

**Expression of T4 Lysozyme gene in *Escherichia coli* and
*Trichoderma reesei***

A thesis presented to
The Faculty of Graduate Studies
of
Lakehead University
by
ALAA AL-HAZMI

In partial fulfillment of requirements
for the degree of
Master of Science in Biology

December 11, 2012

Abstract

The T4 lysozyme, from bacteriophage T4, has the ability to lyse bacterial cell walls; thus it has attracted immense attention as an antimicrobial agent. Nonetheless, the production of T4L cannot meet our needs. All of the work presented here in this thesis, focuses on the potential exploitation of synthetic gene-coding T4L protein to improve the economic feasibility of antimicrobial agent production by producing a natural antimicrobial agent. Synthetic *t4l* was constructed to enhance the production of recombinant T4L for industrialization purposes after its successful expression in the two hosts. Therefore, comparing the wild type and synthetic *t4l* from nucleotide sequences until protein 3D structure was unavoidable. Two T4L genes that have been studied are the wild type and the synthetic one. The later, new synthetic (*t4l*) gene was produced by using codon-based modification with alternation of triplet nucleotides that encodes two specific amino acid residues at position 6 and 12 in a polypeptide chain of T4L protein. Wild type and synthetic *t4l* have variants in the size of nucleotides, 459 bp and 491 bp, respectively, but both have equal amino acids size, 164 a.a. The comparison of 3D protein structure showed us how both protein structures have structural similarity by calculating RMSD; 0.17 for backbone coordinate and 0.14 for carbon alpha coordinate. Also, structure sequence identity was 98.87% by using the alignment of editor Jalview. Additionally, developed T4L gene with codon-based modification has been cloned into three vectors with different promoters in order to be expressed in selected strain of *E. coli* then *T. reesei*. Expression of a synthetic *t4l* by using *E. coli* system was by transforming of the three plasmids-bearing-synthetic *t4l* separately, into two strains of *E. coli*, JM109 and BL21(DE3)pLysS. The later had been tested for its sensitivity and/or resistance to the lysis action of lysozyme enzyme before the transformation-taking place by using agar halo assay. Resistance of BL21(DE3)pLysS strain to lysozyme activity has been confirmed with our halo assays, which keep this strain to be a good choice for expression of this lysis enzyme. Also, the results of both enzyme digestion and PCR showed that the synthetic *t4l* had successfully transformed into selected *E. coli* strain. Furthermore, the expression of that gene had been showed on SDS-PAGE assay with an apparent molecular mass of approximately 18 kDa. However,

determine and measure the recombinant T4L production is still needs more research to work on it. Additionally, *T. reesei* has the capacity to produce extracellular enzymes in very large amounts. Invention relates method for efficiently expressing and producing recombinant protein of T4L from *T. reesei* QM6a by using protoplast-based transformation method. T4 lysozyme synthesis operon (*t4l*) cloned in three different plasmids was successfully integrated in *T. reesei* QM6a genome, to form recombinant strains, termed as ST4L1, ST4L2, and ST4L3. T4 lysozyme encoded by the *t4l* was successfully expressed in two of our recombinant strains, which are ST4L2 and ST4L3, as confirmed by SDS-PAGE. We successfully constructed a recombinant *T. reesei* strain expressing exogenous gene, specifically *t4l* that originated from phage T4 and integrated into the genome. Recombinant T4L was stably produced by those strains with a selected antibiotic supplementation. We develop a novel system by which foreign gene can be successfully inserted into the chromosome of *T. reesei*, where transcription can be driven by *cbhI* promoter, with the need for addition of inducer that is carbon source. T4 lysozyme production by this system was stable and the system may be useful in biotechnological fermentation processes. Success of such a technology will facilitate production of T4L for industrialization and medical purposes.

Lay Summary

Alexander Fleming discovered the antibacterial action of lysozyme in 1922. Since then, scientists have been increasingly studying lysozyme and determining further applications related to medical and/or industrial fields, which ended up with positive evaluations from World Health Organization and Food Drug Administration. In the current study, I worked on one of the most active lysozymes in sterilization, which is T4 lysozyme (T4L). Likewise, as an essential factor in the advancement of genetically engineered DNA techniques has been the development of genetic transformation systems. One of the significant challenges of this study is the development of a genetic transformation system for a species in which none exists. Collecting and applying methods that have been successfully applied to other species and determining how best to apply these methods to fungal organism in this study: *Trichoderma reesei*, was the major challenge. Successfully, modified gene encoding T4L, which studied bioinformatically and expressed in *E. coli* system, was introduced into *Trichoderma reesei*'s genome; moreover, expression of that gene developed a novel strain. However, T4L is an enzyme that is essential for protection against bacterial infection. The success of such technology developments will facilitate the production of T4L for industrialization purposes.

Acknowledgements

I would like to express my sincere appreciation and deepest gratitude to my supervisors Dr. Wensheng Qin for his enthusiasm in research, effective guidance, and continuous encouragement and support throughout this thesis. His guidance and patience was most helpful particularly when this project research experienced any difficulties. Completely, I would like to show my great appreciation for his help in my study, research, and life during the time period of this thesis work in Canada.

Furthermore, I am full of gratitude to the supervisory committee members; Dr. Kam Leung and Dr. Zi-Hua Justin Jiang, and external examiner Dr. Tim Xing for donating their valuable time and effort by lending their knowledge and support to my thesis. I want to extend heartfelt thanks to many other professors and the graduate students for their invaluable comments and their great patience in my graduate studies at Lakehead University.

I want to thank many colleagues, Dr. Miranda Maki and Dr. YuanYuan, just to name a few, for their help in some experiments, consultation and comments on the work, and preparation of the thesis. Special thank to Dr. Xi Chen for the donation of the synthetic *t4l* gene, without which this work could not have been possible. I would also like to thank Dr. Wely Floriano, the instructor of Advanced Bioinformatics course, for her support and generous effort along with valued inputs, advices and helpful suggestions throughout the course and the applied bioinformatics study.

Finally, financial support provided by Jazan University through Saudi Arabian Cultural Bureau in Canada, as well as research funding are highly appreciated.

Table of Contents

Table of Contents	V
List of Tables	VII
List of Figures.....	VIII
Chapter I: Introduction.....	1
Abstract	1
1.1. History	1
1.2. Biomolecular Structure and Function	3
1.2.1 Biological and Molecular Functions of Lysozyme	3
1.2.2 The Structure of Lysozyme	4
1.2.3 Lysozyme in Different Species.....	6
1.2.3.1 T4 Lysozyme	7
1.2.3.2 Hen Egg White Lysozyme.....	8
1.2.3.3 Human Lysozyme.....	8
1.2.3.4 Lysozyme in Other Vertebrates.....	10
1.2.4 Bioengineering and Protein Modifications	10
1.3. Lysozyme Applications.....	13
1.3.1 Medical Applications of Lysozyme.....	13
1.3.2 Industrial Applications of Lysozyme.....	14
1.3.2.1 Antimicrobial Food Packing.....	15
1.3.2.2 Gram Positive/Negative Bacteria Versus Sensitivity/Resistance to Lysozyme.....	17
1.3.2.3 Mechanism of Resistance to Lysozyme Activity	18
1.3.2.4 Inhibitory Activity of Lysozyme Against Microorganisms Other than Bacteria	21
1.4. Lysozyme Detection.....	21
1.4.1 Factors Influencing Lysis Efficiency	22
1.4.2 Agar Plate Bioassays	22
1.4.3 Activity in Media Measurements (Turbidimetric).....	23
1.4.4 Thermostability of Lysozyme on SDS-PAGE	25
1.4.5. Immunoassays	26
1.4.6. Biosensor.....	28
1.4.7 Flow injection Chemiluminescence System.....	29
Thesis objectives	30
Figure and Table legend.....	32
References	34
Chapter II: Computational Analysis of T4 Lysozyme.....	39
Abstract	39
2.1. Introduction	40
2.1.1. Approach	43
2.2. Methodology	43
2.2.1 VNC Software	43
2.2.2. Exhaustive comparison between two versions of T4 lysozyme	44
2.2.2.1. Retrieving nucleotide and amino acid sequence	44
2.2.2.2. Database similarity and multiple sequence alignment.....	45
2.2.2.3. Motifs profile.....	46
2.2.2.4. Analyzing protein 3D structure.....	47
2.2.3. Comparison of T4 lysozyme 3D protein structures	48
2.3. Results and Discussion.....	48
2.3.1. Retrieving nucleotide sequence and protein translation	48
2.3.2. Protein database similarities and multiple sequence alignment.....	49
2.3.2.1 EMBOSS package analysis.....	49
2.3.3. Motifs analysis	50
2.3.4. Analyzing T4 lysozyme 3D protein structure	51

2.3.5. Comparison of T4 lysozyme 3D protein structures	52
Conclusion.....	52
Figure and Table legend.....	54
References	62
Chapter III: Expression of a Synthetic T4 Lysozyme Gene in <i>Escherichia coli</i>	63
Abstract	63
3.1. Introduction	63
3.2. Materials and Methods.....	65
3.2.1. Synthetic T4 Lysozyme Gene	65
3.2.2. Bacterial Strains and Media.....	66
3.2.3. Halo Assay for Analysis of Host Strains.....	66
3.2.4. Electrotransformation of <i>E. coli</i> JM109 and BL21(DE3)pLysS	67
3.2.5. PCR Amplification of Synthetic <i>t4l</i>	68
3.2.6. Expression of T4 Lysozyme from Transformants and SDS-PAGE.....	69
3.3. Results and Discussions	70
3.3.1. Halo Assay.....	70
3.3.2. Synthetic <i>t4l</i> amplification	71
3.3.3. Detection of Recombinant T4 Lysozyme.....	71
Conclusions.....	72
Figure legend.....	73
References	75
Chapter IV: Expression of T4 Lysozyme gene in <i>Trichoderma reesei</i>	77
Abstract	77
4.1. Introduction	77
4.2. Materials and Methods.....	80
4.2.1. Strains, Media, and Cultivation Conditions.....	80
4.2.2. Antifungal Plate Assay	80
4.2.3. Protoplasts Preparation and Genetic Transformation.....	81
4.2.4. Confirmation of Positive Clones.....	82
4.2.4.1. Genomic DNA Isolation and <i>st4l</i> Amplification.....	82
4.2.4.2. Transformants Identification and Sequencing.....	83
4.2.5. Detection of Recombinant T4 Lysozyme.....	83
4.3. Results and Discussions	84
4.3.1. Strain optimization and Halo Assay.....	84
4.3.2. Production of Protoplast and Genetic Transformation	85
4.3.3. Confirmation of Positive Clones	86
4.3.3.1. Characterization of T4 Lysozyme Transformants.....	86
4.3.3.2. Positive clone Identification and Relevancy.....	87
4.3.4. Detection of Recombinant T4 Lysozyme Production.....	87
Conclusions.....	88
Figure and table legend	90
References	94
Chapter V: General Discussion and Future Directions.....	96
References	103

List of Tables**Chapter I: Introduction**

Table 1	33
Table 2	33
Table 3	33

Chapter II: Computational Analysis of T4 Lysozyme

Table 1	58
Table 2	58
Table 3	59
Table 4	59
Table 5	60
Table 6	60
Table 7	61
Table 8	61

Chapter IV: Expressing and Producing T4 Lysozyme through Recombinant *Trichoderma reesei*

Table 1	93
---------------	----

List of Figures**Chapter I: Introduction**

Figure 1.....	32
Figure 2.....	32
Figure 3.....	32

Chapter II: Computational Analysis of T4 Lysozyme

Figure 1.....	54
Figure 2.....	54
Figure 3.....	55
Figure 4.....	55
Figure 5.....	56
Figure 6.....	57
Figure 7.....	57

Chapter III: Expression of a Synthetic T4 Lysozyme Gene in *Escherichia coli*

Figure 1.....	73
Figure 2.....	73
Figure 3.....	73
Figure 4.....	74
Figure 5.....	74
Figure 6.....	67

Chapter IV: Expressing and Producing T4 Lysozyme through Recombinant *Trichoderma reesei*

Figure 1.....	90
Figure 2.....	90
Figure 3.....	90
Figure 4.....	91
Figure 5.....	91
Figure 6.....	92

List of Abbreviations

°C	Celsius degree
bp	Base pair
ddH ₂ O	Double-distilled water
DNA	Deoxyribonucleic acid
dNTPs	deoxyribonucleoside triphosphates
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethyl diphosphate tetra acetic acid
et al.	et alia (and other; Latin)
g	gram or relative centrifugal force (RCF)
h	hour
Kb	kilo bases
kDa	kilo Dalton
L, l	litre(s)
LB	Luria-Bertani
M	molar(s)
<i>M. luteus</i>	<i>Micrococcus luteus</i>
mg	milligram(s)
min	minute(s)
ml	millilitre(s)
mM	millimolar(s)
MW	molecular weight
OD	optical density
PCR	polymerase chain reaction
PEG	polyethylene glycol
PG	peptidoglycan
pH	negative logarithm of hydrogen ion concentration
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
Sec	second(s)
<i>T. reesei</i>	<i>Trichoderma reesei</i>
<i>t4l</i>	T4 lysozyme gene
U1	micro litre(s)
UV	Ultraviolet
V	voltage(s)

Chapter I: Introduction

Abstract

Lysozyme has attracted immense attention as an antimicrobial agent because of its ability to lyse the bacterial cell wall. Further, it is found in a wide variety of body fluids and in cells of the innate immune system. Lysozyme can act as muramidase or as a Cationic Antimicrobial Peptide (CAMP). This literature review covers a fundamental introduction to lysozyme in terms of biological and molecular structure and functions. Based on enzyme nomenclature, the lysozyme is classified as glycosylases under hydrolases. Therefore, the structure of this enzyme is discussed here to better understand its domains, the interface between domains, ligand binding domains, and active sites. Additionally, several studies have attempted to generate a modified lysozyme gene, either for large-scale production or that which is more suitable for industrialization purposes. Nonetheless, the production of recombinant lysozyme has not yet met our needs. Bioengineering and protein modifications that have been applied in other studies have been outlined. Additionally, lysozymes in different species have been summarized. Also, mechanisms of lysozyme-resistance with few bacteria have been discussed. As lysozyme has many applications in medical and industrial principles, it is important to review and study some of these applications. Some studies have used different methods for lysozyme detection; therefore, some of those methods are considered in this review. Ultimately, this type of review gives a basic understanding for researchers; likewise, it is considered a valuable starting point.

Key words: *Lysozyme Biomolecular Structure and Functions, Medical and Industrial Applications, Detection*

1.1. History

Lysozyme (1, 4-*N*-acetylmuramidase, E.C.3.2.1.17) is a small cationic protein first reported by Laschtschenko in 1909 (Burgess, 1973). Lysozyme's discovery, however, is attributed to Alexander Fleming. Fleming's report stated that lysozyme was a "powerful bacteriolytic element found in human

tissues and secretions” (Fleming, 1922). Fleming was at the time suffering from a cold and he is reported to have allowed drops of his nasal secretions to fall onto an agar culture plate, which was thickly colonized with bacteria. The plates were incubated at a certain temperature for a number of hours; bacteria grew extensively. There was complete radial inhibition beyond the nasal mucus on the plate. This experiment was later termed the lysoplate. Fleming concluded that the nasal secretions contained an enzyme capable of bacterial lysis; it was later named lysozyme.

With the success of his preliminary experiment, Fleming continued to work with lysozyme, testing its antibacterial properties with several different bacteria. Fleming reported the discovery of a small round bacterium particularly vulnerable to the effects of lysozyme. This bacterium was named *Micrococcus lysodeikticus* (now referred to as *Micrococcus luteus*) due to its ability to display lysis (Fleming, 1922). Fleming diluted the nasal secretions in saline and added them to a thick suspension of the *M. luteus*. Within minutes of incubation at 37°C the opaque bacterial solution had cleared. This experiment was later termed the Turbidimetric test (Fleming & Allison, 1927). Fleming (1922) is quoted as saying his experiments “clearly demonstrated the very powerful inhibitory and lytic action which nasal mucus has upon *Micrococcus lysodeikticus*.” Fleming also showed increased levels of lysozyme in patients with pyogenic infections (pus producing) such as meningitis, the first indication that lysozyme could be a marker for sepsis.

Since Fleming’s initial discovery in 1922 of lysozyme in tears and nasal secretions, a great deal of literature has accumulated on its structure, function, genetics, biosynthesis, regulation, enzyme activity and properties. Lysozyme is one of the most important factors of innate immunity, possessing antibacterial, antiviral, antitumor and immune modulatory activities. There has been interest in lysozyme as a “natural” antibiotic. Lysozyme is a unique enzymatic in that it exerts not only antibacterial activity, but also antiviral, anti-inflammatory, anticancer and immunomodulatory activities (Helal *et al.*, 2010). The ordinary lysozyme, which has already been industrialized, originates from chicken egg whites, while T4 lysozyme (T4L) originates from phage T4. Until now, there has been no T4L industrialization globally

(“BRI of CAAS,” 2011, p. 4). Success of an advanced technology will facilitate production of T4L for industrialization and medical purposes.

