTCAGA
TAATGTAAGCTCATCCACCAAGAAGCCTGCACCATGTTTTGAGGTTGAGTGAGAATTTCC
GGTTCGACTCGAGCGGCCGCATCGTGACTGACTGACGATCTGCCTCGCGCGTTTCGGTGA
TGACGGTGAAAACCCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCG
GATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTGGCGGGTGTGCGGGC
GCAGCCATGACCCAGTCACGTAGCGATAGCGGAGTGTATAATTCTTGAAGACGAAAGGGC
CTCGTGATACGCCTATTTTTTATAGGTTAATGTATGATAATAATGGTTTTCTTAGACGTCA
GGTGGCACTTTTTCGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACAT
TCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAAATGCTTCAATAATATTGAAAA
AGGAAGAGTATGAGTATTTCAACATTTCCGTGTGCGCCCTTATTCCCTTTTTTTGCGGCATTT
TGCCTTCTGTTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAG
TTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGT
TTTCGCCCCGAAGAACGTTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCG
GTATTATCCCGTGTGACGCCGGGCAAGAGCAACTCGGTGCGCCGCATACACTATTCTCAG
AATGACTTGGTTGAGTACTCACAGTACAGAAAAGCATCTTACGGATGGCATGACAGTA

AGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAAACACTGCGGCCAACTTACTTCTG
ACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTA
ACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATAACCAAACGACGAGCGTGAC
ACCACGATGCCTGCAGCAATGGCAACAACGTTGCGCAAACCTATTAACCTGGCGAACTACTT
ACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCA
CTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAACTGGAGCCGGTGAG
CGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTA
GTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAG
ATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTT
TAGATTGATTTAAAACCTTCATTTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGAT
AATCTCATGACCAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTA
GAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAA
ACAAAAAACCCCGCTACCAGCGGTGGTTTGTGGCCGATCAAGAGCTACCAACTCTT
TTCCGAAGGTAACCTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCTTCTAGTGTA
CCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCCGCTACATAACCTCGCTCTGCTA
ATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCA
AGACGATAGTTACCGGATAAGGCGCAGCGGTGGGGCTGAACGGGGGGTTCGTGCACACAG
CCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAA
AGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGA
ACAGGAGAGCGCACGAGGGAGCTTCCAGGGGAAACGCCTGGTATCTTTATAGTCCTGTC
GGTTTTCGCCACCTCTGACTTGAAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGC
CTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCTGGCCTTTTTGCTGGCCTTTTT
GCTCACATGTTCTTTTCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTT
GAGTGAGCTGATACCGCTCGCCGACCGCAACGACCGAGCGCAGCGAGTCAGTGAGCGAG
GAAGCGGAAGAGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTACAC
CGCATAAATTCCGACACCATCGAATGGTGCAAAACCTTTTCGCGGTATGGCATGATAGCGC
CCGGAAGAGAGTCAATTCAGGGTGGTGAATGTGAAACCAGTAACGTTATACGATGTCGCA
GAGTATGCCGGTGTCTCTTATCAGACCGTTTTCCCGCGTGGTGAACCAGGCCAGCCACGTT
TCTGCGAAAACGCGGGAAAAAGTGGAAAGCGGCGATGGCGGAGCTGAATTACATCCCAAC
CGCGTGGCACAACAACCTGGCGGGCAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGT
CTGGCCCTGCACGCGCCGTCGAAATTTGTCGCGGCGATTAAATCTCGCGCCGATCAACTG
GGTGCCAGCGTGGTGGTGTGATGGTAGAACGAAGCGGCGTCGAAGCCTGTAAAGCGGCG
GTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATCATTAACCTATCCGCTGGATGAC
CAGGATGCCATTGCTGTGGAAGCTGCCTGCACATAATGTTCCGGCGTTATTTCTTGATGTC
TCTGACCAGACACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGGC
GTGGAGCATCTGGTGCATTGGGTCAACAGCAAATCGCGCTGTTAGCGGGCCCATTAAGT
TCTGTCTCGGCGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAATT
CAGCCGATAGCGGAACGGGAAGGCGACTGGAGTGCCATGTCCGGTTTTTCAACAAACCATG
CAAATGCTGAATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCCAACGATCAGATGGCG
CTGGGCGCAATGCGCGCCATTACCGAGTCCGGGCTGCGCGTTGGTGCAGGATATCTCGGTA
GTGGGATACGACGATACCGAAGACAGCTCATGTTATATCCCGCCGTCAACCACCATCAAA
CAGGATTTTTCGCCTGCTGGGGCAAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGC
CAGGCGGTGAAGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAAGAAAAACCCCTG
GCGCCCAATACGCAAACCGCCTCTCCCGCGCGTGGCCGATTCATTAATGCAGCTGGCA
CGACAGGTTTTCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCT
CACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAAT
TGTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTACGGATTCACTGG
CCGTGCTTTTTACAACGTCGTGACTGGGAAAACCCCTGGCGTTACCCAACCTAATCGCCTTG
CAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTT
CCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCTTTGCCTGGTTTTCCGGCACCAGAAG
CGGTGCCGAAAGCTGGCTGGAGTGCGATCTTCTGAGGCCGATACTGTGCTCGTCCCT

CAAACCTGGCAGATGCACGGTTACGATGCGCCCATCTACACCAACGTAACCTATCCCATTA
 CGGTCAATCCGCCGTTTGTTCACGGAGAATCCGACGGGTTGTTACTCGCTCACATTTA
 ATGTTGATGAAAGCTGGCTACAGGAAGGCCAGACGCGAATTATTTTTGATGGCGTTGGAA
 TT



Map template of pGEX-6p-1- Wild Type AGTR1

Gene (AGTR1) was cloned into the *Bam*H1 and *Eco*R1 cloning site of pGEX-6p-1, as shown. Map created using PlasMapper (Dong *et al.*, 2004)

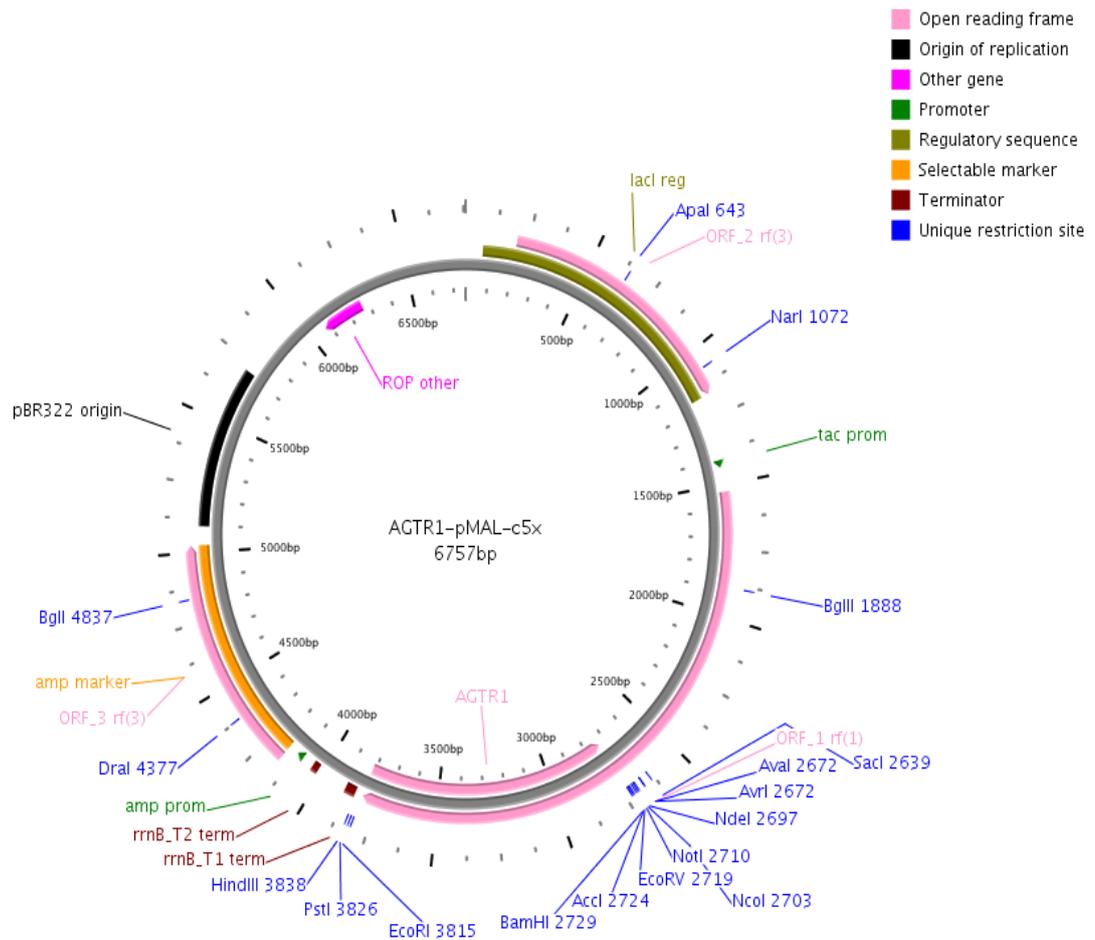
Number 4

Cloning vector pMAL-c5X-BamHI-AGTR1-EcoRI, complete sequence.

```
CCGACACCATCGAATGGTGCAAAACCTTTTCGCGGTATGGCATGATAGCGCCCGGAAGA
GAGTCAATTCAGGGTGGTGAATGTGAAACCAGTAACGTTATACGATGTCGCAGAGTAT
GCCGGTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACCAGGCCAGCCACGTTTCTG
CGAAAACGCGGGAAAAAGTGAAGCGGCGATGGCGGAGCTGAATTACATTCCCAACCG
CGTGGCACAACAACCTGGCGGGCAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGT
CTGGCCCTGCACGCGCCGTCGCAAATGTGCGCGGCGATTAAATCTCGCGCCGATCAAC
TGGGTGCCAGCGTGGTGGTGTGATGGTAGAACGAAGCGGCGTCAAGCCTGTAAAGC
GGCGGTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATCATTAACTATCCGCTG
GATGACCAGGATGCCATTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTT
TTGATGTCTCTGACCAGACACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTAC
GCGACTGGGCGTGGAGCATCTGGTTCGATTGGGTCCACCAGCAAATCGCGCTGTTAGCG
GGCCCATTAAGTTCTGTCTCGGCGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCA
CTCGCAATCAAATTCAGCCGATAGCGGAACGGGAAGGCGACTGGAGTGCCATGTCCGG
TTTTCAACAAACCATGCAAATGCTGAATGAGGGCATCGTTCCCCTGCGATGCTGGTT
GCCAACGATCAGATGGCGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGGCTGCGCG
TTGGTGCAGATATTTTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATAT
CCCGCCGTTAACCACCATCAAACAGGATTTTCGCTGCTGGGGCAAACCAGCGTGGAC
CGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGAAGGGCAATCAGCTGTTGCCCGTCT
CACTGGTGAAGAAGAAAACCACCCTGGCGCCCAATACGCAAACCGCCTCTCCCCGCGC
GTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAG
TGAGCGCAACGCAATTAATGTAAGTTAGCTCACTCATTAGGCACAATTCTCATGTTT
ACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCGTCAGGCAGCCATCGGA
AGCTGTGGTATGGCTGTGCAGGTCGTAATCACTGCATAAATTCGTGTGCTCAAGGCG
CACTCCCCTTCTGGATAATGTTTTTTGCGCCGACATCATAACGGTTCTGGCAAATATT
CTGAAATGAGCTGTTGACAATTAATCATCGGCTCGTATAATGTGTGGAATTGTGAGCG
GATAACAATTTACACAGGAAACAGCCAGTCCGTTTAGGTGTTTTACAGCAATTGA
CCAACAAGGACCATAGATTATGAAAATCGAAGAAGGTAAACTGGTAATCTGGATTAAC
GGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGATAACCG
GAATTAAGTCAACGTTGAGCATCCGGATAAACTGGAAGAGAAATTCACACAGGTTGC
GGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTAC
GCTCAATCTGGCCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAGCTGT
ATCCGTTTACCTGGGATGCCGTACGTTACAACGGCAAGCTGATTGCTTACCCGATCGC
TGTTGAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACC
TGGGAAGAGATCCCGGCGCTGGATAAAGAAGTGAAGCGAAAGGTAAGAGCGCGCTGA
TGTTCAACCTGCAAGAACCGTACTTCACCTGGCCGCTGATTGCTGCTGACGGGGGTTA
TGCGTTCAAGTATGAAAACGGCAAGTACGACATTAAGACGTGGGCGTGGATAACGCT
GGCGCGAAAGCGGGTCTGACCTTCTGGTTGACCTGATTAATAAAACAAACACATGAATG
CAGACACCGATTACTCCATCGCAGAAGCTGCCTTTAATAAAGGCGAAACAGCGATGAC
CATCAACGGCCCCTGGGCATGGTCCAACATCGACACCAGCAAAGTGAATTATGGTGT
ACGGTACTGCCGACCTTCAAGGGTCAACCATCAAACCGTTCGTTGGCGTGTGAGCG
CAGGTATTAACGCCCGCCAGTCCGAACAAAGAGCTGGCAAAGAGTTCCTCGAAAATA
TCTGCTGACTGATGAAGGTCTGGAAGCGGTTAATAAAGACAAACCGCTGGGTGCCGTA
GCGCTGAAGTCTTACGAGGAAGAGTTGGTGAAGATCCGCGTATTGCCGCCACTATGG
AAAACGCCCAGAAAGGTGAAATCATGCCGAACATCCCGCAGATGTCCGCTTTCTGGTA
TGCCGTGCGTACTGCGGTGATCAACGCCCGCCAGCGGTGCTCAGACTGTGATGAAGCC
CTGAAAGACGCGCAGACTAATTTCGAGCTCGAACAACAACAATAACAATAACAACA
```

ACCTCGGGATCGAGGGAAGGATTTACATATGTCCATGGGCGGCCGCGATATCGTCSA
CGGATCCATGATTCTCAACTCTTCTACTGAAGATGGTATTAAAAGAATCCAAGATGAT
TGTCCCAAAGCTGGAAGGCATAATTACATATTTGTCATGATTCCCTACTTTATACAGTA
TCATCTTTGTGGTGGGAATATTTGGAAACAGCTTGGTGGTGATAGTCATTTACTTTTA
TATGAAGCTGAAGACTGTGGCCAGTGTTTTTCTTTGAATTTAGCACTGGCTGACTTA
TGCTTTTTACTGACTTTGCCACTATGGGCTGTCTACACAGCTATGGAATACCGCTGGC
CCTTTGGCAATTACCTATGTAAGATTGCTTCAGCCAGCGTCAGTTTCAACCTGTACGC
TAGTGTGTTTTCTACTCACGTGTCTCAGCATTGATCGATACCTGGCTATTGTTACCCA
ATGAAGTCCCGCCTTCGACGCACAATGCTTGTAGCCAAAGTCACCTGCATCATCATTT
GGCTGCTGGCAGGCTTGGCCAGTTTGCCAGCTATAATCCATCGAAATGTATTTTTTCAT
TGAGAACACCAATATTACAGTTTGTGCTTTCCATTATGAGTCCCAAATTC AACCTC
CCGATAGGGCTGGGCCTGACCAAAAATATACTGGGTTTCTGTTCCTTTTCTGATCA
TTCTTACAAGTTATACTCTTATTTGGAAGGCCCTAAAGAAGGCTTATGAAATTCAGAA
GAACAAACCAAGAAATGATGATATTTTTTAAGATAATTATGGCAATTGTGCTTTTCTTT
TTCTTTTCTGGATTCCCCACCAATATTCACTTTTCTGGATGTATTGATTCAACTAG
GCATCATACGTGACTGTAGAATTGCAGATATTGTGGACACGGCCATGCCTATCACCAT
TTGTATAGCTTATTTTAACAATTGCCTGAATCCTCTTTTTTTATGGCTTTCTGGGGAAA
AAATTTAAAAGATATTTTTCTCCAGCTTCTAAAATATATTTCCCCAAAAGCCAAATCCC
ACTCAAACCTTTCAACAAAAATGAGCACGCTTTCTACCGCCCCTCAGATAATGTAAG
CTCATCCACCAAGAAGCCTGCACCATGTTTTGAGGTTGAGTGAGAATTCCTGCAGGT
AATTAATAAGCTTCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTT
TATCTGTTGTTTGTTCGGTGAACGCTCTCCTGAGTAGGACAAATCCGCCGGGAGCGGAT
TGAACGTTGCGAAGCAACGGCCCGGAGGGTGGCGGGCAGGACGCCGCCATAAACTGC
CAGGCATCAAATTAAGCAGAAGGCCATCCTGACGGATGGCCTTTTTGCGTTTTCTACAA
ACTCTTTTCGGTCCGTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGA
GACAATAACCCTGATAAATGCTTCAATAATATTGAAAAGGAAGAGTATGAGTATTCA
ACATTTCCGTGTGCCCCTTATTCCTTTTTTTGCGGCATTTTGCCTTCTGTTTTTTGCT
CACCCAGAAACGCTGGTGAAGTAAAAGATGCTGAAGACAGTTGGGTGCACGAGTGGG
TTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTTCGCCCCGAAGAA
CGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTG
TTGACGCCGGGCAAGAGCAACTCGGTGCGCCGATACACTATTCTCAGAATGACTTGGT
TGAGTACTCACAGTACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTA
TGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGA
TCGGAGGACCGAAGGAGCTAACCCTTTTTTGCACAACATGGGGGATCATGTAACCTCG
CCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACC
ACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACCTATTAAGTGGCGAACTACTTA
CTCTAGCTTCCCAGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACC
ACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTTATTGCTGATAAATCTGGAGCCGGT
GAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTA
TCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGAT
CGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAAGTGTGACACCAAGTTTACTCA
TATATACTTTAGATTGATTTCTTAGGACTGAGCGTCAACCCCGTAGAAAAGATCAA
GGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAC
CACCGCTACCAGCGGTGGTTTTGTTTGCAGGATCAAGAGCTACCAACTCTTTTTCCGAA
GGTAAGTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCTTCTAGTGTAGCCGTAG
TTAGGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCTCTGCTAATCC
TGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAG
ACGATAGTTACCGGATAAGGCGCAGCGGTGCGGCTGAACGGGGGGTTTCGTGCACACAG
CCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAG
AAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGT
CGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGAAACGCCTGGTATCTTTATAGT

CCTGTCGGGTTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGG
GGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCCTTTTTACGGTTCCTGGCCTTTTG
CTGGCCTTTTTGCTCACATGTTCTTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGT
ATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGAGCCGAACGACCGAGCGCAGCG
AGTCAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTTCTCCTTACGCATCT
GTGCGGTATTTACACCCGCATATAAGGTGCACTGTGACTGGGTTCATGGCTGCGCCCCG
ACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCT
TACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTACCGTCAT
CACCGAAACGCGCGAGGCAGCTGCGGTAAAGCTCATCAGCGTGGTCGTGCAGCGATTC
ACAGATGTCTGCCTGTTTCATCCGCGTCCAGCTCGTTGAGTTTCTCCAGAAGCGTTAAT
GTCTGGCTTCTGATAAAGCGGGCCATGTTAAGGGCGTTTTTTTCTGTTTGGTCACTG
ATGCCTCCGTGTAAGGGGGATTTCTGTTTCATGGGGTAATGATACCGATGAAACGAGA
GAGGATGCTCACGATACGGGTACTGATGATGAACATGCCCGGTTACTGGAACGTTGT
GAGGGTAAACAACCTGGCGGTATGGATGCGGGCGGGACCAGAGAAAAATCACTCAGGGTC
AATGCCAGCGCTTCGTTAATACAGATGTAGGTGTTCCACAGGGTAGCCAGCAGCATCC
TGCGATGCAGATCCGGAACATAATGGTGCAGGGCGCTGACTTCCGCGTTTCCAGACTT
TACGAAACACGGAAACCGAAGACCATTCATGTTGTTGCTCAGGTTCGAGACGTTTTGC
AGCAGCAGTCGTTTACGTTTCGCTCGCGTATCGGTGATTCATTCTGCTAACCAGTAAG
GCAACCCCGCCAGCCTAGCCGGTCTCAACGACAGGAGCACGATCATGCGCACCCGT
GGCCAGGACCCAACGCTGCCCGAAAT



Map template of pMAL-c5x- Wild Type AGTR1

Gene (AGTR1) was cloned into the *Bam*H1 and *Eco*R1 cloning site of pMAL-c5x, as shown. Map created using PlasMapper (Dong *et al.*, 2004).

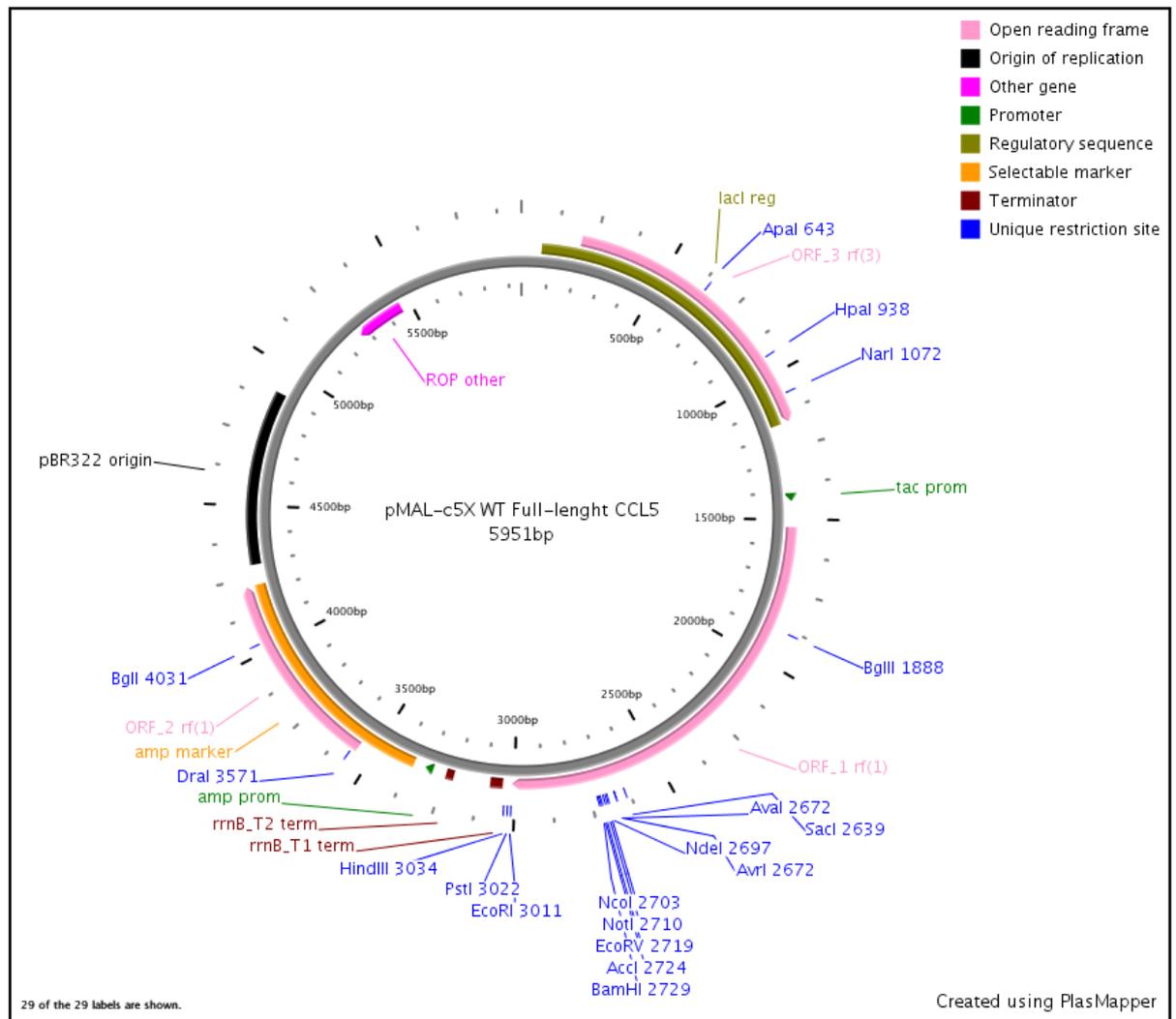
Number 5

Cloning vector pMAL-c5X-BamHI- WT Full-LenghtCCL5-EcoRI
complete sequence

```
CCGACACCATCGAATGGTGCAAAACCTTTTCGCGGTATGGCATGATAGCGCCCGGAAGA
GAGTCAATTCAGGGTGGTGAATGTGAAACCAGTAACGTTATACGATGTCGCAGAGTAT
GCCGGTGTCTCTTATCAGACCGTTTTCCCGCGTGGTGAACCAGGCCAGCCACGTTTTCTG
CGAAAACGCGGGAAAAAGTGGAAGCGGCGATGGCGGAGCTGAATTACATTCCCAACCG
CGTGGCACAACAACCTGGCGGGCAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGT
CTGGCCCTGCACGCGCCGTCGCAAATGTGCGCGGCGATTAAATCTCGCGCCGATCAAC
TGGGTGCCAGCGTGGTGGTGTGATGGTAGAACGAAGCGGCGTCAAGCCTGTAAAGC
GGCGGTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATCATTAACTATCCGCTG
GATGACCAGGATGCCATTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTT
TTGATGTCTCTGACCAGACACCCATCAACAGTATTATTTTTCTCCCATGAAGACGGTAC
GCGACTGGGCGTGGAGCATCTGGTCGCATTGGGTCCACCAGCAAATCGCGCTGTTAGCG
GGCCCATTAAGTTCTGTCTCGGCGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCA
CTCGCAATCAAATTCAGCCGATAGCGGAACGGGAAGGCGACTGGAGTGCCATGTCCGG
TTTTCAACAAACCATGCAAATGCTGAATGAGGGCATCGTTCCCCTGCGATGCTGGTT
GCCAACGATCAGATGGCGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGGCTGCGCG
TTGGTGC GGATATTTTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATAT
CCCGCCGTTAACCACCATCAAACAGGATTTTCGCTGCTGGGGCAAACCAGCGTGGAC
CGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGAAGGGCAATCAGCTGTTGCCCGTCT
CACTGGTGAAAAGAAAAACCACCCTGGCGCCAATACGCAAACCGCCTCTCCCCGCGC
GTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAG
TGAGCGCAACGCAATTAATGTAAGTTAGCTCACTCATTAGGCACAATTCTCATGTTTG
ACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCGTCAGGCAGCCATCGGA
AGCTGTGGTATGGCTGTGCAGGTCGTAATCACTGCATAAATTCGTGTGCTCAAGGCG
CACTCCCCTTCTGGATAATGTTTTTTGCGCCGACATCATAACGGTTCTGGCAAATATT
CTGAAATGAGCTGTTGACAATTAATCATCGGCTCGTATAATGTGTGGAATTGTGAGCG
GATAACAATTTACACAGGAAACAGCCAGTCCGTTTAGGTGTTTTTACGAGCAATTGA
CCAACAAGGACCATAGATTATGAAAATCGAAGAAGGTAAACTGGTAATCTGGATTAAC
GGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGATAACG
GAATTAAGTCAACGTTGAGCATCCGGATAAACTGGAAGAGAAATTCACACAGGTTGC
GGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTAC
GCTCAATCTGGCCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAGCTGT
ATCCGTTTACCTGGGATGCCGTACGTTACAACGGCAAGCTGATTGCTTACCCGATCGC
TGTTGAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACC
TGGGAAGAGATCCCGGCGCTGGATAAAGAAGTGAAGCGGAAAGGTAAGAGCGCGCTGA
TGTTCAACCTGCAAGAACCGTACTTCACTGGCCGCTGATTGCTGCTGACGGGGGTTA
TGCGTTCAAGTATGAAAACGGCAAGTACGACATTAAGACGTGGGCGTGGATAACGCT
GGCGCGAAAGCGGGTCTGACCTTCTGGTTGACCTGATTAATAAACAACACATGAATG
CAGACACCGATTACTCCATCGCAGAAGCTGCCTTTAATAAAGGCGAAACAGCGATGAC
CATCAACGGCCCGTGGGCATGGTCCAACATCGACACCAGCAAAGTGAATTATGGTGTG
ACGGTACTGCCGACCTTCAAGGGTCAACCATCAAACCGTTCGTTGGCGTGCTGAGCG
CAGGTATTAACGCCGCCAGTCCGAACAAGAGCTGGCAAAGAGTTCCTCGAAAACATA
TCTGCTGACTGATGAAGGTCTGGAAGCGGTTAATAAAGACAAACCGCTGGGTGCCGTA
GCGCTGAAGTCTTACGAGGAAGAGTTGGTGAAAGATCCGCGTATTGCCGCCACTATGG
AAAACGCCCGAAAGGTGAAATCATGCCGAACATCCCGCAGATGTCCGCTTTCTGGTA
TGCCGTGCGTACTGCGGTGATCAACGCCGCCAGCGGTGCTCAGACTGTCGATGAAGCC
```