1.2. Biomolecular Structure and Function

1.2.1 Biological and Molecular Functions of Lysozyme

Lysozyme is an outstanding bacteriolytic enzyme that has the ability to hydrolyse glycosidic bonds of 1,4-beta-linkages between *N-acetylglucosamine (NAG)* and *N-acetylmuramic acid (NAM)* in peptidoglycan (PG), which is form the cell walls of prokaryotes (e.g., bacteria) (Mir, 1977; Akinalp *et. al.*, 2007) (Figure 1). In addition, its lysis activity has also been reported to inhibit viruses and eukaryotes, including fungi and parasites, in the absence of typical PG in their cell walls. However, the lysis of yeast and mold has been explained by the presence of an important component of their cell walls called chitin. Indeed, chitin has the same β -(1-4) glycosidic bonds as that of the bacteria PG, except that the chitin bond links two *NAG* residues rather than *NAG* and *NAM*. In other words, the lysozyme also possesses chitinase activity. Furthermore, the inhibition of *Entamoeba histolytica* by lysozyme can be explained by the presence in its membrane of lipopeptidophosphoglycan (LPPG), which can react with the enzyme in a similar manner as the PG (Benkerroum *et. al.*, 2008).

Lysozyme represents an important class of polysaccharide-hydrolyzing enzymes. Lysozyme hydrolyses chitin, a linear polymer found in insects, crustaceans, and fungi cell walls, while chitinase (EC 3.2.1.14) enzyme catalyzes the breakdown of chitin also. Furthermore, chitinase is a key enzyme in the plant defense system against fungal infection. Nonetheless, the chemical similarity between the two-polysaccharide substrates leads to the fact that some lysozymes can hydrolyse chitin; however, they do so less efficiently than their natural substrate and vice versa. Thus, some lysozymes could be considered good chitinases and, reciprocally, some chitinases can cleave PG, the natural substrate of lysozymes. Nevertheless, there is no obvious amino acid sequence similarity found between these two types of enzymes (Wohlkonig *et. al.*, 2010).

The antimicrobial activity of lysozyme has been extensively demonstrated *in vitro* or in physiological

fluids and secretions, including milk, blood serum, saliva, and urine. The antimicrobial activity of milk lysozyme, as part of the unspecific innate defense mechanism, is also well established. Specifically, it acts independently by lysing sensitive bacteria or acts as a component of complex immunological reactions to enhance the phagocytosis of bacteria by macrophages. Therefore, it contributes to the innate protection from milk spoilage after drawing (i.e., bacteriostasis period) (Benkerroum *et. al.*, 2008). Lysozyme also has a Cationic Anti-Microbial Peptide (CAMP) function that leads to bacterial death via the destabilization of the cytoplasmic membrane (Le Jeune *et. al.*, 2010).

Lysozyme is recognized to be non-dialyzable, soluble in water and a weak saline, insoluble in alcohol and ether, resistant to heat and desiccation, and stable at room temperature (Burgess, 1973), stable at acidic pH and labile at alkaline pH (Jolles, 1969). The understanding of the biochemistry and genetics of lysozyme enzymes, their phylogenetic relationships, and methods of estimation will make them more useful in a variety of processes in the future (Patil *et. al.*, 2000). Overall, lysozyme is one type of enzyme capable of hydrolyzing cell membrane polymers such as in fungal, some algae, and most bacteria; generally playing important roles in protecting plants and animals against infections and pathogens (Wohlkonig *et. al.*, 2010).

1.2.2 The Structure of Lysozyme

Lysozyme was one of the first enzymes for which the X-ray structure was determined. Furthermore, several classes of lysozymes have been identified based on their sequence similarities. The best known ones includes C-type (chicken-type), G-type (goose-type), and V-type (viral type) (Wohlkonig *et. al.*, 2010). Since the discovery of lysozyme, many investigators have studied this enzyme. For example, Wolff (1927) reported on its chemistry; Hoder (1931) on its relation to immunology; Anderson (1932) on its importance in vitamin A deficiency; Corper (1932) on its relation to tuberculosis; Meyer *et. al.* (1936) on its purification and properties; and Meyer (1936) and his group on the mechanism of its action (Daly, 1938).

Despite the variability in the amino acid composition and sequence of lysozyme molecules, amino acids of the catalytic center of active sites are well conserved to maintain the hydrolytic function of the enzyme from different sources. In particular, glutamic acid (Glu35) and aspartic acid (Asp52) residues are directly involved in the breakdown of the glycosidic bond between *NAG* and *NAM*. Their presence in the catalytic center is, thus, crucial for the hydrolytic activity of the enzyme. Glutamic acid acts as a proton donor through the free carbonyl group of its side chain, while aspartic acid acts as a nucleophile that generates a glycosyl-enzyme intermediate, which then reacts with a water molecule to give the product of hydrolysis and release the enzyme uncharged. However, the amino acid sequence of known lysozyme has revealed that aspartic acid is not consistently present in the active sites of lysozyme molecules. In contrast, substitution of glutamic acid has resulted in a complete inactivation of the enzyme, which confirms the critical role of this amino acid in the enzymatic activity of lysozyme, regardless of its origin or the class to which it belongs (Benkerroum, 2008).

Furthermore, for lysozyme action mechanisms, there is one polysaccharide ligand and two structural domains. Structural domains are primarily alpha helical and beta sheet. These two domains form a cleft and the ligand binds in the cleft between the two domains with polysaccharide (Alberts *et. al.*, 2010). In other words, the polysaccharide serves as a joint between the two active sites of lysozyme, Glu35 and Asp52. Following this process, lysozyme hydrolyses glycosidic bonds in that polysaccharide.

Since Fleming's report, the potential of lysozyme has attracted considerable interest over many years (Gorin *et. al.*, 1971). As an example, hen egg white lysozyme (HEWL) is very similar to human lysozyme, with 129 amino acids, while human lysozyme has ~148 amino acids and is three to four times more reactive than HEW lysozyme as tested with the turbidimetric test (Jolles, 1969; Lollike *et. al.*, 1995). However, due to its structural similarity, availability and inexpensiveness, HEWL has been used as a model for human lysozyme experiments, making it one of the most well studied enzymes (Lollike *et. al.*, 1995).

There are at least two types of glycosidase: those such as HEWL that have transglycosidase activity, and those that do not, such as T4 phage lysozyme (T4L). The mechanistic difference between these two types of enzymes is not well understood. Indeed, the transglycosylation reaction can be more effective than hydrolysis. In contrast, T4L has no detectable transglycosidase activity. Early structural studies showed that three different classes of lysozymes, typified by HEWL, T4L, and goose egg white lysozyme (GEWL), have overall similarities in three-dimensional structure, although their amino acid sequences are almost entirely different. The active sites of these enzymes have a glutamic acid that is highly conserved (Glu-35 in HEWL, Glu-11 in T4L, Glu-73 in GEWL). Otherwise, the active site residues show substantial variability. For example, the position of Asp-20 in T4L does not correspond to the “essential” Asp-52 of HEWL, and GEWL has no apparent counterpart at all to either of these acidic residues (Kuroki *et. al.*, 1999).

1.2.3 Lysozyme in Different Species

While lysozyme is widely distributed in prokaryotes and eukaryotes, it has distinct types based on its amino acid sequence, protein structure, and catalytic characters, although its catalytic function is still consistent (Akinalp *et. al.*, 2007; Dong *et. al.*, 2008). The wide variability in origin and structural, antigenic, chemical, and enzymatic properties of lysozyme molecules, requires them to be classified in different classes, or types. The conventional, or chicken-type (i.e., c-lysozyme), in which the lysozyme is derived from the egg white of a domestic chicken (*Gallus gallus*), is considered the most studied and best known. Although c-lysozyme is typically found in the egg white of birds, it is also purified from various tissues and secretions of mammals. Other types of lysozyme that are well-known include the g-type, derived from the egg white of the domestic goose (*Anser anser*); h-type lysozyme from plants; i-type from invertebrates; b-type from bacteria (*Bacillus*); and v-type from viruses (Benkerroum, 2008).

Fleming stressed that, while lysozyme is most active against nonpathogenic bacteria, it can attack pathogenic organisms in some parts of the body when allowed to act in full strength. In addition, he showed that lysozyme from different sources affect different organisms: *Pneumococcus* was attacked only

by the lysozyme of pus and *Streptococcus haemolyticus* only by the lysozyme of tears; *Staphylococcus* and *Streptococcus viridans* were affected more by the lysozyme of nasal mucus than by that of tears or pus. Nevertheless, other researchers worked with egg white lysozyme (which may act differently from lysozyme derived from body tissues) and they found that none of a number of *Staphylococcus albus* strains and *Staphylococcus aureus* tested, showed marked susceptibility to the lytic action of the lysozyme (Daly, 1938).

1.2.3.1 T4 Lysozyme

Phage T4 is a bacterial virus that infects only bacteria; specifically, it invades and controls bacterial metabolism, then terminates its incubation by lyse of infected-bacterial cell wall by T4L (Sulakvelidze *et al.*, 2001; Trun & Trempy, 2004). T4L is similar to the HEWL in the structure of its active sites and sites of attack on the PG cell wall. Surprisingly, T4L apparently lyses T4-infected cells by attacking the cell wall from the cytoplasmic side (Akinalp *et al.*, 2007). In other words, it could have active sites on both sides of the bacterial cell wall. T4L is a small, well-characterized protein with a molecular weight of approximately 18 kDa and dimensions of 5.4 x 2.8 x 2.4 nm (Bower *et al.*, 1998); it has been the subject of extensive structural and genetic studies (Poteete *et al.*, 1997). T4L is a single polypeptide chain with a single NH₂-terminal methionine residue and a single COOH-terminal leucine residue. The amino acid composition of the protein has been determined (Tsugita *et al.*, 1968).

In the 3D protein structure of T4L, the backbone of T4L has two-domain structures, the N-domain and C-domain (Baase *et al.*, 2010), which are linked by a single long α -helical chain. The relative placement of the N- and C-terminal domains forms a deep pocket where the ligand can bind. Ligand binding and release is associated with motions that are involved in opening and closing of this binding pocket relative to those domains as evidenced from experiments and computational studies (Ramanathan *et al.*, 2011). The native structure of this protein has been determined to 1.7Å resolution by X-ray crystallography (Poteete *et al.*, 1997). Systematic probing of the T4L structure has also shown that 74 of the 164 positions in its sequence are sensitive to single amino acid substitutions. That is, at least one single amino acid

substitution at one of these positions results in at least a 50-fold reduction in function. Not sparingly, the most critical amino acid residues in the protein have been found to be buried or solvent-exposed in the active site cleft (Poteete *et. al.*, 1997).

Although lysozymes from T4 phage and hen egg whites have similar catalytic activity, they have non-homologous amino acid sequences. A comparison of the 3D structure of the phage enzyme with that of hen egg whites has been determined in other studies. The 3D structures of HEW and T4L have also been seen to be quite different (Matthews & Remington, 1974); it is not clear whether their respective mechanisms of catalysis are related.

However, T4L, produced by the virus phage T4, consists of 164 amino acid residues with a molecular weight of 18,700 Da. This enzyme is devoid of disulfide bridges and the only occurring Cys residue is located at positions 48 and 97. Further, T4L consists of two domains, C- and N-terminal domains; the active site is located in the region between these domains (Malmsten, 2003).

1.2.3.2 Hen Egg White Lysozyme

Lysozyme enzyme isolated from hen egg white was the first enzyme submitted for X-ray crystallographic analysis and for which a detailed mechanism of action was proposed. It is used extensively as a model in protein chemistry, enzymology, protein crystallography, molecular biology and genetics, immunology, and evolution. The progress made by the study of these different aspects has been reviewed in other studies (Van Dael, 1998). However, the HEWL remains, by far, the richest source of this enzyme, with a concentration that ranges between 3400 and 5840 mg/l (Benkerroum, 2008). HEWL is a 14,600 Da large monomeric protein that consists of 129 amino acid residues. It has been classified as a “hard” protein because the adsorption of HEWL to solid surfaces does not involve any major conformational changes (Malmsten, 2003).

1.2.3.3 Human Lysozyme

Lysozyme has also been identified from humans in numerous exocrine secretions such as colostrum, genito-urinary, gastroduodenal, middle ear, upper and lower airway secretions, saliva, and tears

(Hinnrasky *et. al.*, 1990). This enzyme is also found in a wide variety of body fluids, such as respiratory secretions, and in cells of the innate system, including neutrophils, monocytes, macrophages, and epithelial cells (Le Jeune *et. al.*, 2010).

Human lysozyme is a key component of the innate immune system and recombinant forms of the enzyme represent promising leads in the search for therapeutic agents that are able to treat drug-resistant infections. In addition to its catalytic hydrolysis of cell wall PG, human lysozyme also exhibits catalysis-independent antimicrobial properties. The dual functionality results in a protein that attacks Gram-positive and Gram-negative bacterial pathogens; it has been shown to be the most effective cationic anti-pseudomonal agent in human airways fluids (Gill *et. al.*, 2011). Also, lysozyme has been found in high concentrations in human airway secretions collected from patients with chronic pulmonary diseases (Hinnrasky *et. al.*, 1990). In addition, one of the major proteins secreted by an established human colon adenocarcinoma was identified as lysozyme. Human colon adenocarcinoma cells synthesize and secrete several specific proteins, including remarkably large amounts of lysozyme. The lack of diagnostic or prognostic markers, as well as the refractory nature of human colon carcinoma to standard chemoimmuno- and radio-therapy, makes this an extremely relevant system when investigating the chemistry and biology of such secreted macromolecules (Fett *et. al.*, 1985).

Human lysozyme is a ~148 amino acid single polypeptide chain with a low molecular weight of around 14.6 kDa (Cabellero *et. al.*, 1999). Human lysozyme is present in the lysosomes of phagocytic cells, granulocytes and monocytes (Burgess *et. al.*, 1994). It is released as part of the non-specific immune response, and exists among cells of the blood, especially leukocytes. Lysozyme is found in all stages of the maturation of the myelocytic series, but not in the myeloblast, eosinophil or basophil (Davis, 1971). Monocytes contain large amounts of lysozyme, yet none is found in the lymphocytes. Moreover, in tissues lysozyme is mainly found in bone marrow, lungs, intestines, spleen, and kidneys. Lysozyme exists here due to the breakdown of neutrophilic granulocytes in these organs (Hansen *et. al.*, 1972). Tissue

macrophages release lysozyme into serum, nasal and lacrimal secretions along with various other bodily secretions.

1.2.3.4 Lysozyme in Other Vertebrates

In higher vertebrates, lysozymes are involved in a broad battery of defense mechanisms and embrace several actions such as opsonization, immune response potentiation, and restricted anti-viral and antineoplastic activities. Moreover, several studies have indicated that lysozyme may play a role in fish as a defense mechanism against infectious disease (Lie *et. al.*, 1989). In addition to their defense role, lysozymes also play a role in digestive tasks, for example a dominant gene for high lysozyme activity has been detected in cattle (Lie *et. al.*, 1989). Also, high levels of intestinal lysozyme occur in some species of house mice in which the enzyme may have assumed a digestive role. This trait appears in only one of two closely related lysozyme genes in the same species, resulting in a specific overexpression of lysozyme P in the small intestine and leaves a normal expression of lysozyme M in other tissues (Cross *et. al.*, 1988).

Furthermore, the milk of virtually all mammals contains lysozyme, either as a free soluble protein or within leucocytes and lysosomes. Although all lysozyme from milk has been reported as belonging to the c-type, it varies widely in terms of structure and physicochemical properties; this includes the folding and unfolding status, ability to bind calcium ions, stability to heat or pH and isoelectric point (Table 1). Based on lysozyme content in milk, human, equine, and canine, milks are the main representatives of high levels of lysozyme (averages 200 to 1330 mg/L group), while bovine, ovine, and caprine milks (Table 2) represent low levels of lysozyme (3000 to 6000 times less than the milks of the first group) (Benkerroum, 2008). Additionally, the concentration of soluble lysozyme in milk varies considerably from one species to another, as well as within the same species, depending on various factors such as the breed, stage of lactation, parturition, nutrition, udder health, and season of the year.

1.2.4 Bioengineering and Protein Modifications

The gene encoding lysozyme is cloned from mammalian tissues and secretions, insects, plants, protozoa, bacteria, and viruses and then expressed in some bacteria, fungi, yeast, and plants (Akinalp *et.*

al., 2007). For example, T4L is one of the most thoroughly characterized enzymes through site-specific mutagenesis; more than 2,000 single-amino acid substitutions have been done (Malmsten, 2003). Studies using T4L variants, which differed from each other by only one amino acid residue or more, provided data that clearly related a change in enzyme structure (Bower *et. al.*, 1998). Numerous variants of T4L have also been produced through site-directed mutagenesis, including a set of variants that differ by substitution of the isoleucine residue at position 3 (Ile 3). Since the Ile 3 residue contributes to a major hydrophobic core of T4L, any substitution, at this point, can significantly alter the enzyme's stability (Bower *et. al.*, 1998).

In another work, wild-type T4L enzyme and two site-directed mutants have been used. The site of mutation was position 3, which, in the wild-type enzyme, is occupied by an isoleucine (Ile3). Notably, one of the mutants increased thermal stability compared to the wild-type enzyme. This mutant had the Ile residue in position 3 that was replaced by Cysteine (Cys), which led to the formation of a disulfide bond with another Cys residue at position 97. Thermal unfolding of the Cys mutant revealed that the formation of the disulfide bond stabilized the mutant by 1.2 kcal/mol relative to the wild-type. However, in the other mutant, Ile in position 3 was replaced by the larger amino acid Tryptophan (Trp). This decreased the thermal stability of this mutant by 2.8 kcal/mol relative to the wild-type enzyme. This mutant is one of the least stable T4L mutants characterized (Malmsten, 2003). Therefore, this residue contributes to the major hydrophobic core of the C-terminal domain and helps link the N- and C-terminal domains.

In other studies, researchers examined the structure of a mutant T4L with a single amino acid substitution of 2.5 nm from the active site. The change (replacement of glutamate-128 by a lysine residue) caused very little change in the structure of the lysozyme molecule; however, the enzyme had only 4% of the catalytic activity as the native enzyme. These results suggest that glutamate-128 participates directly in substrate binding or catalysis. The glutamate residue is located on the C-terminal domain, an area that has no counterpart in the HEWL molecule. Therefore, it has been suggested that the role of this C-terminal

domain is to bind the peptide crosslink that connects neighbouring saccharide strands in the *E. coli* cell wall (Jones, 1984; Grutter & Matthews, 1982).

Generally, a major goal of protein engineering is to design proteins that demonstrate enhanced stability and activity. Among the physical forces that maintain the tertiary structure of proteins, disulfide bonds make a substantial contribution. However, the addition of new disulfides has not always increased stability. One aspect of this problem is the limited knowledge of the mechanism for which crosslinks such as disulfide bonds, stabilize, or destabilize proteins (Mastsumura *et. al.*, 1989).

From a historical perspective, one of the main reasons for undertaking structural studies of lysozyme was to take advantage of genetic screening for temperature-sensitive (TS) mutants. When the lysozyme project began, methods for the generation of substitutions by site-directed mutagenesis were yet to be developed. Also, the sequence changes in TS lysozymes were determined by classical protein chemistry rather than at the DNA level. For example, five TS T4Ls characterized by following approach included Arg96 → His (R96H), Met102 → Thr (M102T), Ala146 → Thr (A146T), Gly156 → Asp (G156D), and Thr157 → Ile (T157I). On average, these variants reduced the melting temperature of the protein by 10°C (Baase *et. al.*, 2010). In a similar study, a multiple mutant was engineered from T4L: I3L/S38D/A41V/A82P/N116D/V131A/N144D (PDB Code 189L) and was shown to increase T_m by 8.3°C (Zhang *et. al.*, 1995).

Currently, different methods have been tested to either select or design mutants that retain activity and have increased thermostability. These include the use of entropic effects, hydrophobic stabilization, helical propensity, salt-bridge interactions, metal binding, S-S bridges, methionine replacement, computational procedures, and genetic selection. In cases where disulfide bridges could be engineered into T4L, these methods have been very effective in increasing the stability of the protein. A good example is the Cys21-Cys142 bridge across the active site cleft of T4L. In this case, “hinge-bending” allowed the protein to adjust to accommodate the S-S bridge and the melting temperature increased by 11°C at pH 2 (Baase *et. al.*, 2010). Furthermore, the genetics of T4 phage lysozyme have been studied extensively and a

variety of mutant enzymes have been isolated and characterized. A previous study suggested that the residues, Asp 20, Glu 22, Glu 105, Trp 138, Asn 140, and Glu 141, are all essential for full catalytic activity on the grounds that changes in any of these amino acids drastically reduce the catalytic effectiveness of the enzyme (Matthews & Remington, 1974).

In contrast, the close proximity of the substituted residue Ile 3 to the active site Glu 11 may also contribute to the observed differences in activity among T4L variants because both residues are located on the same α -helix. The α -helix content of a protein in a solution is known to increase or decrease in response to changes in its environment (e.g., pH, ionic strength, and hydrophobicity of the solvent). A T4L variant that is forced to accommodate a relatively large, bulky tryptophan residue in position 3 may also make conformational adjustments in adjacent residues. If these effects are propagated along the α -helix, significant changes in the orientation of the active site may result in an accompanying decrease in catalytic activity (Bower *et. al.*, 1998).

Moreover, the loss of catalytic activity of the T4L variants might also relate to the degree of interference caused by mutation in the hinge region of the enzyme. These amino acid substitutions are located directly opposite the active site in an area of the molecule that is responsible for the opening and closing movement of the active cleft. As such, an enzyme would retain full activity only if the substrate were able to force its way past the “hinged” cleft and reach the active site (Bower *et. al.*, 1998).

1.3. Lysozyme Applications

1.3.1 Medical Applications of Lysozyme

Lysozyme has been credited with being an indicator in some hematological malignancy; such as acute myeloid leukemia. In addition, an elevated serum lysozyme level could be helpful in differentiating between the types of cancer that affect blood, bone marrow, and lymph nodes. However, lysozyme has been associated with different abnormalities than those associated with an elevated level of this protein. There could be a casual relationship, because activity increases in the gastric juice of patients with peptic ulcer and in the colonic sections of patients with ulcerative colitis and regional enteritis (Mir, 1977).

Lysozyme lends significant insight, as it is associated with any abnormality in humans who demonstrate an elevated level of this important enzyme.

Hence, deviations in lysozyme levels of blood and urine, for example, have been correlated to particular myeloid and renal abnormalities and lysozyme levels have subsequently been used to monitor the success of therapy. Although the lysozyme protein can be found in most tissues, studies at the cellular level have revealed high levels of lysozyme, specifically in phagocytic cells (including macrophages and granulocytes) and the Paneth cells (including the intestine and the proximal tubules of the kidney). While the kidney accumulates at least some of its lysozyme from the blood, endogenous lysozyme synthesis has been detected in mammalian myeloid cells (Cross *et. al.*, 1988). Additionally, lysozymuria has been reported, in various renal tubular disorders, as a manifestation of the disease. Furthermore, discrepancies could be caused by differences in methods, patient populations, and stages of disease (Mir, 1977). Future studies may reveal various subtypes of lysozyme based on more sophisticated physical and biochemical characteristics; these subtypes may reveal more about the cell functions and their sensitivity to antitumor drugs. Of note, it was reported that lysozyme is an effective agent for killing HIV *in vitro* (Song and Hou, 2003). Thus, lysozyme has potential for future applications in the medical field.

1.3.2 Industrial Applications of Lysozyme

Lysozyme enzyme is gaining importance for its biotechnological applications. Specifically, such applications are used in agriculture to control plant pathogens, for example. Its high microbicidal activity remains, by far, the main virtue explaining its high attention with scientists and industrial stakeholders for its practical applications in medical and food industries. Mainly, it is known for its antibacterial activity and has been used in food preservation. Although the egg white is the primary source for lysozyme production at an industrial scale, other sources, such as mammals' milk, should not be overlooked, as they may contain lysozyme molecules with specific properties that are not present in the conventional egg white lysozyme (Benkerroum, 2008). The success in employing lysozyme for different aspects depends on the supply of highly active preparations at a reasonable cost (Patil *et. al.*, 2000).

Chicken egg white lysozyme characterizes its use in oenology. While it has been used for several years in the pharmaceutical and food-processing industries, it has been recently added in some oenology processes (e.g. wine maturation). Moreover, lysozyme has recently been introduced to the wine industry to control contamination by lactic bacteria; in such industry, lysozyme has been permitted to the level of 100 mg/L (Benkerroum, 2008, Ferreira *et. al.*, 2011). Furthermore, from a biochemical perspective, lysozyme has an optimum activity between 40°C and 45°C and can remain active up to 62°C, which is entirely compatible with the temperature used for enzymatic hydrolysis. Also, with regards to the optimum pH value, the enzymatic activity of this protein is slightly disturbed within a pH value range between 3.5 and 7; however, it can remain active between 2 and 10 (Ferreira *et. al.*, 2011).

Lysozyme also presents the advantage of limiting the use of sulfur dioxide (SO₂), a product that is often used today to limit microbiological contamination in industrial processes. However, this process is expensive and has health-related side effects. Further, lysozyme plays a protective role in difficult fermentation. In fact, nitrogen deficiencies make alcoholic fermentation difficult with the risk of developing lactic bacteria that degrade fermented food products. However, the addition of lysozyme accommodates with this type of problem with high efficiency (Ferreira *et. al.*, 2011).

Additionally, lysozyme is used as a preservative in many food products that are likely to contain lactic bacteria such as cheese, tofu, or sake. The use of lysozyme in brewing processes has won attention and there are now many applications for lysozyme in the brewing stages. Lysozyme is effective against most lactic bacteria. At the practical level, lysozyme has found applications in food preservation and has already been used successfully as an antimicrobial in many foods, particularly cheese. This has received positive evaluations from international regulatory agencies such as the World Health Organization (WHO) and the Food and Drug Administration (FDA) (Benkerroum, 2008).

1.3.2.1 Antimicrobial Food Packing

In order to control microbial growth in foods and improve their shelf life and safety, different antimicrobial agents are mixed with the initial food formulations or are applied to food surfaces by

dusting, dipping or spraying (Min and Krochta 2005). In recent years there has been greater interest in antimicrobial food packaging technologies due to increased food-borne microbial outbreaks caused by minimally processed fresh and refrigerated products. Traditionally, antimicrobial additives are mixed into initial food formulations to control microbial growth and extend shelf life. However, this strategy is not always effective because the protective ability of the antimicrobial agent ceases once it is neutralized in reactions and/or interactions in the complex food system and the quality of food degrades at an increased rate. In addition, the antimicrobial compound directly added into the food cannot selectively target the food surface where spoilage reactions occur more intensively. Antimicrobial packaging is an alternative method to overcoming these limitations because the agent slowly releases from the film onto the food surface during storage, hence maintaining the critical concentration necessary for inhibiting the microbial growth (De Roever 1998, Devlieghere *et. al.*, 2004).

To prepare antimicrobial films, different synthetic antimicrobial chemicals have been incorporated into packaging materials including organic or inorganic acids, metals, alcohols, ammonium compounds or amines (Appendini and Hotchkiss 2002; Suppakul *et. al.*, 2003). However, the increasing consumer health concern and growing demand for healthy foods have stimulated the use of natural biopreservatives such as antimicrobial enzymes and bacteriocins (Suppakul *et. al.*, 2003; Labuza & Breene, 1989). Lysozyme is one of the most studied biopreservatives for antimicrobial packaging applications (Han 2000; Quintavalla and Vicini 2002).

Antimicrobial packaging is one of the many applications of active packaging. It is a packaging system that is able to kill or inhibit spoilage and pathogenic microorganisms that are contaminating foods (Han 2000). Also, it is specifically designed to control growth of microorganisms, unlike conventional food packaging systems, which are used for shelf-life extension, quality maintenance, and safety assurances that could be achieved by various methods. Table 3 shows the application area, when lysozyme is used as an antimicrobial agent and packaging materials are used in antimicrobial food packaging.

Many researchers have studied lysozyme for biopreservation of foods because of its lytic activity on bacteria. Most of the studies related to lysozyme have been focused on HEWL because the enzyme is commercially purified from egg white (Chang *et. al.*, 2000). The antimicrobial activity of the enzyme is mainly effective against Gram-positive bacteria; however, because the PG layer is surrounded by a protective lipopolysaccharide (LPS) membrane, it is ineffective against some Gram-negative bacteria (Nakamura *et. al.*, 1991, Ibrahim *et. al.*, 1991).

With that being said, many researchers have been interested in increasing the antimicrobial spectrum of lysozyme. For instance, the combination of lysozyme with EDTA makes lysozyme highly effective on Gram-negative bacteria (Mecitoglu *et. al.*, 2006), and conjugates of lysozyme with dextran, galactomannan or xyloglucan have good antimicrobial activity against both Gram-positive and Gram-negative bacteria when applied in combination with mild heating at 50°C (Nakamura *et. al.*, 1992). Nonetheless, in several studies lysozyme has been incorporated as a preservative in many packaging materials to extend the shelf life of the foods (Appendini and Hotchkiss 1997; Buonocore *et. al.*, 2004; Mecitoglu *et. al.*, 2006).

1.3.2.2 Gram Positive/Negative Bacteria Versus Sensitivity/Resistance to Lysozyme

While PG in Gram-positive bacterial cell walls is freely accessible to lysozyme, on the contrary Gram-negative bacterial PG is shielded by the lipopolysaccharidic (LPS) layer of the outer membrane, some studies have suggested that resistance of bacteria to lysozyme is not exclusively related to the presence of the LPS layer. However, the presence of an outer membrane in Gram-negative bacteria does not provide absolute protection against the hydrolytic action of lysozyme. Sensitive Gram-negative bacteria have been described, along with mechanisms not hindered by LPS. In addition to the protective effect of LPS against the hydrolytic action of lysozyme, Gram-negative bacteria have recently been shown to use another strategy that involves specific protein-inhibitors with high affinity to lysozyme (Benkerroum, 2008).

Specifically, in 1927, Fleming and Allison were first to report resistant strains of *Streptococcus*. Fleming and Allison's work into the nature of the bacterial resistance continued. Several strains of

bacteria were tested with the tears from Fleming's patients. Fleming found lysozyme was not effective against a range of bacteria. These included *Haemophilus influenzae*, *Neisseria meningitidis*, *Escherichia coli*, *Streptococcus pneumoniae*, *Klebsiella pneumoniae* and group B *Streptococci*. From these studies, lysozyme resistance was characterized as an inability to break down bacteria that are encapsulated with a gelatinous polysaccharide layer. This dramatically limits the effectiveness of lysozyme to bacteria that have no capsule. Lysozyme is therefore of little value as a therapeutic agent against bacterial growth, but the release of lysozyme as a product of phagocytosis and white blood cell turnover may be a useful marker for sepsis (Fleming and Allison, 1927).

1.3.2.3 Mechanism of Resistance to Lysozyme Activity

The complexity of the composition and structure of peptidoglycan is well known (Le Jeune *et. al.*, 2010). The pathological effects of peptidoglycan are greatly enhanced by various modifications and substitutions to its basic composition and structure (Clarke & Dupont, 1992). While some bacterial species are sensitive to lysozyme, the majority of bacteria (like some important human pathogens such as *Staphylococcus aureus* or *Neisseria gonorrhoeae*, *E. faecalis*, and *Proteus mirabilis*) are resistant to lysozyme (Le Jeune *et. al.*, 2010; Herbert *et. al.*, 2007).

In fact, the exact mechanism of lysozyme resistance is not fully understood and may vary according to the bacterial strain or species (Benkerroum, 2008). Various mechanisms involved in this resistance have been well described in different species of bacteria. Such mechanisms of resistance have been suggested such as interference of lysozyme action by surface attachment polymers (e.g., capsular polysaccharides and teichoic acids) (Benkerroum, 2008). For example, a high degree of peptide cross-linkage (Bera *et. al.*, 2005; Herbert *et. al.*, 2007; Benkerroum, 2008), this is a modification of the net negative charge of the bacterial cell surface by adding positively charged residues genes, teichoic, and lipoteichoic acids, which help bacteria to avoid being killed by antimicrobial peptides or CAMP activity of lysozyme (Herbert *et. al.*, 2007; Le Jeune *et. al.*, 2010), and *N-acetylmuramic* acid residues being modified by a tetrapeptide composed of both L and D amino acids. In most cases, these peptide side chains are involved in the direct

cross-linking between adjacent glycan strands; however, in some instances, a more complicated arrangement exists. For example, a pentaglycine peptide bridges the neighboring peptide side chains within the peptidoglycan sacculus of *Staphylococcus aureus* (Clarke & Dupont, 1992); this is a modification of different sites of the PG structure by some type of enzyme (e.g., the peptidoglycan-specific *O*-acetyltransferase (*OatA*) of *S. aureus*) (Le Jeune *et. al.*, 2010). Similarly, it has been described for 11 Gram-positive and Gram-negative species (Bera *et. al.*, 2005), which results in *O*-acetylation of hexosamine residues of the cell wall PG (Benkerroum, 2008) and the N-acetylglucosamine deacetylase (*PgdA*) of *Streptococcus pneumonia*. These both prevent the binding of lysozyme to its substrate and contribute to the muramidase resistance. *N*-deacetylation of the acetamido group of the hexosamine residues (Benkerroum, 2008), incorporation of D-Aspartic acid in the bacterial PG cross bridge, as was demonstrated in *Lactococcus lactis*, and the production of protein-inhibitors specific to lysozyme, *N*-non-substituted glucosamine residues in the peptidoglycan, as in *Bacillus cereus* and *Streptococcus pneumonia* (Bera *et. al.*, 2005).

Moreover, evading the killing action of lysozyme may also consist of the production of lysozyme inhibitors such as the Streptococcal Inhibitor of Complement (SIC) in *streptococci*, the inhibitor of vertebrate lysozyme in *E. coli*, the periplasmic (or membrane) lysozyme inhibitor of c-type lysozyme (for example: Periplasmic Lysozyme Inhibitor of c-type lysozyme in *Salmonella enteritidis*, or Membrane-bound Lysozyme Inhibitor of c-type lysozyme in *E. coli*), which have protective functions (Le Jeune *et. al.*, 2010).