CTGAAAGACGCGCAGACTAATTCGAGCTCGAACAACAACAATAACAATAACAACA
ACCTCGGGATCGAGGGAAGGATTTACATATGTCCATGGGCGGCCGCGATATCGTTCGA
CGGATCCATGAAGGTCTCCGCGGCAGCCCTCGCTGTCATCCTCATTGCTACTGCCCTC
TGCGCTCCTGCATCTGCCTCCCATATTCCTCGGACACCACACCCTGCTGCTTTGCCT
ACATTGCCCCGCCACTGCCCCGTGCCACATCAAGGAGTATTTCTACACCAGTGGCAA
GTGCTCCAACCCAGCAGTCGTCTTTGTACCCGAAAGAACCGCCAAGTGTGTGCCAAC
CCAGAGAAGAAATGGGTTCGGGAGTACATCAACTCTTTGGAGATGAGCTAGGAATTC
CTGCAGGTAATTAATAAGCTTCAAATAAAAACGAAAGGCTCAGTCGAAAGACTGGGCC
TTTCGTTTTATCTGTTGTTTGTTCGGTGAACGCTCTCCTGAGTAGGACAAATCCGCCGG
GAGCGGATTGAACGTTGCGAAGCAACGGCCCGGAGGGTGGCGGGCAGGACGCCGCCA
TAAACTGCCAGGCATCAAATTAAGCAGAAGGCCATCCTGACGGATGGCCTTTTTGCGT
TTCTACAAACTCTTTTCGGTCCGTTGTTTATTTTTCTAAATACATTCAAATATGTATCC
GCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATG
AGTATTCAACATTTCCGTGTGCCCCATTATCCCTTTTTTTGCGGCATTTTGCCTTCCTG
TTTTTGTCTACCCAGAAACGCTGGTGAAGTAAAGATGCTGAAGACAGTTGGGTGCA
CGAGTGGGTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCC
CCGAAGAACGTTTCCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATT
ATCCCGTGTGACGCCGGCAAGAGCAACTCGGTGCGCCGCATACACTATTCTCAGAAT
GACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAA
GAGAATTATGCAGTGTGCCATAACCATGAGTGATAAACTGCGGCCAACTTACTTCT
GACAACGATCGGAGGACCGAAGGAGCTAACCCTTTTTTGCACAACATGGGGGATCAT
GTAACCTGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGC
GTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACATTAACTGGCGA
ACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTT
GCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTTATTGCTGATAAATCTG
GAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCC
CTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAAT
AGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAG
TTTACTCATATATACTTTAGATTGATTTCTTAGGACTGAGCGTCAACCCCGTAGAAA
AGATCAAAGGATCTTCTTGAGATCCTTTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAAC
AAAAAAACCACCGCTACCAGCGGTGGTTTTGTTTGC CGGATCAAGAGCTACCAACTCTT
TTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGT
AGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCT
GCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTG
GACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGT
GCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGA
GCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGC
GGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATC
TTTATAGTCTGTGCGGTTTTCGCCACCTCTGACTTGAGCGTCGATTTTTTGTGATGCTC
GTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTG
GCCTTTTTGCTGGCCTTTTTGCTCACATGTTCTTTCTGCGTTATCCCCTGATTCTGTGG
ATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGA
GCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTTCTCCTT
ACGCATCTGTGCGGTATTTACACCCGCATATAAGGTGCACTGTGACTGGGTTCATGGCT
GCGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGG
CATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTGAGAGGTTTTT
ACCGTCATCACCGAAACGCGCGAGGCAGCTGCGGTAAAGCTCATCAGCGTGGTCTGTC
AGCGATTACAGATGTCTGCCTGTTTATCCGCGTCCAGCTCGTTGAGTTTTCTCCAGAA
GCGTTAATGTCTGGCTTCTGATAAAGCGGGCCATGTTAAGGGCGGTTTTTTTTCTGTTT
GGTCACTGATGCCTCCGTGTAAGGGGATTTCTGTTTATGGGGTAATGATACCGATG
AAACGAGAGAGGATGCTCACGATACGGGTACTGATGATGAACATGCCCGGTTACTGG

AACGTTGTGAGGGTAAACAACACTGGCGGTATGGATGCGGCGGGACCAGAGAAAAATCAC
 TCAGGGTCAATGCCAGCGCTTCGTTAATACAGATGTAGGTGTTCCACAGGGTAGCCAG
 CAGCATCCTGCGATGCAGATCCGGAACATAATGGTGCAGGGCGCTGACTTCCGCGTTT
 CCAGACTTTACGAAACACGGAAACCGAAGACCATTTCATGTTGTTGCTCAGGTCGCAGA
 CGTTTTGCAGCAGCAGTCGCTTCACGTTTCGCTCGCGTATCGGTGATTTCATCTGCTAA
 CCAGTAAGGCAACCCCGCCAGCCTAGCCGGTCTCAACGACAGGAGCACGATCATGC
 GCACCCGTGGCCAGGACCCAACGCTGCCCGAAATT



Map of pMAL-c5x- Wild Type CCL5

Gene (CCL5) was cloned into the *Bam*H1 and *Eco*R1 cloning site of pMAL-c5X as shown. Map created using PlasMapper. (Dong *et al.*, 2004).