Over recent years, these mechanisms have been gaining attention in research. While there is a general agreement that surface attachment polymers and the degree of peptide cross-linking do not account for lysozyme-resistance, it is not clear whether there is only a single mechanism of resistance for all bacteria or if more specific mechanisms may be used by specific strains or species. However, the above mechanisms may be modulated by other factors that are not directly involved in lysozyme resistance. For example, the presence of teichoic acid or a high degree of peptide cross-linking, though shown not to have

an intrinsic effect on lysozyme resistance, has enhanced significantly the effect of *O*-acetylation on lysozyme resistance in *S. aureus* (Benkerroum, 2008). One such modification is the presence of acetyl moieties at the C-6 hydroxyl group of *N*-acetylmuramyl residues and some species of eubacteria, including some important human pathogens, such as *Neisseria gonorrhoeae*, *Proteus mirabilis*, and *Staphylococcus aureus*, which are known to possess *O*-acetylated peptidoglycan (Clarke & Dupont, 1992). *O*-acetylation of peptidoglycan occurs at the C-6 hydroxyl group of *N*-acetylmuramyl residues, producing corresponding 2, 6-diacetylmuramyl derivatives. The extent of naturally occurring *O*-acetylation in microorganisms range between < 10 and 70%, while a spontaneous mutant of *Micrococcus luteus*, cultured in the presence of HEWL, has been reported to have a molar ratio of 1:1 for *N*-acetylmuramic acid and *O*-acetyl, respectively. However, the role of *O*-acetylation in conferring resistance to the hydrolytic activity of HEWL was discerned soon after the initial discovery of the modification. Subsequent studies have confirmed the resistance of most muramidase to *O*-acetylated peptidoglycan; however, it should be noted that the efficacy of others, such as the N, O-diacetylmuramidases of *Chalaropsis* and *Streptomyces globisporus*, is not affected by O-acetyl groups, rather its role in resistance to muramidase exists (Clarke & Dupont, 1992).

Moreover, as with other types of modifications to peptidoglycan, the chloramphenicol-induced increase in the degree of 6-*O*-acetylation of the *N. gonorrhoeae* PG was observed to immediately follow the cessation of protein synthesis; this subsequent acetylation was limited to newly incorporated PG. Therefore, while the process of *O*-acetylation occurs outside the cytoplasm, it would appear that it is regulated by agents that act directly on cytoplasmic targets. Of therapeutic concern, the increased 6-*O*-acetyl content of cell walls in bacteria that are treated with this antibiotic serve to increase its resistance to lysozyme (Clarke & Dupont, 1992).

To overcome the killing action of lysozyme, bacteria have developed different mechanisms, among which some have been well dissected. Further, these are mainly based on the modification of the PG

structure (Le Jeune *et. al.*, 2010). There is still a need for further investigation to explore mechanisms of lysozyme resistance to improve the use of lysozyme in therapeutics.

1.3.2.4 Inhibitory Activity of Lysozyme Against Microorganisms Other than Bacteria

To date, no clear explanation has been provided for the sensitivity of some viruses to lysozyme. Cisani suggested that the inhibition of the herpes virus is due to interference with antiviral activity because of the basic nature of lysozyme, rather than to the hydrolytic activity (Cisani *et. al.*, 1984). Moreover, *in vitro* inhibition of human immunodeficiency virus (HIV) by lysozyme was attributed to the hydrolysis of viral polysaccharides and RNA transcripts, or genomic RNA. Therefore, lysozyme may not always act by its hydrolytic activity; however, it could inhibit the growth of some microorganisms either by permeabilizing the plasma membrane or acting on intracellular components by virtue of its cationic hydrophobic nature, as has been described for a variety of antimicrobial peptides. In fact, sublethal concentrations (10 µg/ml) of lysozyme were shown to accumulate in the cytoplasm of *Candida albicans* and reduce the production and activity of aspartic proteinase, a putative virulence factor of yeast. This finding indicates that lysozyme acts at the transcriptional or translational level of DNA expression and, while at high concentrations, it induces a cell-swelling and invaginations near bud scars. This suggests that interference with the synthesis of cell-wall components may be alternative targets for this enzyme (Benkerroum, 2008). Further research is needed to develop a clearer understanding of the mechanism of action by which lysozyme acts against sensitive microorganisms in general, and eukaryotic and non-cellular microorganisms in particular.

1.4. Lysozyme Detection

The need for a gentle method to lyse a particular mechanism is common to many areas of research in microbiology; however, a few examples of those procedures with a gentle procedure for removing or weakening the cell wall have best initiated (Chassy & Giuffrida, 1980). Numerous methods have been reported for the determination of lysozyme. A traditional method was reported in 1946, in which an Ostwald or Uberlohde's viscometer was employed in the determination of lysozyme by monitoring the

decrease in viscosity of various substrates (Song and Hou, 2003). Until now, the most widely used method was chromatography involving HPLC and affinity chromatography. Furthermore, the nephelometric immunoassay has also been developed and has a detection limit of 0.58 mg/ml for assay of lysozyme. In addition, several other useful methods have been proposed, including fluorimetry and acoustic wave viscosity sensor and response surface methodology (Song and Hou, 2003).

1.4.1 Factors Influencing Lysis Efficiency

A number of factors that influence lysis (lysozyme activity) efficiency have been discussed in other studies. Specifically, increases in the amount of lysozyme used were not particularly effective in increasing extent of lysis; however, it was found that lengthening the period of incubation improved lysis. Another factor that is likely to produce a large change in lysis efficiency is the growth medium. Researchers found that supplementation of the growth medium with L-threonine, L-lysine, or both, usually produced *streptococci* or *lactobacilli* cells that were more easily lysed. The action of threonine can be explained by its known interference in the establishment of cell wall cross-links. The mode of action of L-lysine is more difficult to explain and it is obvious that the composition of the growth medium can affect the susceptibility of bacteria to lysozyme; therefore, it should be evaluated with strains that are difficult to lyse (Chassy & Giuffrida, 1980). In addition, the choice of buffers is likely a major factor in determining the success of many attempted lysis experiments. Other studies have relied on higher pH values, higher buffer concentrations, or phosphate buffers. Thus, the use of NaCl, MgCl₂, other salts, or chelating anions should be avoided because these additives appear to interfere with the lysis of Gram-positive, asporogenous bacteria. Finally, another factor that contributes to the effectiveness is that the enzymatic stabilizer polyethylene glycol (PEG) appears to have more value as a stimulant of lysis than as an osmotic stabilizer (Chassy & Giuffrida, 1980).

1.4.2 Agar Plate Bioassays

The lysoplate method, Fleming's second experiment, has also been used as the basis for assay development. However, Fleming's experiments were more complex than the methods used today. Fleming

began by boring a well into the agar in a petri dish. A mixture of molten agar and nasal secretions containing lysozyme were added to the wells and a second layer of agar was applied, covering the whole plate. The top layer of agar was then inoculated with a layer of *M. luteus* bacteria and incubated for 24 hours. Fleming's experiment showed that lysozyme was able to diffuse through the agar and prevent growth of the bacteria (Fleming & Allison, 1922). Osserman *et. al.* (1966) showed that Fleming's lysoplate technique still worked even when simplified. In his experiments the agar was mixed with heat killed inactivated bacteria so that it appeared turbid; the agar was allowed to solidify and then wells were bored into it. Lysozyme samples were added to these wells and allowed to diffuse over 12-18 hours. The observed zone of turbidity clearing is proportional to the concentration of lysozyme (Osserman & Lawlor, 1966). A big advantage of using this method is its simplicity. Very little training is required to complete the test in a clinical environment.

Similarly, the measurement of lysozyme activity by agar plate bioassay with *Micrococcus lutues* cells, as a test strain, is an indicator test that has been used with many other studies. Enzyme-action produced a clear zone with diameters that were proportional to the catalytic activities of the enzyme (Bower, Xu, & McGuire, 1998), by growth inhibition of *M. luteus* on the plate. Therefore, the lysozyme enzyme is found to be active on *M. lutues* cells, which prevented their growth on assay plates (Akinalp *et. al.*, 2007). Therefore, the use of agar plate bioassays is a quick and efficient method for screening of initial lysozyme activity.

1.4.3 Activity in Media Measurements (Turbidimetric)

Fleming's turbidimetric assay is the first foundation for the turbidimetric assay of lysozyme. His assay is founded on a clearing phenomenon, which is based on preparing a bacterial suspension and followed by the addition of lysozyme sample. The rate of optical density reduction was measured to determine the lysozyme content of the lysozyme sample (Ronan *et. al.*, 1975). Many of the turbidimetric methods used are relatively similar and differ only with respect to the lysozyme activity and sample preparation. These differences might include buffer composition, pH, ionic strength, concentration of *M. luteus* substrate,

temperature, duration of incubation and preparation of enzyme (Houser, 1983). In example of this method, *M. luteus* cells were dissolved in phosphate buffer (e.g. 0.01 M; pH 7.3) to give a stable suspension with an optical density of 0.9 at 450 nm. The activity of the lysozyme sample was determined by the decrease in turbidity of the *Micrococcus* suspension (at 450 nm) using a spectrophotometer. In this study, for each activity assay, 100 µl of T4L (0.5 mg/ml) was added to 3.0 ml of substrate and the decrease in optical density (absorbance) was recorded every 0.1 s for 60 s (Bower *et. al.*, 1998).

Moreover, in a review of turbidimetric studies, Klass *et. al.* (1977) described the turbidimetric technique as having a high sensitivity and a rapid turnover rate; each sample took only one minute to process. The Klass *et. al.* (1977) method has a detection limit of 1 µg/ml whereas the Ronan *et. al.* (1975) method has a detection limit of 1.5 mg/ml (Klass *et. al.*, 1977; Ronan *et. al.* 1975). The Klass *et. al.* method is the preferred method for routine testing of clinical samples because it has time effective; yet, it is not the most reliable as it produces a high inter-batch variability of results, with a precision of 2.5% as was obtained by Gorin *et. al.* (1971). Gorin actually did not look favourably on this precision as they compared their data with that of Bergmeyer (1965), who achieved a higher rate of precision. As well, the study by Gorin *et. al.* (1971) also showed that not all commercially available lysozyme substrates (e.g. lyophilised *M. luteus* from Sigma, Calbiochem etc.) produce the same effects, even though they are labeled *M. luteus*. It is important to note that different companies produce different preparations of the bacteria, giving varying degrees of clearing rates, with a reported difference of up to 30%, as shown in this study. This is caused by the various physical and chemical treatments to which they subject their bacteria, which consequently alters the vulnerability of *M. luteus* to lysozyme and ultimately leads to unreliable results for the diagnosis of diseases. Regardless of its unreliability, this assay is easy to perform, requires small sample volumes, can be used with serum, urine and tears and is sensitive (Gorin *et. al.*, 1971; Bergmeyer, 1965).

Also, in 1971 Terry *et. al.* published work using an automated turbidimetric assay. This assay involved the use of two colorimeters, which correct the urine and serum discoloration problems experienced with

controls (causing turbid solutions for blank samples). The bacterial suspension and clinical samples were continuously stirred and standards were run simultaneously. This automated assay was able to process 20 samples per hour and had an increased (yet undisclosed) sensitivity (Terry *et. al.*, 1971). Furthermore, a method built on the principles of the turbidimetric assay was devised by Caballero *et. al.* (1999).

Cabellero described a micro-particle enhanced nephelometric immunoassay using serum and urine patient samples. This assay involved the use of polystyrene particles covalently bound with anti-lysozyme antibodies. The polystyrene particles formed larger particles due to the binding with free lysozyme in the sample, which resulted in scattered light. The scattered light at the start of the reaction was compared with that at the end and used to calculate the lysozyme concentration in the patient sample. The assay detection limit was 0.58 mg/l. However, even with this assay Caballero recognized there was still room for improvement due to the long incubation periods, need for biological fluid pre-treatment and poor detection limit (Caballero *et. al.*, 1999).

Finally, a commercial assay currently on the market that employs the principle of turbidimetric assays is the EnzChek® Lysozyme by Molecular Probes (Leiden, Netherlands). This test measures lysozyme in solution at levels as low as 20 U/ml (equivalent to <0.5 µg/ml). The assay measures lysozyme activity using *M. luteus* cell walls on substrate. The bacterial cell walls are specially labeled with a fluorophore in such a way that the fluorescence is quenched. Thus, activated lysozyme reduces the quenching while increasing the fluorescence. The fluorescence therefore is proportional to the lysozyme activity. The increase in fluorescence is measured using a spectrofluorometer, mini fluorometer or a fluorescence microplate reader. Each assay takes around 30 minutes with less than an hour preparation time. The assay lays claim to being “simple and sensitive”

(<https://tools.invitrogen.com/content/sfs/manuals/mp22013.pdf>).

1.4.4 Thermostability of Lysozyme on SDS-PAGE

To determine the thermostability of lysozyme enzymes, Akinalp *et. al.* (2007) tested the supernatant of recombinant bacteria by exposing it to various temperatures (from 37°C to 100°C) for 15 min before

centrifugation at 15,000 rpm to remove denatured proteins. Then supernatant was mixed with an equal volume of trichloroacetic acid and the protein was collected by centrifugation. Following, protein analysis was performed using a denaturing polyacrylamide gel (SDS-PAGE, 12% (w/v)). After electrophoresis, protein bands were visualized by a coomassie blue staining procedure (Akinalp *et. al.*, 2007). The thermostability of these enzymes from the recombinant bacteria was also found to differ from each other. In this study, lysozyme expressed by *S. salivarius* subsp. *thermophilus* cells seemed to increase its capacity for thermoresistance and was not denatured at 70°C for 15 min. In contrast, the enzyme expressed by *L. lactis* and *E. coli* cells was easily denatured when exposed to the same temperature treatment (Akinalp *et. al.*, 2007).

1.4.5. Immunoassays

Researchers have raised great concern for the lack of sensitivity and lengthy performance times of Fleming's turbidimetric and lysoplate techniques; thus, immunoassays may offer an important alternative to the determination of lysozyme by addressing some of these issues. Immunoassays rely on the reaction between the target analyte and a specific binding molecule of biological decent (the antibody). They can produce both quantitative and qualitative results and have shown considerable improvements in sensitivity (Ekins & Chu, 1997). For example, Porstmann *et. al.* (1989) developed an enzyme immunoassay for the detection of lysozyme in patients with Crohn's disease and rheumatoid arthritis. Urine samples were taken from patients and were tested using three variations of the same method. The method showing highest sensitivity involved pre-coating a microtitre plate with anti-lysozyme IgG overnight at 4°C, followed by two hours incubation at room temperature with the clinical sample and IgG-HRP (Horseradish Peroxidase) conjugate. Then, the assay was stopped by addition of o-henylenediamine with 15 minutes incubation. The detection limit for this assay was 0.2 µg/l (Porstmann *et. al.* 1989).

Moreover, Francina *et al.* (1986) reported an immunoassay developed to test lysozyme secretion in serum of acute myeloid leukemia patients. A microtitre plate was incubated at 4°C overnight with anti-lysozyme IgG. After a wash, the plate was then incubated for one hour at 37°C with the clinical sample.

Consecutively, biotinylated antilysozyme was added and the plate was again incubated for one hour further. After, the plate was washed and incubated with avidin peroxidase solution for 10 minutes at room temperature. This assay was stopped by the addition of enzyme substrate for five minutes incubation. The total assay time was ~14.5 hours. To reduce assay time the two one-hour incubations were reduced to 20 minutes and the enzyme substrate incubation to five minutes, yet the total assay time was still over 12 hours. The detection limit for the standard assay was 0.1 ng/ml and for the rapid assay 1 ng/ml (Francina *et. al.*, 1986). In another example, Taylor *et. al.* (1992) also present an immunoassay for the detection of lysozyme. This test was developed to measure lysozyme in healthy adult's serum and urine samples. The method involved pre-coating a microtitre plate with rabbit anti-human lysozyme, followed by addition of the clinical sample and incubation for 90 minutes at room temperature. Additionally, a conjugated sheep anti-human lysozyme was added and incubated for a further 90 minutes at room temperature. Finally, enzyme substrate *p*-nitrophenyl sodium phosphate was added and incubated for 30 minutes at room temperature. The total assay time was 15 hours and the detection limit was 1 µg/litre (Taylor *et. al.*, 1992).

As well, some immunoassay kits are available on the market and include the Human Lysozyme EIA kit from Biomedical Technologies Inc. (Stoughton, USA). Referred to as sandwich ELISA, this method detects lysozyme from serum, plasma, urine, tears and saliva, where the reference value for human lysozyme from serum ranges from 3-10 µg/ml. In this assay, specific lysozyme antibodies bound to polystyrene wells were incubated with a sample and a second human lysozyme-specific antibody is then added; as well as a horseradish peroxidase conjugated secondary antibody. The total test time is 4 ¼ hours (<http://www.funakoshi.co.jp/data/datasheet/BTI/BT-630.pdf>).

Similarly, the company Orgentec (Mainz, Germany) supplies a kit known as Anti-lysozyme kit for lysozyme in serum and plasma. This test only requires 10 µl of patients' sample and involves a plate pre-coated with antibody. The antigen from patients' sample is added along with a horseradish peroxidase conjugate; followed by TMB (a color substrate). The reaction is stopped using Hydrochloric acid and the total time for this assay is two hours, due to the plates being purchased pre-coated. During this time 96

patient samples can be processed

(http://www.orgentec.com/products/pdfs/ELISA_en_IFU_ORG_526.pdf). All these kits have the advantage of conducting analysis on several patient samples using minimal biological fluid. This not only alleviates discomfort and distress to patients, but also the costs to lab facilities.

1.4.6. Biosensor

Biosensor and related bioarray techniques represent the end product of a rapidly growing field, which combines fundamental biological, chemical, and physical sciences with engineering and computer science to satisfy needs in a broad range of application areas. Not surprisingly, the term biosensor has different connotations depending on the field. For example, to the biologist, a biosensor is a device that translates biological variables (e.g., electric potentials, movement, or chemical concentrations) into electrical signals. To the chemist, a definition might be a device that uses specific biochemical reactions as mediated by isolated enzymes, immunosystems, tissues, organelles, or whole cells to detect chemical compounds, usually by electrical, thermal, or optical signals. While, the physicist might define a biosensor as a device that detects, records, and transmits information about a physiological change or process (Marks *et. al.*, 2007). Together, biosensor can be defined as an analytical device that converts the concentration of the target substances, the analyte, into an electrical signal via a combination of biological or biologically-derived recognition systems, either integrated within or intimately associated with a suitable physico-chemical transducer. Basically, biosensors consist of three parts including recognition, transducer, and signal output. A biological sensing element with a transducer will produce a signal that is proportional to target analyses. When biological molecules interact with target analyses, there is change in one or more physical-chemical parameters that are associated with the interaction such as generation of ions, gases, electrons, heat, or mass. The quantities of these indicating signals are converted into electrical signals (Figure 2).

In addition, biological or biologically-derived elements are capable of recognizing the presence, activity, or concentration of a specific target analyte in a complex mixture of components. The recognition

element may comprise one of three different types: affinity biosensors (based on ligand-receptor interactions such as those involving antibodies, nucleic acid, aptamers, peptides, protein, or cell receptor); binding and catalysis (involves enzymes, microorganisms, organelles, plant or animal cells or tissue slices); biomimetic receptors (based on various synthetic binding or catalytic systems). The interaction, or subsequent reaction, of the recognition element with the analyte in a sample matrix results in a reactant or formation of a product that is immediately proximal to the transducer. The latter converts the change in solution property into a quantifiable and processable electrical signal. The transducer is a device, usually electronic, electroacoustic, electro-optical, electromagnetic, electrothermal, or electromechanical, that converts one type of energy (electricity, sound, light, magnetism, heat, or mechanical) into another (usually electrical) for various purposes including measurement or information transfer (Marks *et. al.*, 2007). More research is needed to develop an electrochemical approach for the detection of lysozyme; such biosensor development will offer an alternative, sensitive, and versatile method for protein detection. Currently, most proteins are detected mostly by antibodies in analytical formats like ELISA, immunobead assay, western blotting, and microarrays.

1.4.7 Flow injection Chemiluminescence System

A novel analytical procedure, based on chemiluminescence (CL) detection, was described for the determination of lysozyme at ng/ml level using controlled-reagent-release technology in a flow injection system. The analytical reagents involved in the CL reaction, including luminol and periodate, were immobilized on the anion-exchange resins in the flow injection system. Through water injection, luminol and periodate were eluted from the anion-exchange column to generate the chemiluminescence, which was inhibited in the presence of lysozyme. By measuring the decrease of CL intensity, one could analyze the lysozyme concentration in the range of 30-100 ng/ml (Figure 3). A typical analytical procedure, which includes sampling and washing, could be performed in 0.5 min at a flow rate of 2.0 ml/min, giving a throughput of 120 h with a relative standard deviation of less than 3.0%. The proposed method was

applied successfully to the determination of lysozyme in human tear and saliva samples and recovery was from 92.0% to 105.7% (Song and Hou, 2003).

F1 chemiluminescence (CL) analysis is becoming increasingly important in various fields because of its high sensitivity, rapidity, simplicity and feasibility. The fast oxidation reaction between luminol and periodate in alkaline medium produces a strong CL signal. Further, CL is greatly inhibited by lysozyme; thus, a sensitive CL inhibition assay for lysozyme combined with F1 technology was described in that paper (Figure 4). The proposed method offered simple and cheap instrumentation and rapid and reproducible means of detection; that was demonstrated in its application to determine lysozyme in human tear and saliva (Song and Hou, 2003). Within that work, a sensitive and convenient analytical method, using control-reagent-release technology in flow injection system, was developed to determine lysozyme in human tears and saliva by its inhibitory effect on CL. In addition, incorporation of the immobilized reagents column into the flow injection analysis (FIA) manifold offers satisfactory stability, good reproducibility, and precision in the analysis (Song and Hou, 2003).

Thesis objectives

T4L, the most active lysozyme in sterilization, constitutes a natural defense mechanism. In addition to its anti-bacterial action, lysozyme may have anti-tumor, anti-metastatic, and anti-inflammatory activities in the medical field. Also, it has some industrial application usages including: food-processing, food packing, wine maturation, infant formulas and baby foods, and enhancing plant resistance to diseases. To date, there has been no T4L industrialization globally. This success of such technology development in this thesis study will facilitate the production of recombinant T4L for industrialization purposes.

The principal goal of this study was to express a synthetic *t4l* in two microorganism hosts, which are *E. coli* and *T. reesei*. That was through sequential steps in this study, first studying synthetic *t4l* in bioinformatics level by comparing it to wild-type T4L. Second, transforming three plasmids-bearing-synthetic *t4l* into *E. coli*, then producing a novel strain of *T. reesei*, followed by optimizing culture conditions of transformants in order to enhance the production of recombinant T4L, and evaluation of

target gene expression of transformants. To accomplish that primary goal, there were several specific objectives generated and classified accordingly into each chapter.

Figure and Table legend

Figures:

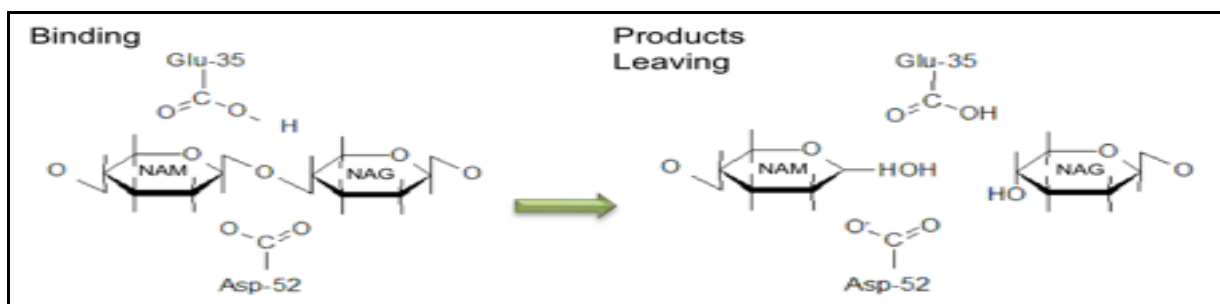


Figure 1 The catalytic mechanism of lysozyme. The substitution on the sugar rings have been omitted to improve the clarity of the diagram. The diagram drew by using ACD/ChemSketch Freeware software, which is a free comprehensive chemical drawing package (<http://www.acdlabs.com>). It was originally adapted from (Smith and Wood, 1991).

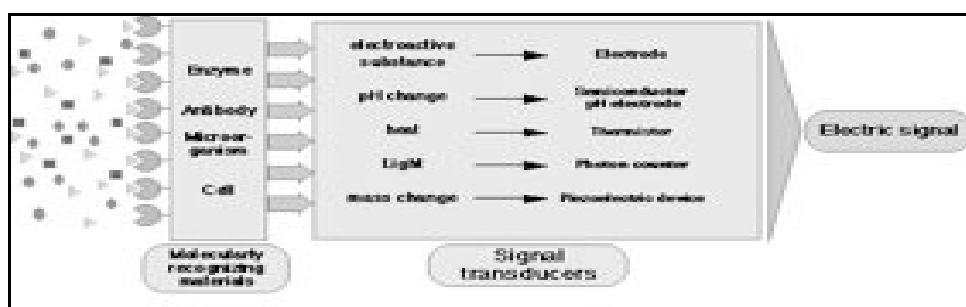


Figure 2 Principle of biosensor (<http://www.jaist.ac.jp/~yokoyama/images/biosensor.gif>)

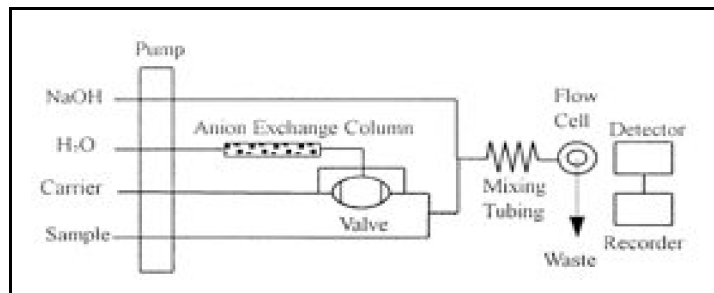


Figure 3 Schematic diagram of the flow-injection system for lysozyme determination (Song and Hou, 2003)

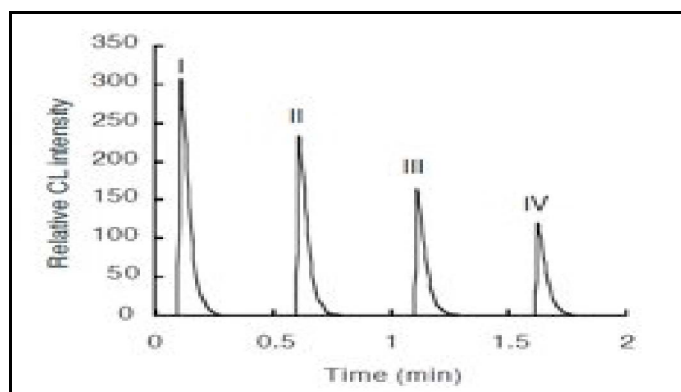


Figure 4 CL time profile in the batch system. I, CL intensity in the absence of lysozyme; II, CL intensity in the presence of lysozyme (50 ng ml^{-1}); III, CL intensity in the presence of lysozyme (150 ng ml^{-1}); IV, CL intensity in the presence of lysozyme (500 ng ml^{-1}) (Song and Hou, 2003)

Tables:**Table 1** Main chemical properties of milk lysozymes of different mammals: chicken egg-white lysozyme properties are also given as a reference for the c-type lysozyme, (Benkerroum, 2008).

Origin	MW (kDa)	Folding properties	Catalytic center residues	Reference
Camel milk	14.4	NA	NA	El Agamy et al. (1996)
Cow milk	14.4	Two-state	Glu ₃₅ Asp ₅₃	El Agamy et al. (2000)
Ewe milk	16.2	Two-state	Glu ₅₃ Asp ₂₁	Maroni et Cuccuri, (2001)
Human	15	Two-state	---	Parry et al. (1960)
Goat	14.4	Two-state	Glu ₃₅ Asp ₅₃	Jolles et al. (1990)
Buffalo	16.0	Two-state	NA	El Agamy (2000)
Mare	14.7	Three-state	Glu ₃₅ Asp ₅₃	Jauregui-adell (1974); Sarwar et al. (2001)
Canine	14.5	Three-state	Glu ₃₅ Asp ₅₃	Grobler et al. (1994)
Egg-white	14.3	Two-state	Glu ₃₅ Asp ₅₂	Matagne and Dobson, (1998)

Table 2 Reported concentrations (mg/l) of lysozyme in the milk of different mammals, (Benkerroum, 2008)

Animal Species	Average concentration	References
Human	400	Mathur et al. (1990)
	320	Montagne et al. (1998)
	270-890	Montagne et al. (2001) and Chandan et al. (1968)
	224-426	Hennart et al. (2004)
Ass	1428	Salimei et al. (2004)
Mare	790	Jauregui-Adell (1975)
	1330	Sarwar et al. (2001)
Cow	0.13	Chandan et al. (1968)
	0.07	El Agamy et al. (1996)
	0.05 – 0.21	Piccinni et al. (2005)
Buffalo	0.0012	Priyadarshini and Kansal (2003)
Ewe	0.1	Chandan et al. (1968)
Goat	0.25	Chandan et al. (1968)
Sow	6.8	Schultz and Müller (1980)
Camel	0.15	El Agmay et al. (1996)

Table 3 Application area, antimicrobial agents and packaging materials used in antimicrobial food packaging (Source: Han 2000, Appendini and Hotchkiss 2002).

Antimicrobial Agent	Packaging Material	Application area
Lysozyme	LDPE*	Bell pepper
	LDPE	Cheese

* LDPE: low-density polyethylene;

References

- Akinalp, A.S., M. Asan and N. Ozcan. (2007). Expression of T4 lysozyme gene (gene e) in *Streptococcus salivarius* subsp. *thermophilus*. *Afr. J. Biotechnol.*, 6: 963-966.
- Alberts, B., Bray, D., Hopkin, K., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2010), *Essential Cell Biology* (3rd ed.). London: Taylor & Francis Group. ISBN: 978-0-8153-4129-1
- Appendini, P. and Hotchkiss, J. H. (1997). Immobilization of lysozyme on food contact polymers as potential antimicrobial films. *Packag. Technol Sci.* 10, 271-279.
- Appendini, P. and Hotchkiss, J. H. (2002). Review of antimicrobial food packaging. *Innov. Food Set Emerg. Technol.* 3,113-126.
- Baase, W. A., Liu, L., Tronrud, D. E., & Matthews, B. W. (2010). Lessons from the lysozyme of phage T4. *Protein Sci*, 19(4), 631-641. doi: 10.1002/pro.344
- Benkerroum, N. (2008). Antimicrobial activity of lysozyme with special relevance to milk. *Afr. J. Biotechnol.* 7: 4856–4867.
- Bera, A., Herbert, S., Jakob, A., Vollmer, W., & Gotz, F. (2005). Why are pathogenic staphylococci so lysozyme resistant? The peptidoglycan O-acetyltransferase OatA is the major determinant for lysozyme resistance of *Staphylococcus aureus*. *Mol Microbiol*, 55(3), 778-787. doi: 10.1111/j.1365-2958.2004.04446.x
- Bio-products: BRI of CAAS awaiting investments in T4 lysozyme industrialization. (2011). *Industrial Biotechnologies China News*, 3(01), 4
- Bower, C. K., Sananikone, S., Bothwell, M. K., & McGuire, J. (1999). Activity losses among T4 lysozyme charge variants after adsorption to colloidal silica. *Biotechnol Bioeng*, 64(3), 373-376.
- Bower, C. K., Xu, Q., & McGuire, J. (1998). Activity losses among T4 lysozyme variants after adsorption to colloidal silica. *Biotechnol Bioeng*, 58(6), 658-662.
- Buonocore, G. G., Sinigaglia, M., Corbo, M. R., Bevilacqua, A., La Notte, E., & Del Nobile, M. A. (2004). Controlled release of antimicrobial compounds from highly swellable polymers. *J Food Prot*, 67(6), 1190-1194.
- Burgess P (1973) *Lysozyme: History, Methods of Assay and Some Applications*, Industrial Year Report, Hatfield Polytechnic
- Burgess, P., Appel, S. H., Wilson, C. A., & Polk, H. C., Jr. (1994). Detection of intraabdominal abscess by serum lysozyme estimation. *Surgery*, 115(1), 16-21.
- Caballero, M., Ruiz, R., Marquez de Prado, M., Seco, M., Borque, L., & Escanero, J. F. (1999). Development of a microparticle-enhanced nephelometric immunoassay for quantitation of human lysozyme in pleural effusion and plasma. *J Clin Lab Anal*, 13(6), 301-307.
- Chang, H. M., Yang, C. C., & Chang, Y. C. (2000). Rapid separation of lysozyme from chicken egg white by reductants and thermal treatment. *J Agric Food Chem*, 48(2), 161-164.
- Chassy, B. M., & Giuffrida, A. (1980). Method for the lysis of Gram-positive, asporogenous bacteria with lysozyme. *Appl Environ Microbiol*, 39(1), 153-158.
- Cisani, G., Varaldo, P.E., Ingianni, A., Pompei, R., & Satta, G., (1984). Inhibition of herpes simplex virus-induced cytopathic effect by modified hen egg-white lysozymes. *Current Microbiology*. 10: 35–40. DOI: 10.1007/BF01576045
- Clarke, A. J., & Dupont, C. (1992). O-acetylated peptidoglycan: its occurrence, pathobiological significance, and biosynthesis. *Can J Microbiol*, 38(2), 85-91.