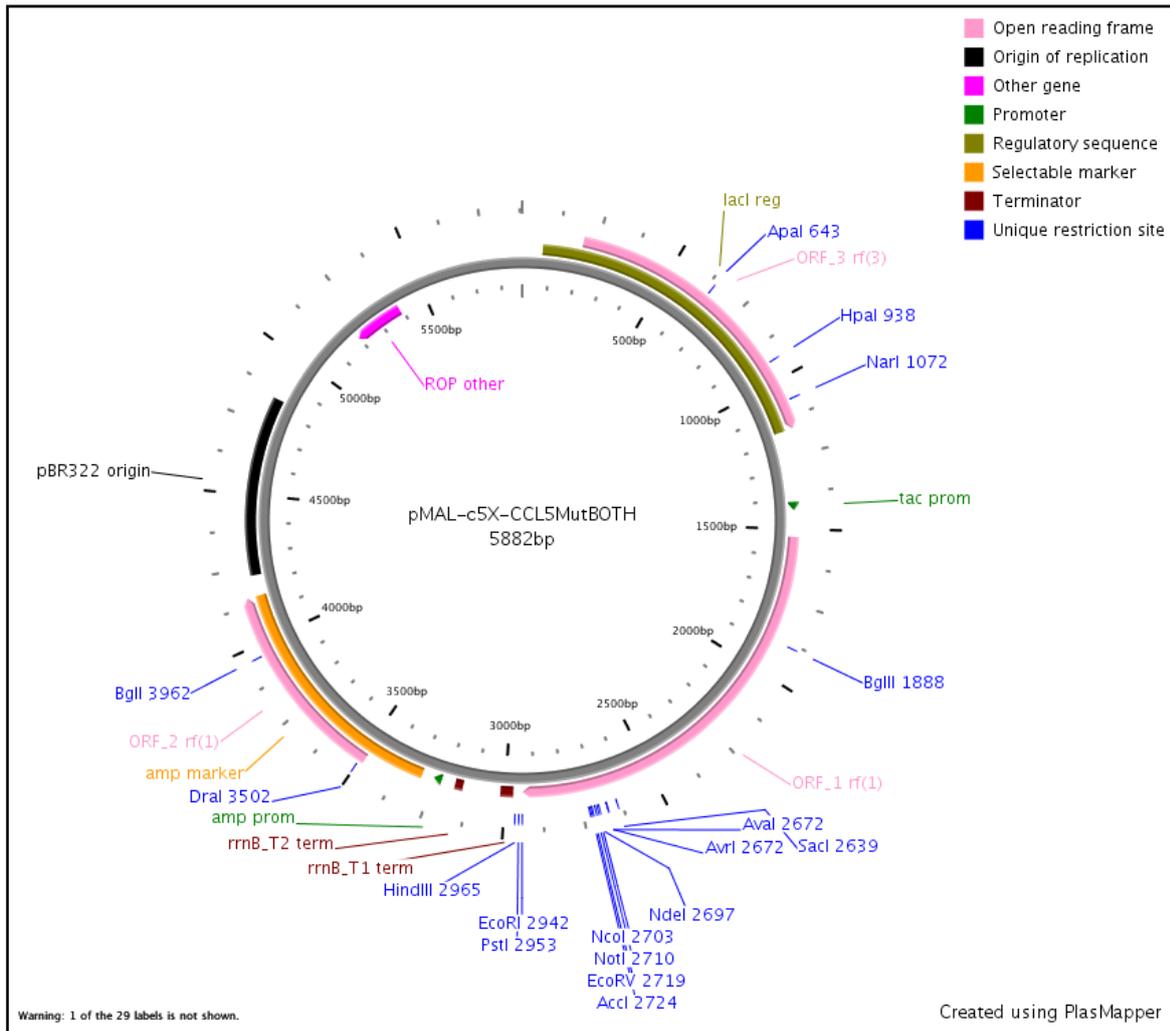
Number 6

Cloning vector pMAL-c5X-BamHI- CCL5MutBOTH-EcoRI complete sequence

```
CCGACACCATCGAATGGTGCAAAACCTTTTCGCGGTATGGCATGATAGCGCCCGGAAGA
GAGTCAATTCAGGGTGGTGAATGTGAAACCAGTAACGTTATACGATGTCGCAGAGTAT
GCCGGTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACCAGGCCAGCCACGTTTCTG
CGAAAACGCGGGAAAAAGTGGAAGCGGCGATGGCGGAGCTGAATTACATTCCCAACCG
CGTGGCACAACAACCTGGCGGGCAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGT
CTGGCCCTGCACGCGCCGTCGCAAATGTGCGCGGCGATTAAATCTCGCGCCGATCAAC
TGGGTGCCAGCGTGGTGGTGTGATGGTAGAACGAAGCGGCGTCAAGCCTGTAAAGC
GGCGGTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATCATTAACTATCCGCTG
GATGACCAGGATGCCATTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTT
TTGATGTCTCTGACCAGACACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTAC
GCGACTGGGCGTGGAGCATCTGGTTCGATTGGGTCCACCAGCAAATCGCGCTGTTAGCG
GGCCCATTAAGTTCTGTCTCGGCGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCA
CTCGCAATCAAATTCAGCCGATAGCGGAACGGGAAGGCGACTGGAGTGCCATGTCCGG
TTTTCAACAAACCATGCAAATGCTGAATGAGGGCATCGTTCCCCTGCGATGCTGGTT
GCCAACGATCAGATGGCGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGGCTGCGCG
TTGGTGC GGATATTTTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATAT
CCCGCCGTTAACCACCATCAAACAGGATTTTCGCTGCTGGGGCAAACCAGCGTGGAC
CGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGAAGGGCAATCAGCTGTTGCCCGTCT
CACTGGTGA AAAAGAAAACCACCCTGGCGCCCAATACGCAAACCGCCTCTCCCCGCGC
GTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAG
TGAGCGCAACGCAATTAATGTAAGTTAGCTCACTCATTAGGCACAATTCTCATGTTG
ACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCGTCAGGCAGCCATCGGA
AGCTGTGGTATGGCTGTGCAGGTGCTAAATCACTGCATAAATTCGTGTGCTCAAGGCG
CACTCCCCTTCTGGATAATGTTTTTTGCGCCGACATCATAACGGTTCTGGCAAATATT
CTGAAATGAGCTGTTGACAATTAATCATCGGCTCGTATAATGTGTGGAATTGTGAGCG
GATAACAATTTACACAGGAAACAGCCAGTCCGTTTAGGTGTTTTACAGCAATTGA
CCAACAAGGACCATAGATTATGAAAATCGAAGAAGGTAAACTGGTAATCTGGATTAAC
GGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGATAACCG
GAATTAAGTCAACGTTGAGCATCCGGATAAACTGGAAGAGAAATCCCACAGGTTGC
GGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTAC
GCTCAATCTGGCCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAGCTGT
ATCCGTTTACCTGGGATGCCGTACGTTACAACGGCAAGCTGATTGCTTACCCGATCGC
TGTTGAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACC
TGGGAAGAGATCCCGGCGCTGGATAAAGAAGTGAAGCGGAAAGGTAAGAGCGCGCTGA
TGTTCAACCTGCAAGAACCGTACTTCACCTGGCCGCTGATTGCTGCTGACGGGGGTTA
TGCGTTCAAGTATGAAAACGGCAAGTACGACATTAAGACGTGGGCGTGGATAACGCT
GGCGCGAAAGCGGGTCTGACCTTCCCTGGTTGACCTGATTA AAAACAAACACATGAATG
CAGACACCGATTACTCCATCGCAGAAGCTGCCTTTAATAAAGGCGAAACAGCGATGAC
CATCAACGGCCCGTGGGCATGGTCCAACATCGACACCAGCAAAGTGAATTATGGTGTG
ACGGTACTGCCGACCTTCAAGGGTCAACCATCAAACCGTTCGTTGGCGTGTGAGCG
CAGGTATTAACGCCCGCCAGTCCGAACAAAGAGCTGGCAAAGAGTTCCTCGAAAATA
TCTGCTGACTGATGAAGGTCTGGAAGCGGTTAATAAAGACAAACCGCTGGGTGCCGTA
GCGCTGAAGTCTTACGAGGAAGAGTTGGTGAAGATCCGCGTATTGCCGCCACTATGG
AAAACGCCCAGAAAGGTGAAATCATGCCGAACATCCCGCAGATGTCCGCTTTCTGGTA
```

TGCCGTGCGTACTGCGGTGATCAACGCCGCCAGCGGTTCGTCAGACTGTTCGATGAAGCC
CTGAAAGACGCGCAGACTAATTTCGAGCTCGAACAACAACAATAACAATAACAACA
ACCTCGGGATCGAGGGAAGGATTTTACATATGTCCATGGGCGGCCGCGATATCGTCGA
CGGATCCAGCCCCTATTCAAGTGACACCACTCCTTGCTGTTTTCGCTTATATCGCTCGC
CCTTTACCACGCGCCACATTAAGGAGTATTTTTTACACTAGCGGTAAATGTAGTAACC
CCGCCGTGGTCTTCGTTACCGCTGCAAATGCCAGGTCTGCGCGAATCCAGAAGCAGC
ATGGGTAGCCGAGTATATCAATTCCTTAGAAATGTCCTAGGAATTCCTGCAGGTAAT
TAAATAAGCTTCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTAT
CTGTTGTTTGTTCGGTGAACGCTCTCCTGAGTAGGACAAATCCGCCGGGAGCGGATTGA
ACGTTGCGAAGCAACGGCCCCGAGGGTGGCGGGCAGGACGCCCGCCATAAACTGCCAG
GCATCAAATTAAGCAGAAGGCCATCCTGACGGATGGCCTTTTTTTCGTTTTCTACAACT
CTTTCGGTCCGTTGTTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGAC
AATAACCCTGATAAATGCTTCAATAATATTGAAAAGGAAGAGTATGAGTATTCAACA
TTTCCGTGTCGCCCTTATTCCTTTTTTTCGCGCATTTTTGCCTTCTGTTTTTTCGTCAC
CCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGACAGTTGGGTGCACGAGTGGGTTA
CATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTTCGCCCCGAAGAAGT
TTCCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTGTG
ACGCCGGGCAAGAGCAACTCGGTTCGCCGCATACACTATTCTCAGAATGACTTGGTTGA
GTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGC
AGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCG
GAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACCTCGCCT
TGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACG
ATGCCTGTAGCAATGGCAACAACGTTGCGCAAACCTATTAACGGCGAACTACTTACTC
TAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACT
TCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTTATTGCTGATAAATCTGGAGCCGGTGAG
CGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCG
TAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGC
TGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATAT
ATACTTTAGATTGATTTCTTAGGACTGAGCGTCAACCCCGTAGAAAAGATCAAAGGA
TCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAACAC
CGCTACCAGCGGTGGTTTTGTTTGC CGGATCAAGAGCTACCAACTCTTTTTCCGAAGGT
AACTGGCTTTCAGCAGAGCGCAGATAACAAATACTGTCCTTCTAGTGTAGCCGTAGTTA
GGCCACCACTTCAAGAACTCTGTAGCACC GCCTACATACCTCGCTCTGCTAATCCTGT
TACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACG
ATAGTTACCGGATAAGGCGCAGCGGTTCGGGCTGAACGGGGGGTTTCGTGCACACAGCCC
AGCTTGGAGCGAACGACCTACACCGAACTGAGATACTACAGCGTGAGCTATGAGAAA
GCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTTCGG
AACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCT
GTCGGGTTTTCGCCACCTCTGACTTGAGCGTCGATTTTTTGTGATGCTCGTCAGGGGGGC
GGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCTGGCCTTTTGTG
GCCTTTTTGCTCACATGTTCTTTTCTGCGTTATCCCTGATTCTGTGGATAACCGTATT
ACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGT
CAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTG
CGGTATTTACACCCGCATATAAGGTGCACTGTGACTGGGTCATGGCTGCGCCCCGACA
CCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTAC
AGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTTACCCTCATCAC
CGAAACGCGCGAGGCAGCTGCGGTAAGCTCATCAGCGTGGTTCGTGCAGCGATTCA
GATGTCTGCCTGTTTATCCGCGTCCAGCTCGTTGAGTTTTCTCCAGAAGCGTTAATGTC
TGGCTTCTGATAAAGCGGGCCATGTTAAGGGCGGTTTTTTTTCTGTTTGGTCACTGATG
CCTCCGTGTAAGGGGATTTCTGTTTATGGGGTAATGATACCGATGAAACGAGAGAG
GATGCTCACGATACGGGTTACTGATGATGAACATGCCCGGTTACTGGAACGTTGTGAG

GGTAAACAACCTGGCGGTATGGATGCGGCGGGACCAGAGAAAAATCACTCAGGGTCAAT
 GCCAGCGCTTCGTTAATACAGATGTAGGTGTTCCACAGGGTAGCCAGCAGCATCCTGC
 GATGCAGATCCGGAACATAATGGTGCAGGGCGCTGACTTCCGCGTTTCCAGACTTTAC
 GAAACACGGAAACCGAAGACCATTTCATGTTGTTGCTCAGGTCGCAGACGTTTTGCAGC
 AGCAGTCGCTTACGTTTCGCTCGCGTATCGGTGATTCATTCTGCTAACCAGTAAGGCA
 ACCCCGCCAGCCTAGCCGGGTCTCAACGACAGGAGCACGATCATGCGCACCCCGTGGC
 CAGGACCCAACGCTGCCCGAAATT



Map of pMAL-c5x- CCL5MutBOTH

Gene (CCL5MutBOTH) was cloned into the *Bam*H1 and *Eco*R1 cloning site of pMAL-c5X as shown. Map created using PlasMapper. (Dong *et al.*, 2004).