- Cross, M., Mangelsdorf, I., Wedel, A., & Renkawitz, R. (1988). Mouse lysozyme M gene: isolation, characterization, and expression studies. *Proc Natl Acad Sci U S A*, 85(17), 6232-6236.
- Daly, S. (1938). Lysozyme of Nasal Mucus: Method of preparation and preliminary report on its effects on growth and virulence of common pathogens of paranasal sinuses. *Arch Otolaryngol*. 27(2): 189-196.
- Davis, C. S. (1971). Diagnostic value of muramidase. [Review]. *Postgrad Med*, 49(4), 51-54.
- De Roever, C. (1998). Microbiological Safety Evaluations and Recommendations on Fresh Produce. *Food Control*. 9, 321-347.
- Devlieghere, F., Vermeiren, L., Debevere, J., 2004. New preservation technologies: possibilities and limitations. *Int. Dairy J.* 14, 273–285.
- Dong, S., Shew, H. D., Tredway, L. P., Lu, J., Sivamani, E., Miller, E. S., & Qu, R. (2008). Expression of the bacteriophage T4 lysozyme gene in tall fescue confers resistance to gray leaf spot and brown patch diseases. *Transgenic Res*, 17(1), 47-57. doi: 10.1007/s11248-007-9073-3
- Ekins, R., & Chu, F. (1997). Immunoassay and other ligand assays: present status and future trends. [Review]. *J Int Fed Clin Chem*, 9(3), 100-109.
- Fett, J. W., Strydom, D. J., Lobb, R. R., Alderman, E. M., Vallee, B. L., Artymiuk, P. J., & Redfield, C. (1985). Lysozyme: a major secretory product of a human colon carcinoma cell line. *Biochemistry*, 24(4), 965-975.
- Fleming, A. (1922). On a remarkable bacteriolytic element found in tissues and secretions. *Proc. Roy. Soc. London*. 93: 306-317.
- Fleming, A., & Allison, V. (1927). On the Development of Strains of Bacteria Resistant to Lysozyme Action & the Relation of Lysozyme Action to Intracellular Digestion. *J. Brit. Exp. Path.* 8, 214-218.
- Fleming, A., & Allison, V. D. (1992). Further observations on a bacteriolytic element found in tissues and secretions. *Proc. Roy. Soc. London. Series B*, 44:142-151.
- Francina, A., Cloppet, H., Guinet, R., Rossi, M., Guyotat, D., Gentilhomme, O., & Richard, M. (1986). A rapid and sensitive non-competitive avidin-biotin immuno-enzymatic assay for lysozyme. [Comparative Study]. *J Immunol Methods*, 87(2), 267-272.
- Gill, A., Scanlon, T. C., Osipovitch, D. C., Madden, D. R., & Griswold, K. E. (2011). Crystal structure of a charge engineered human lysozyme having enhanced bactericidal activity. *PLoS One*, 6(3), e16788. doi: 10.1371/journal.pone.0016788
- Gorin, G., Wang, S. F., & Papapavlou, L. (1971). Assay of lysozyme by its lytic action on *M. lysodeikticus* cells. *Anal Biochem*, 39(1), 113-127.
- Grutter, M. G., & Matthews, B. W. (1982). Amino acid substitutions far from the active site of bacteriophage T4 lysozyme reduce catalytic activity and suggest that the C-terminal lobe of the enzyme participates in substrate binding. *J Mol Biol*, 154(3), 525-535.
- Han, J. H. (2000). Antimicrobial food packaging. *Food Technol*. 3: 56–65.
- Hansen, N. E., Karle, H., Andersen, V., & Olgaard, K. (1972). Lysozyme turnover in man. *J Clin Invest*, 51(5), 1146-1155. doi: 10.1172/JCI106907
- Helal, R., & Melzig, M.F. (2010) In vitro Effects of Selected Saponins on the Production and Release of Lysozyme Activity of Human Monocytic and Epithelial Cell Lines. *Sci Pharm*. 79(2): 337-349

- Herbert, S., Bera, A., Nerz, C., Kraus, D., Peschel, A., Goerke, C., . . . Gotz, F. (2007). Molecular basis of resistance to muramidase and cationic antimicrobial peptide activity of lysozyme in staphylococci. *PLoS Pathog*, 3(7), e102. doi: 10.1371/journal.ppat.0030102
- Hinnrasky, J., Chevillard, M., & Puchelle, E. (1990). Immunocytochemical demonstration of quantitative differences in the distribution of lysozyme in human airway secretory granule phenotypes. [Comparative Study]. *Biol Cell*, 68(3), 239-243.
- Houser, M. T. (1983). Improved turbidimetric assay for lysozyme in urine. [Comparative Study]. *Clin Chem*, 29(8), 1488-1493.
- Houser, M.T. (1983). Improved turbidimetric assay for lysozyme in urine. *Clin. Chem.* 29: 1488-1493
- Ibrahim, H.R., Higashiguchi, S., Jugena, L.R., Kim, M., & Yamamoto, T. (1996). A Structural Phase of Heat-Denaturated Lysozyme with Novel Antimicrobial Action. *J. Agric. Food Chem.* 44: 1416-1423.
- Ibrahim, H.R., Kato, A., Kobayashi, K. (1991). Antimicrobial Effects of Lysozyme Against Gram-Negative Bacteria due to Covalent Binding of Palmitic Acid. *J. Agric. Food Chem.* 39, 2077-2083.
- Jollès P. (1969). Lysozymes: a chapter of molecular biology. *Angew. Chem. Int. Ed. Engl.* 8(4): 227–239.
- Klass, H. J., Hopkins, J., Neale, G., & Peters, T. J. (1977). The estimation of serum lysozyme: a comparison of four assay methods. [Comparative Study]. *Biochem Med*, 18(1), 52-57.
- Kuroki, R., Weaver, L. H., & Matthews, B. W. (1999). Structural basis of the conversion of T4 lysozyme into a transglycosidase by reengineering the active site. *Proc Natl Acad Sci U S A*, 96(16), 8949-8954.
- Labuza, T. P., & Breene, W. (1989). Application of 'active packaging' technologies for the improvement of shelf-life and nutritional quality of fresh and extended shelf-life foods. [Review]. *Bibl Nutr Dieta*(43), 252-259.
- Le Jeune, A., Torelli, R., Sanguinetti, M., Giard, J. C., Hartke, A., Auffray, Y., & Benachour, A. (2010). The extracytoplasmic function sigma factor SigV plays a key role in the original model of lysozyme resistance and virulence of *Enterococcus faecalis*. *PLoS One*, 5(3), e9658. doi: 10.1371/journal.pone.0009658
- Lie, O., Evensen, O., Sorensen, A., & Froyssedal, E. (1989). Study on lysozyme activity in some fish species. *Dis. Aquat. Org.* 6: 1-5.
- Lollike, K., Kjeldsen, L., Sengelov, H., & Borregaard, N. (1995). Purification of lysozyme from human neutrophils, and development of an ELISA for quantification in cells and plasma. *Leukemia*, 9(1), 206-209.
- Malmsten, M. (Ed.) (2003). *Biopolymers at Interfaces* (2nd Ed.). New York: Marcel Dekker.
- Marks, R.S., Cullen, D., Lowe, C., Weetall, H.H., & Karube, I. (2007). *Handbook of Biosensors and Biochips*. Mississauga: John Wiley & Sons Ltd.
- Marx, K. A. (2003). Quartz crystal microbalance: a useful tool for studying thin polymer films and complex biomolecular systems at the solution-surface interface. *Biomacromolecules*, 4(5), 1099-1120. doi: 10.1021/bm020116i
- Matsumura, M., Bechtel, W. J., Levitt, M., & Matthews, B. W. (1989). Stabilization of phage T4 lysozyme by engineered disulfide bonds. *Proc Natl Acad Sci U S A*, 86(17), 6562-6566.
- Matthews, B. W., & Remington, S. J. (1974). The three dimensional structure of the lysozyme from bacteriophage T4. [Comparative Study]. *Proc Natl Acad Sci U S A*, 71(10), 4178-4182.

- Mecitoglu, Ç., Yemeniciolu, A., Arslanolu, A., Elmacı, Z.S., Korel, F., & Çetin, A.E. (2006). Incorporation of Partially Purified Hen Egg White Lysozyme into Zein Films for Antimicrobial Food Packaging. *Food Research International*. 39: 12-21.
- Min, S. & Krochta, J.M. (2005). Inhibition of *Penicillium Commune* by Edible Whey Protein Films Incorporating Lactoferrin, Lactoferrin Hydrosylate, and Lactoperoxidase Systems. *J. Food Sci.* 70: 87-94.
- Mir, M. A. (1977). Lysozyme: a brief review. *Postgrad Med J*, 53(619), 257-259.
- Nakamura, S., Kato, A., & Kobayashi, K. (1991) New antimicrobial characteristics of lysozyme-dextran conjugates. *J. Agric. Food Chem.* 39: 647-650.
- Nakamura, S., Kato, A., & Kobayashi, K. (1992). Bifunctional lysozyme-galactomannan conjugate having excellent emulsifying properties and bactericidal effect. *J. Agric. Food Chem.* 40:735-739.
- O'Sullivan, C.K., & Guilbault, G.G. (1999). Commercial quartz crystal microbalances-theory and applications. *Biosens. Bioelectron.* 14: 663-670.
- Osserman, E. F., & Lawlor, D. P. (1966). Serum and urinary lysozyme (muramidase) in monocytic and monomyelocytic leukemia. *J Exp Med*, 124(5), 921-952.
- Patil, R. S., Ghormade, V. V., & Deshpande, M. V. (2000). Chitinolytic enzymes: an exploration. *Enzyme Microb Technol*, 26(7), 473-483.
- Porstmann, B., Jung, K., Schmechta, H., Evers, U., Pergande, M., Porstmann, T., & Krause, H. (1989). Measurement of lysozyme in human body fluids: comparison of various enzyme immunoassay techniques and their diagnostic application. *Clin Biochem*, 22(5), 349-355.
- Poteete, A. R., Rennell, D., Bouvier, S. E., & Hardy, L. W. (1997). Alteration of T4 lysozyme structure by second-site reversion of deleterious mutations. *Protein Sci*, 6(11), 2418-2425. doi: 10.1002/pro.5560061115
- Quintavalla, S., & Vicini, L. (2002). Antimicrobial food packaging in meat industry. *Meat Sci*, 62(3), 373-380.
- Ramanathan, A., Savol, A. J., Langmead, C. J., Agarwal, P. K., & Chennubhotla, C. S. (2011). Discovering conformational sub-states relevant to protein function. *PLoS One*, 6(1), e15827. doi: 10.1371/journal.pone.0015827
- Reitamo, S., Lalla, M., & Huipero, A. (1981). Serum lysozyme: evaluation of a nephelometric assay. [Comparative Study]. *Scand J Clin Lab Invest*, 41(4), 329-332.
- Ronen, D., Eylan, E., Romano, A., Stein, R., & Modan, M. (1975). A spectrophotometric method for quantitative determination of lysozyme in human tears: description and evaluation of the method and screening of 60 healthy subjects. *Invest Ophthalmol*, 14(6), 479-484.
- Song, Z. H., & Hou, S. (2003). A new analytical procedure for assay of lysozyme in human tear and saliva with immobilized reagents in flow injection chemiluminescence system. *Anal Sci*, 19(3), 347-352.
- Sulakvelidze, A., Alavidze, Z., & Morris, J. G., Jr. (2001). Bacteriophage therapy. *Antimicrob Agents Chemother*, 45(3), 649-659. doi: 10.1128/AAC.45.3.649-659.2001
- Suppakul, P., Miltz, J., Sonneveld, K., & Bigger, SW. (2003). Active packaging technologies with an emphasis on antimicrobial packaging and its applications. *J. Food Sci.* 68(2): 408-420.
- Takahashi, K., Lou, X. F., Ishii, Y., & Hattori, M. (2000). Lysozyme-glucose stearic acid monoester conjugate formed through the Maillard reaction as an antibacterial emulsifier. *J Agric Food Chem*, 48(6), 2044-2049.

- Taylor, D. C., Cripps, A. W., & Clancy, R. L. (1992). Measurement of lysozyme by an enzyme-linked immunosorbent assay. [Comparative Study]. *J Immunol Methods*, 146(1), 55-61.
- Terry, J.M., Blainey, J.D., & Swingler, M.C. (1971). An automated method for lysozyme assay. *Clin Chim Acta*. 35(2): 317-320.
- Trun, N., & Trempey, J (2004). *Fundamental Bacterial Genetics*. Malden: Blackwell Publishing.
- Tsugita, A., & Inouye, M. (1968). Purification of bacteriophage T4 lysozyme. [Comparative Study]. *J Biol Chem*, 243(2), 391-397.
- Van Dael, H. (1998). Chimeras of human lysozyme and alpha-lactalbumin: an interesting tool for studying partially folded states during protein folding. *Cell Mol Life Sci*, 54(11), 1217-1230.
- Wohlkonig, A., Huet, J., Looze, Y., & Wintjens, R. (2010). Structural relationships in the lysozyme superfamily: significant evidence for glycoside hydrolase signature motifs. *PLoS One*, 5(11), e15388. doi: 10.1371/journal.pone.0015388
- Zhang, X. J., Baase, W. A., Shoichet, B. K., Wilson, K. P., & Matthews, B. W. (1995). Enhancement of protein stability by the combination of point mutations in T4 lysozyme is additive. *Protein Eng*, 8(10), 1017-1022.
- Zhang, X.J., Baase, W.A., Shoichet, B.K., Wilson, K.P., & Matthews, B.W. (1995). Enhancement of protein stability by the combination of point mutations in T4 lysozyme is additive. *Protein Eng*, 8: 1017-1022.
- Smith C. A. and Wood E. J. (1991). *Molecular and Cell Biochemistry: Biological Molecules*, Chapman & Hall, London. ISBN 0 412 40780 9. 114-116

Chapter II: Computational Analysis of T4 Lysozyme

Abstract

The field of bioinformatics has experienced explosive growth in the last decade, in both the field of bioinformatics and computational biology; two T4L genes that have been studied are the wild-type and synthetic one. The later, new synthetic *t4l* was produced by codon-based modification with alternation of triplet nucleotides that encodes two specific amino acid residues at position 6 and 12 in a polypeptide chain of T4L protein. The goal of constructing synthetic *t4l* was to enhance the production of recombinant T4L for industrialization purposes after its successful expression in the host. Therefore, comparison of the wild-type and synthetic *t4l* from nucleotide sequences up to 3D protein structure was unavoidable. Bioinformatics tools are used to identify the nucleotide sequences and their protein products, identify homologous sequences, structural information, and analyze the quality of the structures. Also, different programs, web servers, and parameters are used as bioinformatics tools. Likewise, experimentally derived 3D proteins' structures have been studied in detail.

Wild-type and synthetic *t4l* have variants in the size of nucleotides, 459 bp and 491 bp, respectively, but both have equal amino acids size, 164 a.a. The results of the multiple sequence alignment indicate both amino acids sequences have high fully conserved residues to their sequence similarity obtained from the database. In addition, the conserved motifs were high with wild-type *t4l* as well the synthetic one; just one out of seven had a slight difference for position-specific score matrix with T4LYSOZYME1 motif in our study. The comparison of 3D protein structure showed us how both protein structures have structural similarity by calculating RMSD; 0.17 for the backbone coordinate and 0.14 for the carbon alpha coordinate. Also, structure sequence identity was 98.87% by using the alignment of editor Jalview. In conclusion, wild-type and synthetic T4L supposedly share the same function of hydrolysis, yet there is variability at position 6 and 12 in amino acids sequence.

Key words: *Bioinformatics, Wild type and Synthetic T4 Lysozyme,*

2.1. Introduction

Phage T4 is considered one of the most structurally complex viruses. It is a linear double-stranded DNA (dsDNA) tailed virus with a genome that contains 274 open reading frames (ORFs) out of which more than 40 encode structural proteins. The structures of phage T4 head, tail, and fibres have been well studied. Subsequently, a structural model of phage T4 has been determined. The advances in electron microscopy, X-ray crystallography, and computing power have extended the structural knowledge of T4 to higher resolution (Leiman *et al.*, 2003). Phage T4 has been used as an important model system in the development of modern genetics and molecular biology since the 1940s. The advantages of phage T4 as a model system stemmed in part from the virus' total inhibition of host gene expression, which allows investigators to differentiate between host and phage macromolecular syntheses. Phage T4 also produces some important enzymes with widespread commercial applications, including its DNA and RNA ligase, polynucleotide kinase, and DNA polymerase. Many would argue that to know phage T4 is to know the foundations of molecular biology and the essential paradigms of genetics and gene expression (Miller *et al.*, 2003).

Structurally, phage T4 has a chromosome surrounded by a capsid that is attached to a tail structure; all are composed of phage-encoded proteins. The phage genome encodes most of the gene products needed for replicating more phages. It has a 172 kb linear double-stranded DNA genome (Trun & Trempey, 2004). Additionally, by the early 1990s, much of the genome had been sequenced, but extensive regions remained intractable. However, the complete 168,903-bp sequence of the T4 genome is now available on the NCBI Genome website (<http://www.ncbi.nlm.nih.gov/Entrez>), under GenBank accession no. AF158101 and/or entry NC_000866 (Miller *et al.*, 2003).

In molecular biology, glycoside hydrolase family 24 is a family of glycoside hydrolases EC 3.2.1. a widespread group of enzymes that hydrolyse the glycosidic bond between two or more carbohydrates, or between a carbohydrate and a non-carbohydrate moiety. A classification system for glycoside hydrolases, based on sequence similarity, has led to the definition of a number of different families (Henrissat, 1991;

Henrissat & Bairoch, 1993; Henrissat & Bairoch, 1996). This classification is available on the CAZy website (<http://www.cazy.org/Glycoside-Hydrolases.html>) (Henrissat and Coutinho, 2011), and also discussed at the CAZypedia website, which is an online encyclopaedia of carbohydrate active enzymes (http://www.cazypedia.org/index.php/Main_Page). Moreover, the glycoside hydrolase family 24 (GH24) comprises enzymes with only one known activity that is hydrolysis. This family includes *E. coli* endolysin, lambda phage lysozyme, as well as T4L (Weaver & Matthews, 1987). *E. coli* endolysin functions in bacterial cell lysis and acts as a transglycosylase, while lysozyme helps to release mature phage particles from the cell wall by breaking down the peptidoglycan of prokaryotic cell walls. Specifically, T4L structure contains two domains, the interface between which forms the active-site cleft. The N-terminus of the two domains undergoes a “hinge-bending” motion about an axis passing through the molecular waist (Weaver & Matthews, 1987; Faber & Matthews, 1990). This mobility is thought to be important in allowing access of substrates to the enzyme active site.

Currently, molecular biological data are being produced at an unexpected rate. A flood of data means that many of the challenges in biology are now challenges in computing. Thus, the use of techniques from applied mathematics, informatics, statistics and computer sciences to solve biological problems has created the bioinformatics field (Luscombe *et. al.*, 2001). In fact, several new fields within molecular genetics have been created, such as bioinformatics, genomics, and proteomics. Each of these fields integrates the use of computer hardware and software to store, process and analyze biological information, such as DNA or protein sequences (Elrod & Stansfield, 2010). Therefore, bioinformatics could be defined as the application of computational techniques to understand and organize the information associated with biological molecules. The three main aims of bioinformatics involve: firstly, the simplest organization of data in such a way that allows scientific researchers to access existing information and to submit new entries as they are produced. Secondly, bioinformatics requires development of tools and resources to aid in the analysis of those data. Thirdly, these tools analyze the data and interpret the results in a biologically

meaningful manner (Luscombe *et. al.*, 2001). Therefore, correlation of these aims extends the purpose of bioinformatics much further.

More specifically, the field of bioinformatics relies heavily on Linux-based computers and program software, which is a free operating system (OS) for computers. While most bioinformatics programs can be run on Mac OS X and Windows systems, it is often more convenient to install and use the software on a Linux system. And for most Linux users, the simplest way to access a Linux system is by connecting from their primary Mac OS X or Windows machine. This type of arrangement allows several users to run software on a single Linux system, which can be maintained by an experienced systems administrator. In addition, Linux users should be familiar with some basic commands (Stothard, 2011), because it is the only language for communication with this system.

Furthermore, Virtual Network Computing (VNC) is remote control computer software, which allows the user to view and fully interact with one VNC server using a simple VNC viewer on another computer desktop anywhere on the Internet. Additionally, it is not a requirement that the two computers be the same type. While for ultimate simplicity, there is even a Java viewer, so that any desktop can be controlled remotely from within a browser without having to install software. Currently, VNC is in widespread active use by many millions of people throughout industry, academia and privately. There are also several versions from which to choose, including a free version and some substantially enhanced commercial versions (Sasa, 2012).

Here, studying wild-type and synthetic *t4l* genes is crucial to genetic and protein modifications, by means of bioinformatics. While the genetic modification on synthetic *t4l* was highly sensitive and specific, comparison to its wild-type bioinformatically might be essential for a desired expression. The comparison between these two *t4l* genes has been studied on a bioinformatics level, starting from their nucleotide sequences up to three-dimensional (3D) protein structures. Generally in this section of the thesis project, nucleotide and amino acid, conserved and non-conserved motifs, and 3D structures of T4L were analyzed by different popular bioinformatics software.

2.1.1. Approach

Two nucleotide sequences identify the genetic sequence of a particular protein, T4L, they have been assigned by sequential steps. Firstly, a gene sequence corresponding to wild-type and synthetic *t4l* genes was retrieved, using the NCBI genome browser for the former, while the latter was already known. Also, search for synthetic *t4l* gene sequence was important to determine whether there is any previous published research work using the synthetic *t4l* in this study yet. Subsequently, the nucleotide sequence is translated into amino acid sequence, and then the obtained protein is reported. Next, a database similarity search is run using BLAST for amino acid sequences. Then, a multiple sequence alignment is performed using the amino acid sequence file built as mentioned above. There are many reasons to construct a multiple sequence alignment, including highlighting regions of similarity, divergence and mutations. This is done by describing regions/positions as either conserved or variable. Later, a specific database server was used to search for conserved motifs in the protein sequence. Then, those motifs sequences determined where they hit in the synthetic T4L amino acid sequence. The next step is to find out if wild-type and synthetic *t4l* proteins have three-dimensional (3D) protein structures available in the Protein Data Bank (PDB). If any of those proteins do not have experimentally determined 3D structure, one can be created through the use of a modeling web. The protein 3D structure is required to build a graphical representation and get an image file; that is to analyze protein structures and discuss the overall quality of both structures.

2.2. Methodology

2.2.1 VNC Software

This applied project familiarizes one with the computer environment, basic Linux commands and fundamental molecular biology. Furthermore, VNC software, housing the “Advanced Bioinformatics” course, has been used throughout this study. This system allowed several connections (VNC Viewer Software) to the same desktop (VNC Server Software), providing an invaluable tool for collaborative or shared working through secure logon into the Internet. However, VNC server address and password is not provided. Finally, all projects’ reports are created in electronic form only.

2.2.2. Exhaustive comparison between two versions of T4 lysozyme

2.2.2.1. Retrieving nucleotide and amino acid sequence

To begin, genetic sequences for wild-type *t4l* were obtained. Specifically, the National Center for Biotechnology Information (NCBI) website (<http://www.ncbi.nlm.nih.gov/>) was used to acquire the nucleotide sequence of wild-type *t4l*; “T4 Lysozyme” was the search query. From the results list, “e intracellular N-acetyl muramidase; lysis from within [Enterobacteria phage T4]” was chosen as corresponding to wild-type *t4l*. Also from the same entry page, the link “Nucleotide” under “Links” on the right side of the window was selected; from that results list “Enterobacteria phage T4, complete genome” was chosen. Enterobacteria phage T4 linear DNA size is 168,903 bp with accession no. NC_000866.4 and/or GI 29366675. Then, gene with location complement (66503..66997) was clicked on, which is gene “e” (its locus tag is “T4p123” and GeneID 1258585); nucleotide sequence was obtained by scrolling down to the feature table, denoted by the “FEATURES” key in the left-hand column on the window. Following, the nucleotide sequence was obtained in FASTA format, by selecting FASTA. Next, under “All Links from this record,” “Protein” then “e Lysozyme murein hydrolase [Enterobacteria phage T4]” were selected to get protein translation of coding regions from sequence records in the current set. Protein size was 164 a.a.; accession no. NP_049736.1 and/or GI 9632737. The protein sequence in FASTA format was acquired and saved, by selecting FASTA.

Moreover, since synthetic T4L nucleotide sequence was known; BLAST algorithm at NCBI (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) was used directly to run a database similarity for that gene. The first-best matched under alignments of query was “Synthetic T4 Lysozyme gene, complete cds,” with accession no. GI: M18375.1; it has linear DNA with 518 bp and has been selected as an alternative for our synthetic *t4l* for preliminary steps. Then, protein-coding sequence (CDS) 24..514 region was selected and represented Lysozyme. Next, the new window was for synthetic T4L gene, complete cds, and its size, 495 bp, the nucleotide sequence of which was obtained in FASTA format, by selecting FASTA. Thus, this nucleotide sequence was considered the most similar sequence to our synthetic *t4l* in the NCBI database,

which can be used for the initial steps. In addition, this protein size was 164 aa; with protein ID AAA72629. The protein sequence was obtained in FASTA format, by selecting the FASTA link.

2.2.2.2. Database similarity and multiple sequence alignment

BLAST algorithm at NCBI was used to compare a database similarity for “e Lysozyme murein hydrolase [Enterobacteria phage T4]” amino acid sequence, which was selected to represent wild-type T4L, to all other sequences in database. This was done by clicking the “protein blast” link, followed by pasting amino acid sequence in the query sequence textbox. Also, “Non-redundant protein sequences (nr)” under “Database” and “blastp (protein-protein BLAST)” under “Algorithm” were selected before running the BLAST. The ten best-matched homologous proteins under alignments of query were recorded. Subsequently, ten selected homologs amino acid sequences were used in addition to the original amino acid sequence to create multiple sequence alignments. Regions of similarity, divergence and mutations were highlighted using Clustalx application with BLOSUM 30 for substitution matrices. Then, conserved and non-conserved positions were determined. Also, the percentages of identities between the original amino acid sequence and each additional sequence in the alignment clarified distantly related members of a family and the evolution of the target sequence. Similarly, BLAST algorithm at NCBI was used again to run a database similarity for the obtained amino acid sequence, which is “Synthetic T4 Lysozyme gene, complete cds,” to all sequences in a database, using the same approach as with wild-type T4L nucleotide sequence. The ten first-best matched under alignments of query were recorded.

Nevertheless, translating our synthetic T4L nucleotide sequence into a protein was required to perform a multiple sequence alignment and along with further steps. Thus, jemboss application of European Molecular Biology Open Software Suite (EMBOSS) package was used for the gene nucleotide sequence to be translated to amino acid sequence. First, Open Reading Frames (ORFs) were identified using Plot Potential Open Reading Frames (PLOTORF). F1 frame was the longest ORF with 498 nucleotides. Thereafter, Finds and Extracts Open Reading Frames (GETORF) option was used to translate that area obtained from PLOTORF. Therefore, numbers 7 and 498 were start codon and stop codon, respectively.

Also, amino acid sequence of synthetic *t4l* was acquired in FASTA format. Then, the sequence of that region determined above was translated using Translate Nucleotide Acid Sequence (TRANSEQ) with parameters derived from results of previous steps, to get the exact amino acid sequence of synthetic *t4l* in FASTA format. Furthermore, the amino acid sequence obtained was compared using GETORF with the amino acid sequence of Lysozyme [synthetic constructed], and using BLAST algorithm at NCBI, an identity of 99% (162/164) was found. Additionally, *freak* application, reads nucleotide sequences and generates the used GC nucleotide frequency, and the output was in plot.

All amino acid sequences of synthetic *t4l* plus 10 selected homologs amino acid sequences were used to create multiple sequence alignments. Then, the Clustalx application was applied again using BLOSUM 30 for substitution matrices. Also, conserved and non-conserved positions were determined. Finally, the percentages of identities between synthetic *t4l* and the other sequence in the alignment clarified distantly related members of a family and the evolution of the target sequence.

2.2.2.3. Motifs profile

While “e Lysozyme murein hydrolase [Enterobacteria phage T4]” protein with accession no. NP_049736.1 on NCBI is represented in the wild-type T4L protein, scanning its sequence has permitted exploration of its motifs and the search for conserved motifs in the sequence. Therefore, Expert Protein Analysis System (ExPASy) proteomic server of the Swiss Institute Bioinformatics (SIB) database (<http://expasy.org/>) was used to find out whether wild-type T4L has an entry on ExPASy or not. From that server, Protein Knowledge database (UniProtKB) was selected and accession no. NP_049736.1 was used as the inquiry search. However, only “P00720” accession no. for T4L protein was found in that database. Consequently, the link “PR00684” under the section “Family and domain databases” / “PRINTS” block was selected to determine the number of motifs in this protein structure along with recording the sequence of each motif to find their exact location within the wild-type T4L amino acid sequence (Table 7). The number of motifs and their sequences were obtained using wild-type T4L protein and thus, they are

already known; therefore, the search for the sequence of each motif was recorded to find their location in the synthetic *t4l* amino acid sequence. Finally, recording how well they are conserved was reported.

2.2.2.4. Analyzing protein 3D structure

The 3D protein structure of the wild-type T4L is required to study the protein structure. Thus, Protein Data Bank (PDB), a member of the Research Collaboratory for Structural Bioinformatics (RCSB) (<http://www.rcsb.org/pdb/home/home.do>), was explored to find out the 3D protein structure. “Phage T4 lysozyme” was the query search and 2LZM as PDB ID was selected for this protein, because it was the only entry mentioned in its Molecular Description “P00720” entry for UniProtKB. Then, the structure in PDB format was downloaded along with its experimental details.

First, Visual Molecular Dynamics (VMD) software was used to visualize the protein structure by creating graphical representations and determining the number of alpha helix and beta-strand regions. Different models were generated with graphic representation. Also, looking at the amino acid sequence of wild-type T4L protein structure was achieved using Sequence Viewer option with the same software that determined the number of alpha helix regions (marked as pink-purple blocks) and beta-strand regions (represented as yellow blocks).

However, there was no entry in any database for synthetic *t4l*. Therefore, there was no experimentally determined 3D protein structure for it in RCSB PDB. One of the ExPASy tools used to create a model for synthetic *t4l* protein was the Tertiary structure-Homology modeling-Automated homology modeling program using neural networks (ESyPred3D) (<http://www.fundp.ac.be/sciences/biologie/urbm/bioinfo/esypred/>). This modeling web server requires the protein sequence and 3D protein structure to be used as template. A 3D model of that protein sequence was built using the 3D wild-type lysozyme protein structure (PDB ID 2LZM) as template as was selected before, which shared 98.8% identities with the sequence query (using the align program). Also, the structure of protein in PDB format was downloaded with no experimental details. Then finally, the protein structure was visualized using VMD and graphical representations created.

2.2.3. Comparison of T4 lysozyme 3D protein structures

The comparison of wild-type and synthetic T4L proteins' structures was also done using VMD software. This was accomplished by performing structural alignment and calculating Root Mean Sequence Deviation (RMSD). RMSD measures the average distance between the atoms of superimposed proteins, which makes a quantitative comparison between the structure of the target protein and the structure of native state. A structural alignment was performed using Multiple Alignment application. Backbone, all atoms, and carbon alpha coordinators were calculated and reported after superposition of the two 3D protein structures. In addition, the output of multiple alignments read with alignment editor Jalview in jemboss package. This was to calculate the percentage of sequence identity, which was obtained from those 3D protein structures, with pairwise alignment.

2.3. Results and Discussion

2.3.1. Retrieving nucleotide sequence and protein translation

Based on NCBI search, wild-type *t4l* is represented on “e intracellular N-acetyl muramidase; lysis from within [Enterobacteria phage T4]” entry. Some information related to this entry is summarized in Table 1. The size of the gene code is 459 bp: linear DNA. This sequence was the starting point for the study of wild-type *t4l*; the nucleotide sequence in FASTA format is shown in Figure 1-A. Moreover, using NCBI, protein translation of coding regions of wild-type T4L has been acquired; its accession number is NP_049736. The protein is made up of 164 a.a. The rest of the studies all depend on this protein sequence, making it important; also, the protein sequence in FASTA format is shown in Figure 1-B. In comparison, based on BLAST algorithm at NCBI search, the best matched under alignments for the synthetic *t4l* nucleotide sequence was “synthetic T4 lysozyme gene, complete cds.” Information related to this selection is summarized in Table 2. As well, the size of this code is 495 bp, linear DNA for Lysozyme; CDS 24..514 region selected represented Lysozyme. The nucleotide sequence in FASTA format is shown in Figure 2-A. Finally, on NCBI protein translation of coding regions of Lysozyme [synthetic construct]

protein sequence was obtained; it is made up of 164 a.a. and with accession no. AAA72629. The protein sequence in FASTA format is shown in Figure 2-B.

2.3.2. Protein database similarities and multiple sequence alignment

Protein database similarity was performed for wild-type T4L amino acid sequence. The ten first best-matched amino acid sequences similarity are recorded along with their organism name, score, E-value and percentage of identity and similarity to search inquiry, using BLAST algorithm at NCBI (Table 3). All those sequences were also related to bacteriophage T4 organism, and thus, required performance of a multiple sequence alignment for them. Additionally, wild-type T4L amino acid sequence and amino acid sequences greater than the ten first best-matched were used to perform a multiple sequence alignment, using Clustalx application. From the Clustalx result, there were two non-conserved, six strongly conserved, and 156 fully conserved positions identified with the position of residues (Table 4). Also, Clustalx gave 98% percentage of identities between wild-type T4L and each individual sequence of the ten first best-matched amino acid sequences (Table 5. A). Similarly, using BLAST algorithm at NCBI the ten first matched of “Lysozyme [synthetic construct]” amino acid sequence similarity were recorded along with their organism name, score, E-value and percentage of identity and similarity to search inquiry (Table 4). However, the amino acid sequence for synthetic T4L required performance of multiple sequence alignments.

2.3.2.1 EMBOSS package analysis

Firstly, F1 frame had the longest ORFs using PLOTORF application; this allowed a rapid visual overview of the distribution of ORFs in the six frames of our synthetic T4L sequence (Figure 3). As well, there are 498 nucleotides. Thereafter, the GETORF application was used for the translation of region between 7 and 498 codons (Figure 4 A). Then identification of ORFs and the start and end nucleotide positions; the TRANSEQ application identified a protein coding region (Figure 4 B). Moreover, the comparison between the TRANSEQ result and the translation of “Lysozyme [synthetic constructed]” sequence was done using BLAST algorithm at NCBI, resulting in 99% (162/164) identity. In addition, the

plot result of freak application showed the G+C nucleotides frequency of wild-type and synthetic T4L DNA molecules; either guanine (G) or cytosine (C). Basically, the GC pair is bound by three hydrogen bonds, while AT pairs are bound by two hydrogen bonds. DNA with high GC-content is more stable than DNA with low GC-content; however, the hydrogen bonds do not stabilize the DNA significantly, and stabilization is due mainly to stacking interactions. Consequently, in PCR experiments the GC-content of primers is used to predict their annealing temperature to the template DNA. A higher GC-content level indicates a higher melting temperature, and this was used for considerations in the following study (Figure 5). Our synthetic T4L amino acid sequence plus the ten first best-matched amino acid sequences were used to perform a multiple sequence alignment using the Clustalx application. From the Clustalx results, there are two non-conserved, five strongly conserved, and 157 fully conserved positions identified in the multiple sequence alignment (Table 5). Also, Clustalx gave the percentage of identities of 98% between our synthetic T4L and the other sequence in alignment (Table 6).