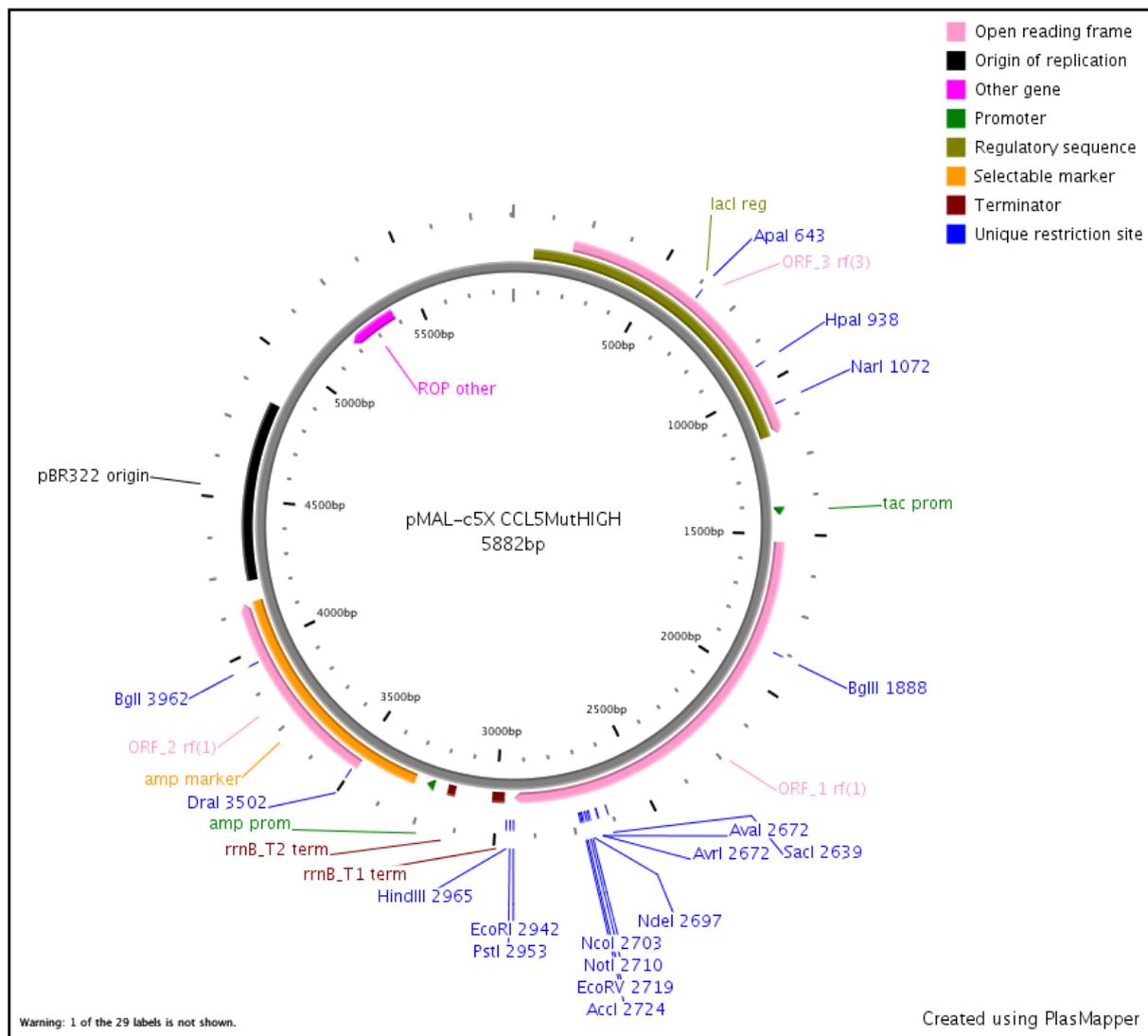
Number 7

Cloning vector pMAL-c5X-BamHI- CCL5MuthIGH-EcoRI complete sequence

```
CCGACACCATCGAATGGTGCAAAACCTTTTCGCGGTATGGCATGATAGCGCCCGGAAGA
GAGTCAATTCAGGGTGGTGAATGTGAAACCAGTAACGTTATACGATGTCGCAGAGTAT
GCCGGTGTCTCTTATCAGACCGTTTTCCCGCGTGGTGAACCAGGCCAGCCACGTTTTCTG
CGAAAACGCGGGAAAAAGTGGAAGCGGCGATGGCGGAGCTGAATTACATTCCCAACCG
CGTGGCACAACAACCTGGCGGGCAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGT
CTGGCCCTGCACGCGCCGTCGCAAATGTGCGCGGCGATTAAATCTCGCGCCGATCAAC
TGGGTGCCAGCGTGGTGGTGTGATGGTAGAACGAAGCGGCGTCAAGCCTGTAAAGC
GGCGGTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATCATTAACTATCCGCTG
GATGACCAGGATGCCATTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTT
TTGATGTCTCTGACCAGACACCCATCAACAGTATTATTTTTCTCCCATGAAGACGGTAC
GCGACTGGGCGTGGAGCATCTGGTCGCATTGGGTCCACCAGCAAATCGCGCTGTTAGCG
GGCCCATTAAGTTCTGTCTCGGCGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCA
CTCGCAATCAAATTCAGCCGATAGCGGAACGGGAAGGCGACTGGAGTGCCATGTCCGG
TTTTCAACAAACCATGCAAATGCTGAATGAGGGCATCGTTCCCACTGCGATGCTGGTT
GCCAACGATCAGATGGCGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGGCTGCGCG
TTGGTGC GGATATTTTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATAT
CCCGCCGTTAACCACCATCAAACAGGATTTTCGCTGCTGGGGCAAACCAGCGTGGAC
CGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGAAGGGCAATCAGCTGTTGCCCGTCT
CACTGGTGAAAAGAAAAACCACCCTGGCGCCAATACGCAAACCGCCTCTCCCCGCGC
GTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAG
TGAGCGCAACGCAATTAATGTAAGTTAGCTCACTCATTAGGCACAATTCTCATGTTTG
ACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCGTCAGGCAGCCATCGGA
AGCTGTGGTATGGCTGTGCAGGTCGTAATCACTGCATAATTCGTGTGCTCAAGGCG
CACTCCCGTCTTGATAATGTTTTTTGCGCCGACATCATAACGGTCTTGCCAAATATT
CTGAAATGAGCTGTTGACAATTAATCATCGGCTCGTATAATGTGTGGAATTGTGAGCG
GATAACAATTTACACAGGAAACAGCCAGTCCGTTTAGGTGTTTTTACGAGCAATTGA
CCAACAAGGACCATAGATTATGAAAATCGAAGAAGGTAAACTGGTAATCTGGATTAAC
GGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGATACCG
GAATTAAGTCAACGTTGAGCATCCGGATAAACTGGAAGAGAAATCCCACAGGTTGC
GGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTAC
GCTCAATCTGGCCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCAGGACAAGCTGT
ATCCGTTTACCTGGGATGCCGTACGTTACAACGGCAAGCTGATTGCTTACCCGATCGC
TGTTGAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACC
TGGGAAGAGATCCCGGCGCTGGATAAAGAAGTGAAGCGGAAAGGTAAGAGCGCGCTGA
TGTTCAACCTGCAAGAACCGTACTTCACTGGCCGCTGATTGCTGCTGACGGGGGTTA
TGCGTTCAAGTATGAAAACGGCAAGTACGACATTAAGACGTGGGCGTGGATAACGCT
GGCGCGAAAGCGGGTCTGACCTTCTGGTTGACCTGATTA AAAACAAACACATGAATG
CAGACACCGATTACTCCATCGCAGAAGCTGCCTTTAATAAAGGCGAAACAGCGATGAC
CATCAACGGCCCCTGGGCATGGTCCAACATCGACACCAGCAAAGTGAATTATGGTGTG
ACGGTACTGCCGACCTTCAAGGGTCAACCATCAAACCGTTCGTTGGCGTGCTGAGCG
CAGGTATTAACGCCGCCAGTCCGAACAAAGAGCTGGCAAAGAGTTCCTCGAAAACATA
TCTGCTGACTGATGAAGGTCTGGAAGCGGTTAATAAAGACAAACCGCTGGGTGCCGTA
GCGCTGAAGTCTTACGAGGAAGAGTTGGTGAAAGATCCGCGTATTGCCGCCACTATGG
AAAACGCCCGAAAGGTGAAATCATGCCGAACATCCCGCAGATGTCCGCTTTCTGGTA
TGCCGTGCGTACTGCGGTGATCAACGCCGCCAGCGGTGCTCAGACTGTCGATGAAGCC
```

CTGAAAGACGCGCAGACTAATTCGAGCTCGAACAACAACAATAACAATAACAACA
ACCTCGGGATCGAGGGAAGGATTTACATATGTCCATGGGCGGCCGCGATATCGTCTGA
CGGATCCTCACCGTATAGCAGTGATACTACACCTTGCTGTTTTGCGTACATCGCACGC
CCTCTGCCCCGCGCCATATTAAGAGTATTTTTTACTTTCGGGGAAGTGTAGTAACC
CCGCTGTTCGTATTCGTTACAGCCGCTAACGCACAAGTGTGTGCAAACCCGGAAAAAAA
ATGGGTGCGTGAGTACATCAACTCCCTGGAGATGTCATAGGAATTCCTGCAGGTAAT
TAAATAAGCTTCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTTCGTTTTAT
CTGTTGTTTGTTCGGTGAACGCTCTCCTGAGTAGGACAAATCCGCCGGGAGCGGATTGA
ACGTTGCGAAGCAACGGCCCCGGAGGGTGGCGGGCAGGACGCCCGCCATAAACTGCCAG
GCATCAAATTAAGCAGAAGGCCATCCTGACGGATGGCCTTTTTGCGTTTTCTACAAACT
CTTTCGGTCCGTTGTTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGAC
AATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACA
TTTCCGTGTTCGCCCTTATTCCCTTTTTTGCGGCATTTCCTTCCCTGTTTTTGTCTCAC
CCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGACAGTTGGGTGCACGAGTGGGTTA
CATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTTCGCCCCGAAGAAGT
TTCCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTGTG
ACGCCGGGCAAGAGCAACTCGGTTCGCCGATACACTATTCTCAGAATGACTTGGTTGA
GTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGC
AGTGTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCG
GAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACCTCGCCT
TGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACG
ATGCCTGTAGCAATGGCAACAACGTTGCGCAAACCTATTAACCTGGCGAACTACTTACTC
TAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACT
TCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTTATTGCTGATAAATCTGGAGCCGGTGAG
CGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCG
TAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGC
TGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATAT
ATACTTTAGATTGATTTCCCTTAGGACTGAGCGTCAACCCCGTAGAAAAGATCAAAGGA
TCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAACAC
CGCTACCAGCGGTGGTTTTGTTTGC CGGATCAAGAGCTACCAAACCTTTTTTCCGAAGGT
AACTGGCTTCAGCAGAGCGCAGATAACAAATACTGTCCTTCTAGTGTAGCCGTAGTTA
GGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGT
TACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGCTTACCAGGGTTGGACTCAAGACG
ATAGTTACCGGATAAAGCGCAGCGGTCGGGCTGAACGGGGGGTTTCGTGCACACAGCCC
AGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAA
GCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCCG
AACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCTGGTATCTTTATAGTCCT
GTCGGGTTTTGCCACCTCTGACTTGAGCGTCGATTTTTTGTGATGCTCGTCAGGGGGGC
GGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCCTGGCCTTTTGTG
GCCTTTTGTCTCACATGTTCTTTTCTGCGTTATCCCCTGATTCTGTGGATAACCGTATT
ACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGT
CAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTTTCTCCTTACGCATCTGTG
CGGTATTTACACCGCATATAAGGTGCACTGTGACTGGGTTCATGGCTGCGCCCCGACA
CCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTAC
AGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTACACCGTCATCAC
CGAAACGCGCGAGGCAGCTGCGGTAAGCTCATCAGCGTGGTTCGTGCAGCGATTCA
GATGTCTGCCTGTTTCATCCGCGTCCAGCTCGTTGAGTTTTCTCCAGAAGCGTTAATGTC
TGGCTTCTGATAAAGCGGGCCATGTTAAGGGCGGTTTTTTTCTGTTTGGTCACTGATG
CCTCCGTGTAAGGGGGATTTCTGTTTCATGGGGTAATGATACCGATGAAACGAGAGAG
GATGCTCACGATACGGGTTACTGATGATGAACATGCCCGGTTACTGGAACGTTGTGAG
GGTAAACAACTGGCGGTATGGATGCGGCGGGACCAGAGAAAAATCACTCAGGGTCAAT

GCCAGCGCTTCGTTAATACAGATGTAGGTGTTCCACAGGGTAGCCAGCAGCATCCTGC
 GATGCAGATCCGGAACATAATGGTGCAGGGCGCTGACTTCCGCGTTTCCAGACTTTAC
 GAAACACGGAAACCGAAGACCATTTCATGTTGTTGCTCAGGTCGCAGACGTTTTTGCAGC
 AGCAGTCGCTTACGTTTCGCTCGCGTATCGGTGATTCATTCTGCTAACCAGTAAGGCA
 ACCCCGCCAGCCTAGCCGGGTCTCAACGACAGGAGCACGATCATGCGCACCCCGTGGC
 CAGGACCCAACGCTGCCCGAAATT



Map of pMAL-c5x- CCL5Mut-HIGH

Gene (CCL5Mut-HIGH) was cloned into the *Bam*H1 and *Eco*R1 cloning site of pMAL-c5X as shown. Map created using PlasMapper. (Dong *et al.*, 2004).

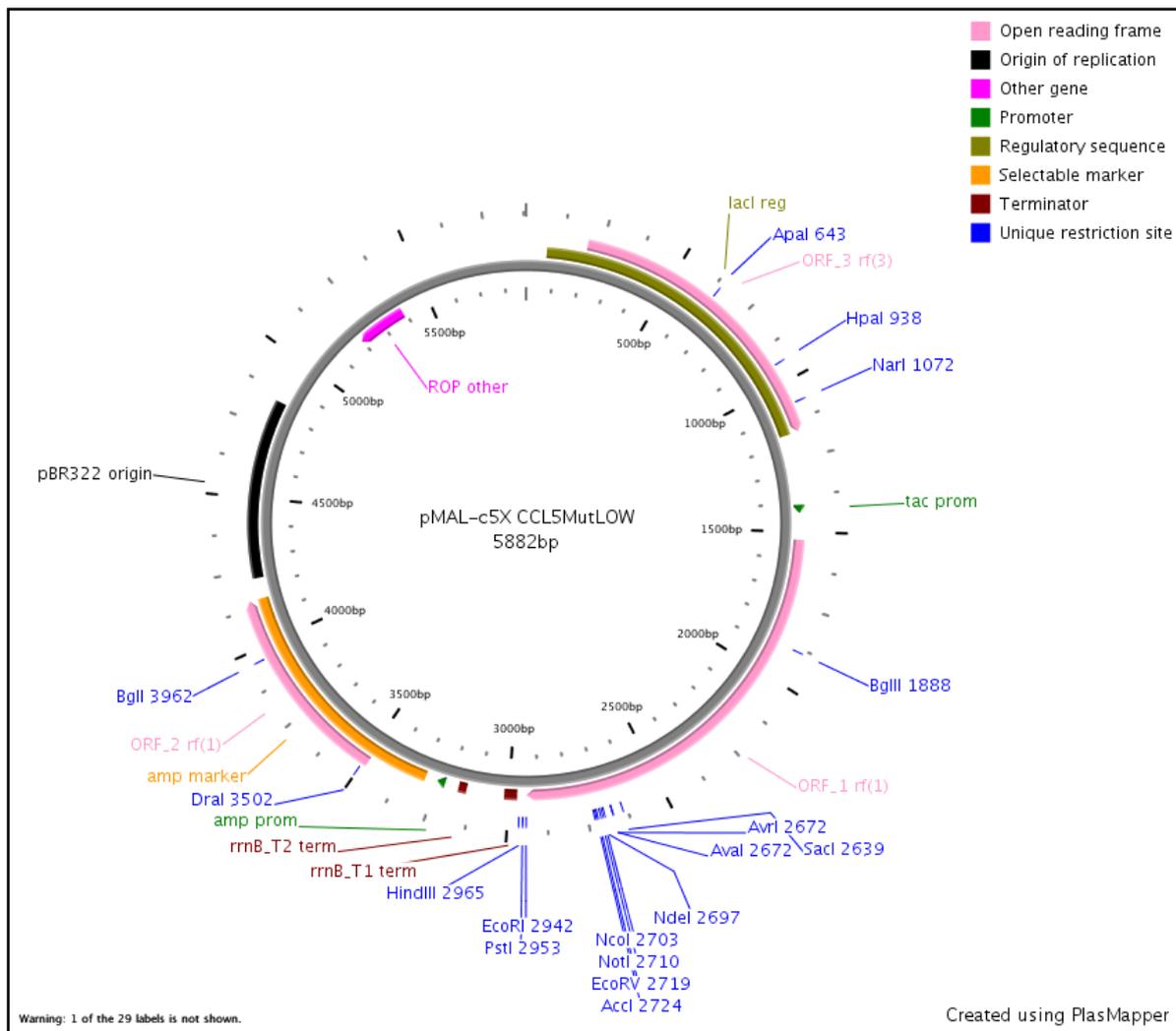
Number 8

Cloning vector pMAL-c5X-BamHI- CCL5MutLOW-EcoRI complete sequence

```
CCGACACCATCGAATGGTGCAAAACCTTTTCGCGGTATGGCATGATAGCGCCCGGAAGA
GAGTCAATTCAGGGTGGTGAATGTGAAACCAGTAACGTTATACGATGTCGCAGAGTAT
GCCGGTGTCTCTTATCAGACCGTTTCCC GCGTGGTGAACCAGGCCAGCCACGTTTCTG
CGAAAACGCGGGAAAAAGTGGAAAGCGGCGATGGCGGAGCTGAATTACATTCCCAACCG
CGTGGCACAACAACCTGGCGGGCAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGT
CTGGCCCTGCACGCGCCGTCGCAAATGTGCGCGGCGATTAAATCTCGCGCCGATCAAC
TGGGTGCCAGCGTGGTGGTGTGCGATGGTAGAACGAAGCGGCGTCAAGCCTGTAAAGC
GGCGGTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATCATTA ACTATCCGCTG
GATGACCAGGATGCCATTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTTC
TTGATGTCTCTGACCAGACACCCATCAACAGTATTATTTTTCTCCCATGAAGACGGTAC
GCGACTGGGCGTGGAGCATCTGGTCGCATTGGGTCACCAGCAAATCGCGCTGTTAGCG
GGCCCATTAAGTTCTGTCTCGGCGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCA
CTCGCAATCAAATTCAGCCGATAGCGGAACGGGAAGGCGACTGGAGTGCCATGTCCGG
TTTTCAACAAACCATGCAAATGCTGAATGAGGGCATCGTTCCC ACTGCGATGCTGGTT
GCCAACGATCAGATGGCGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGGCTGCGCG
TTGGTGC GGATATTTTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATAT
CCCGCCGTTAACCACCATCAAACAGGATTTTCGCTGCTGGGGCAAACCAGCGTGGAC
CGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGAAGGGCAATCAGCTGTTGCCCGTCT
CACTGGTGA AAAAGAAAACCACCCTGGCGCCCAATACGCAAACCGCCTCTCCCCGCGC
GTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAG
TGAGCGCAACGCAATTAATGTAAGTTAGCTCACTCATTAGGCACAATTTCTCATGTTT
ACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCGTCAGGCAGCCATCGGA
AGCTGTGGTATGGCTGTGCAGGTCGTAATCACTGCATAAATTCGTGTGCTCAAGGCG
CACTCCC GTTCTGGATAATGTTTTTTGCGCCGACATCATAACGGTTCTGGCAAATATT
CTGAAATGAGCTGTTGACAATTAATCATCGGCTCGTATAATGTGTGGAATTTGTGAGCG
GATAACAATTTACACAGGAAACAGCCAGTCCGTTTAGGTGTTTTTCACGAGCAATTGA
CCAACAAGGACCATAGATTATGAAAATCGAAGAAGGTAAACTGGTAATCTGGATTAAC
GGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGATAACCG
GAATTAAGTCAACGTTGAGCATCCGATAAACTGGAAGAGAAATCCCACAGGTTGC
GGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTAC
GCTCAATCTGGCCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAGCTGT
ATCCGTTTACCTGGGATGCCGTACGTTACAACGGCAAGCTGATTGCTTACCCGATCGC
TGTTGAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACC
TGGGAAGAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGA
TGTTCAACCTGCAAGAACCGTACTTCACCTGGCCGCTGATTGCTGCTGACGGGGGTTA
TGCGTTCAAGTATGAAAACGGCAAGTACGACATTAAGACGTGGGCGTGGATAACGCT
GGCGCGAAAGCGGGTCTGACCTTCCTGGTTGACCTGATTA AAAACAAACACATGAATG
CAGACACCGATTACTCCATCGCAGAAGCTGCCTTTAATAAAGGCGAAACAGCGATGAC
CATCAACGGCCCCTGGGCATGGTCCAACATCGACACCAGCAAAGTGAATTATGGTGT
ACGGTACTGCCGACCTTCAAGGGTCAACCATCAAACCGTTTCGTTGGCGTGTGAGCG
CAGGTATTAACGCCGCCAGTCCGAACAAAGAGCTGGCAAAGAGTTCCCTCGAAAAC TA
TCTGCTGACTGATGAAGGTCTGGAAGCGGTTAATAAAGACAAACCGCTGGGTGCCGTA
GCGCTGAAGTCTTACGAGGAAGAGTTGGTGAAAGATCCGCGTATTGCCGCCACTATGG
AAAACGCCCGAGAAAGGTGAAATCATGCCGAACATCCCGCAGATGTCCGCTTTCTGGTA
TGCCGTGCGTACTGCGGTGATCAACGCCGCCAGCGGTGCTCAGACTGTCGATGAAGCC
```

CTGAAAGACGCGCAGACTAATTCGAGCTCGAACAACAACAATAACAATAACAACA
ACCTCGGGATCGAGGGAAGGATTTACATATGTCCATGGGCGGCCGCGATATCGTTCGA
CGGATCCAGTCCCTATTTCGTCTGACACAACGCCCTGTTGCTTCGCCTACATCGCTCGC
CCCCTGCCTCGCGCACATATTAAGGAATACTTTTACACTTCTGGGAAATGCAGCAACC
CGGCAGTTGTGTTTGTGACTCGCAAAAACCGTTCAGGTTTGTGCAAATCCTGAAGCCGC
TTGGGTAGCCGAATACATCAATTCGCTGGAAATGTCATAGGAATTCCTGCAGGTAAT
TAAATAAGCTTCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTTCGTTTTAT
CTGTTGTTTGTTCGGTGAACGCTCTCCTGAGTAGGACAAATCCGCCGGGAGCGGATTGA
ACGTTGCGAAGCAACGGCCCCGGAGGGTGGCGGGCAGGACGCCCGCCATAAACTGCCAG
GCATCAAATTAAGCAGAAGGCCATCCTGACGGATGGCCTTTTTGCGTTTTCTACAAACT
CTTTCGGTCCGTTGTTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGAC
AATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACA
TTTCCGTGTGCCCCTTATTCCCTTTTTTGCGGCATTTCCTTCTGTTTTTGTCTCAC
CCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGACAGTTGGGTGCACGAGTGGGTTA
CATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTTCGCCCCGAAGAAGT
TTCCCAATGATGAGCACTTTTAAAGTCTGCTATGTGGCGCGGTATTATCCCGTGTG
ACGCCGGGCAAGAGCAACTCGGTCCGCCATACACTATTCTCAGAATGACTTGGTTGA
GTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGC
AGTGTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCG
GAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACCTGCCT
TGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACG
ATGCCTGTAGCAATGGCAACAACGTTGCGCAAACCTATTAACCTGGCGAACTACTTACTC
TAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACT
TCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTTATTGCTGATAAATCTGGAGCCGGTGAG
CGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCG
TAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGC
TGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATAT
ATACTTTAGATTGATTTCCCTTAGGACTGAGCGTCAACCCCGTAGAAAAGATCAAAGGA
TCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAACAC
CGCTACCAGCGGTGGTTTTGTTTGCCGGATCAAGAGCTACCAAACCTTTTTTCCGAAGGT
AACTGGCTTCAGCAGAGCGCAGATAACAAATACTGTCCTTCTAGTGTAGCCGTAGTTA
GGCCACCACTTCAAGAACTCTGTAGCACCCTACATACCTCGCTCTGCTAATCCTGT
TACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGCTTACCAGGGTTGGACTCAAGACG
ATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTTCGTGCACACAGCCC
AGCTTGGAGCGAACGACCTACACCGAACTGAGATACTACAGCGTGAGCTATGAGAAA
GCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCCG
AACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCTGGTATCTTTATAGTCCT
GTCGGGTTTTGCCACCTCTGACTTGAGCGTCGATTTTTTGTGATGCTCGTCAGGGGGGC
GGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCTGGCCTTTTGTG
GCCTTTTGTCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATT
ACCGCCTTTGAGTGAGCTGATAACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGT
CAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTG
CGGTATTTACACCCGCATATAAGGTGCACTGTGACTGGGTTCATGGCTGCGCCCCGACA
CCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTAC
AGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTACACCGTCATCAC
CGAAACGCGCGAGGCAGCTGCGGTAAGCTCATCAGCGTGGTTCGTGCAGCGATTACA
GATGTCTGCCTGTTTCATCCGCGTCCAGCTCGTTGAGTTTTCTCCAGAAGCGTTAATGTC
TGGCTTCTGATAAAGCGGGCCATGTTAAGGGCGGTTTTTTCTGTTTGGTCACTGATG
CCTCCGTGTAAGGGGGATTTCTGTTTCATGGGGTAATGATACCGATGAAACGAGAGAG
GATGCTCACGATACGGGTTACTGATGATGAACATGCCCGGTTACTGGAACGTTGTGAG
GGTAAACAACTGGCGGTATGGATGCGGCGGGACCAGAGAAAAATCACTCAGGGTCAAT

GCCAGCGCTTCGTTAATACAGATGTAGGTGTTCCACAGGGTAGCCAGCAGCATCCTGC
 GATGCAGATCCGGAACATAATGGTGCAGGGCGCTGACTTCCGCGTTTTCCAGACTTTAC
 GAAACACGGAAACCGAAGACCATTTCATGTTGTTGCTCAGGTCGCAGACGTTTTTGCAGC
 AGCAGTCGCTTACGTTTCGCTCGCGTATCGGTGATTCATTCTGCTAACCAGTAAGGCA
 ACCCCGCCAGCCTAGCCGGGTCTCAACGACAGGAGCACGATCATGCGCACCCCGTGCC
 CAGGACCCAACGCTGCCCGAAATT



Map of pMAL-c5x- CCL5Mut-LOW

Gene (CCL5Mut-LOW) was cloned into the *Bam*H1 and *Eco*R1 cloning site of pMAL-c5X as shown. Map created using PlasMapper. (Dong *et al.*, 2004).