2.3.3. Motifs analysis

To begin, the pattern of domain of T4L is not found in the ExPASy server; however, in the PRINTS server there are seven motifs within T4L protein structure. Therefore, the sequence of those seven motifs was recorded and determined to be in the T4L amino acid sequence using Motifs Search tool of GenomeNet (<http://www.genome.jp/tools/motif/>), along with position-specific score matrix for each motif (Table 8). The PRINTS is a compendium of protein motif fingerprints. While, a fingerprint is a group of conserved motifs used to characterize a protein family. During the search of each entry in the database, a set of blocks (simple fingerprints) is converted to a position-specific score matrix (profile). This matrix is then used to score the motif found in the query sequence. Consecutively, the motif blocks in the PRINTS database are calibrated based on the distribution of matches determined in a search as described by the authors. Each raw score is divided by an estimated score of 99.5% of true negative sequence and multiplied by 1000. Thus, a derived score of 1000 infers that match would be at the 99.5% false matches, which means scores less than and around 1000 are probably not important.

Likewise, synthetic T4L does not have an entry in the ExPASy server. Therefore, the determined seven motifs of wild-type T4L amino acid have been applied to the synthetic one. Also, the sequences of those seven motifs have been determined in the synthetic T4L amino acid sequence using the same application that was used with wild-type T4L protein (Table 8). After comparing the score results of found motifs, one of seven motifs had an insignificant difference with score results, which was T4LYSOZYME1 motif. The difference was inversely related to the modification made to the synthetic T4L nucleotide. In other words, there were two single residues replaced into the same protein motif fingerprint as follows: M6 (Methionine at position 6) in wild-type was replaced by K6 (Lysine) in synthetic one. Also, G12 (Glycine at position 12) in wild-type was substituted by R12 (Arginine) in the synthetic one, and that both alterations developed changing the score from 1265 with wild-type to 1227 with synthetic T4L protein. Additionally, the two amino acids in wild-type T4L protein have nonpolar properties and were replaced with two basic properties, according to amino acids properties.

2.3.4. Analyzing T4 lysozyme 3D protein structure

Protein Data Bank (PDB) ID 2LZM is a representation of T4L 3D protein structure; described as “Structure of bacteriophage T4L refined at 1.7 angstroms resolution.” It is related to bacteriophage T4 and classified as a hydrolase. Additionally, the X-ray diffraction experimental method for it is 1.7 Å resolution. This 3D structure will assist comparison and analysis of the synthetic T4L 3D protein structure. Thus, a graphic representation using VMD software was created to visualize the different chemical structure of T4L 3D protein structure to ultimately understand more about it. While using NewCartoon as Drawing Method and ResType as Coloring Method, the amino acid residues were coloured according to their chemical character, such as white for hydrophobic, green for polar uncharged, red for negatively charged, and blue for positively charged. Then, the sequence viewer application determined two regions for alpha helix and two regions for beta-strand in this protein. Again, synthetic T4L is a constructed gene and has no entry in any database. A model for synthetic T4L was created based on PDB ID 2LZM as a template for wild-type T4L protein structure using the ESyPred3D server. This

template shares 98.8% identities with the sequence query (using the align program); however, there is no experimental detail provided from the server.

2.3.5. Comparison of T4 lysozyme 3D protein structures

The wild-type and synthetic T4L share sequence homology; nevertheless, a residue at position 12 indicated variable conserved status based on multiple sequence alignment results for both wild-type and synthetic T4L amino acids sequence. Also, the residue at position 6 was fully conserved in wild-type yet variable with synthetic T4L amino acid sequence. Therefore, the Sequence Viewer application with VMD software was applied to generate amino acid sequences from 3D protein structures of wild-type and synthetic T4L; then the generated amino acid sequences were aligned to obtain their sequence homology. The results showed there were two altered amino acid residues at positions 6 and 12 with synthetic T4L protein (Figure 6); confirming the previous finding.

Moreover, comparison of both proteins' structures determined how structurally similar they are. After the superposition of 3D protein structures, the reported backbone using RMSD for synthetic T4L compared to wild-type was 0.17. In addition, carbon alpha and all-atoms of RMSD were 0.14 and 5.80, respectively, concluding they are close together structurally. Also, the sequence identity of both structures was 98.78%, as determined using the alignment editor Jalview in jemboss package. Finally, based on all the information related to both RMSD and amino acid sequence, the two altered amino acid residues (M6 \rightarrow K6, G12 \rightarrow R12) were labeled using VMD software for each 3D protein structure, wild-type and synthetic (Figure 7).

Conclusion

In this study, we confirmed that there was no published study that used our synthetic T4L gene, which has been previously worked on. Wild-type and synthetic T4L have variation in the size of nucleotides, 459 bp and 491bp, respectively; however, both have the same amino acid size of 164 a.a. Additionally, the results of multiple sequence alignments indicate both amino acids sequences have high fully conserved residues in their sequences similarity obtained from the database. Most of the sequences similarities

obtained from the database are related to Enterobacteria phage T4 and their percentage of identity to target sequences is high, at approximately 98%. Therefore, more studies are needed for the evolution of T4L. Moreover, the conserved motifs were high with wild-type T4L as well the synthetic; one out of seven had an insignificant difference in the score results, which was T4LYSOZYME1 motif, in this study. Also, the comparison of 3D protein structure exhibited how both protein structures have structural similarity through the calculation of RSMD; 0.17 for backbone coordinate and 0.14 for carbon alpha coordinate. As well, structure sequence identity was 98.87% using the alignment of editor Jalview. In conclusion, wild-type and synthetic T4L supposedly share the same function of hydrolysis, yet there is variability at position 6 and 12 in amino acids sequence.

Figure and Table legend

Figure:

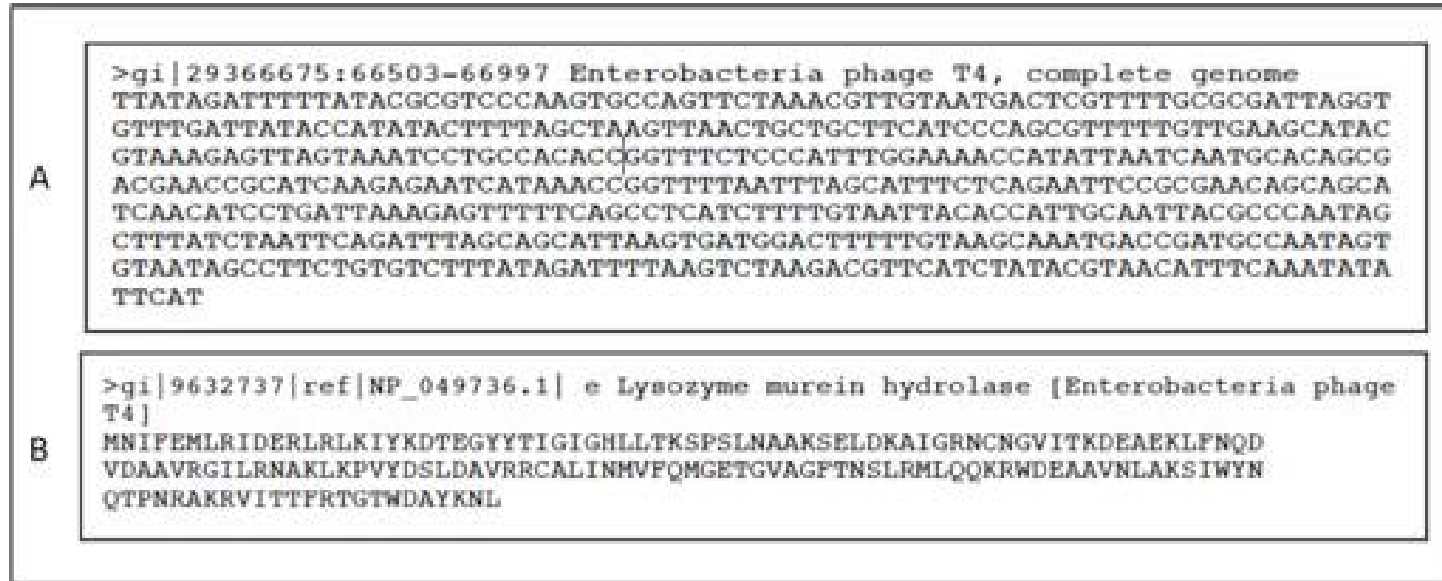


Figure 1: Wild type T4L (A) nucleotide and (B) amino acid sequences in FASTA format

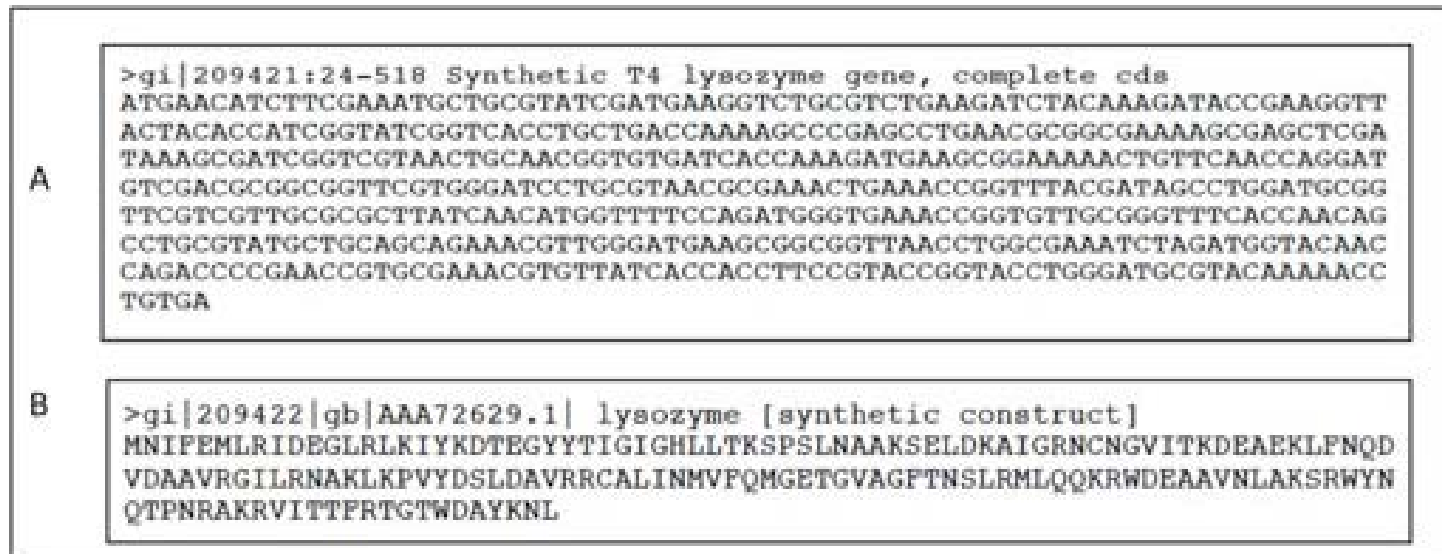


Figure 2: “Synthetic T4 Lysozyme gene, complete cds” (A) nucleotide and (B) amino acid sequences in FASTA format

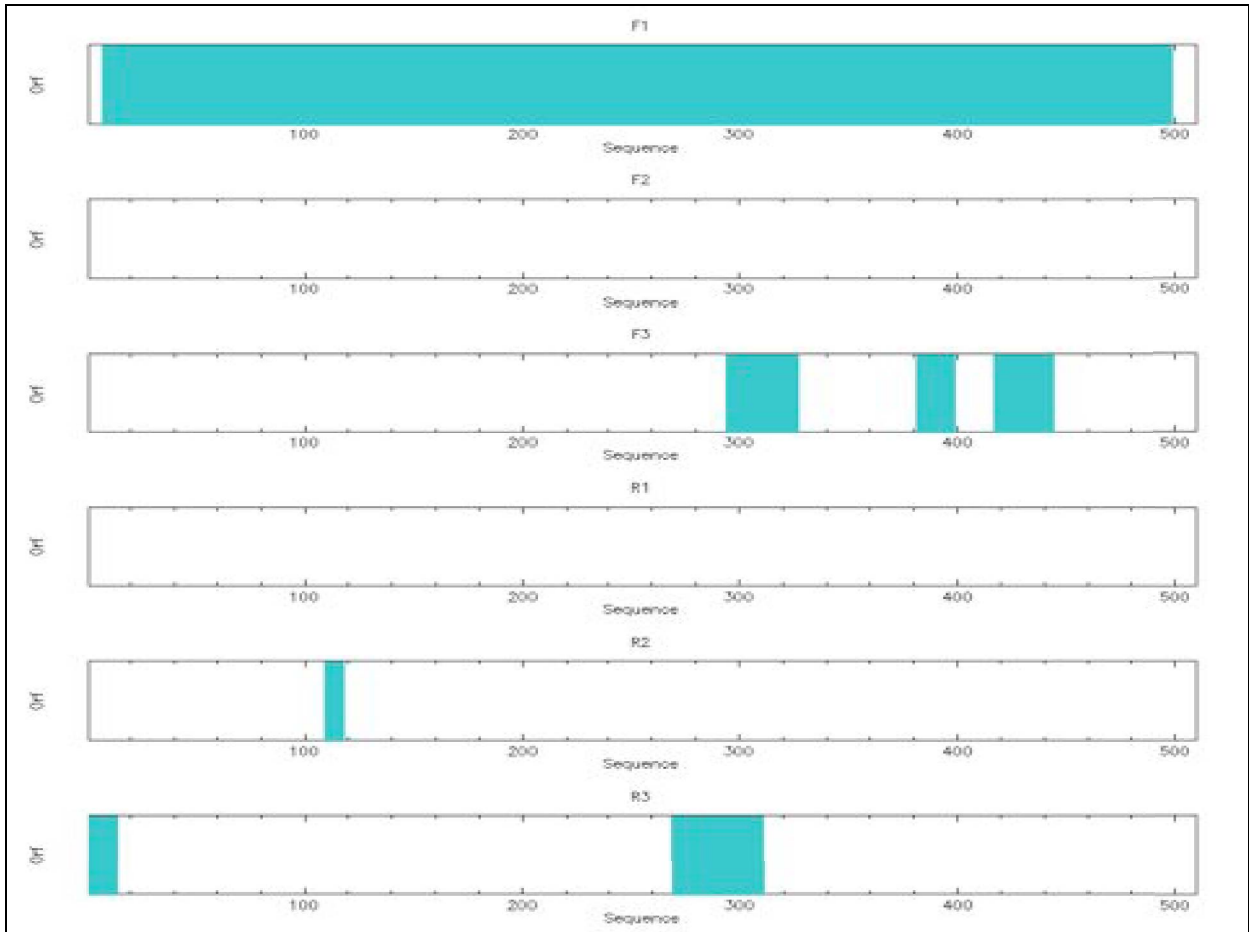


Figure 3: PLOTORF application frames result.

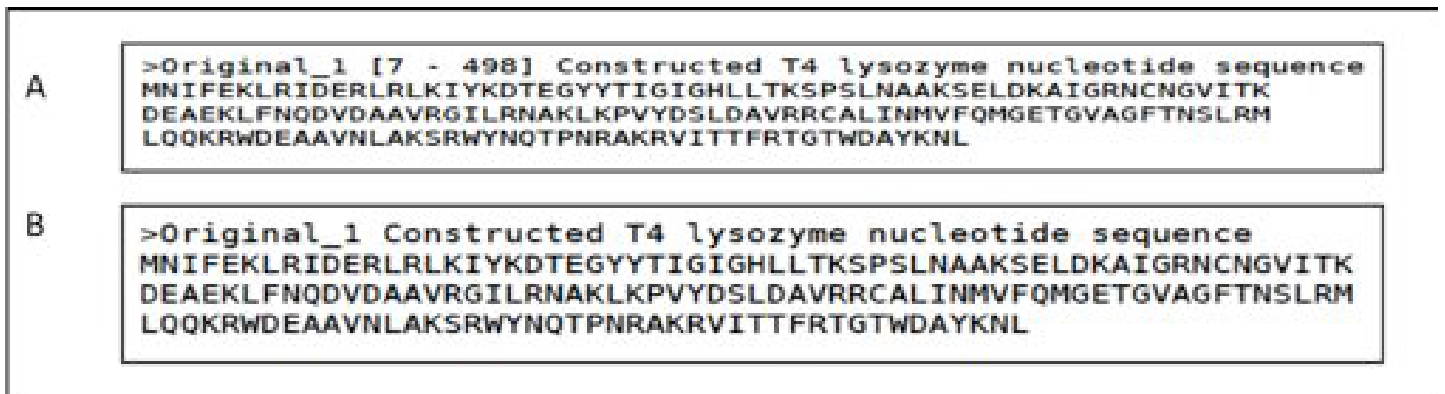


Figure 4: Translation of synthetic T4L nucleotides by (A) GETORF and (B) TRANSEQ

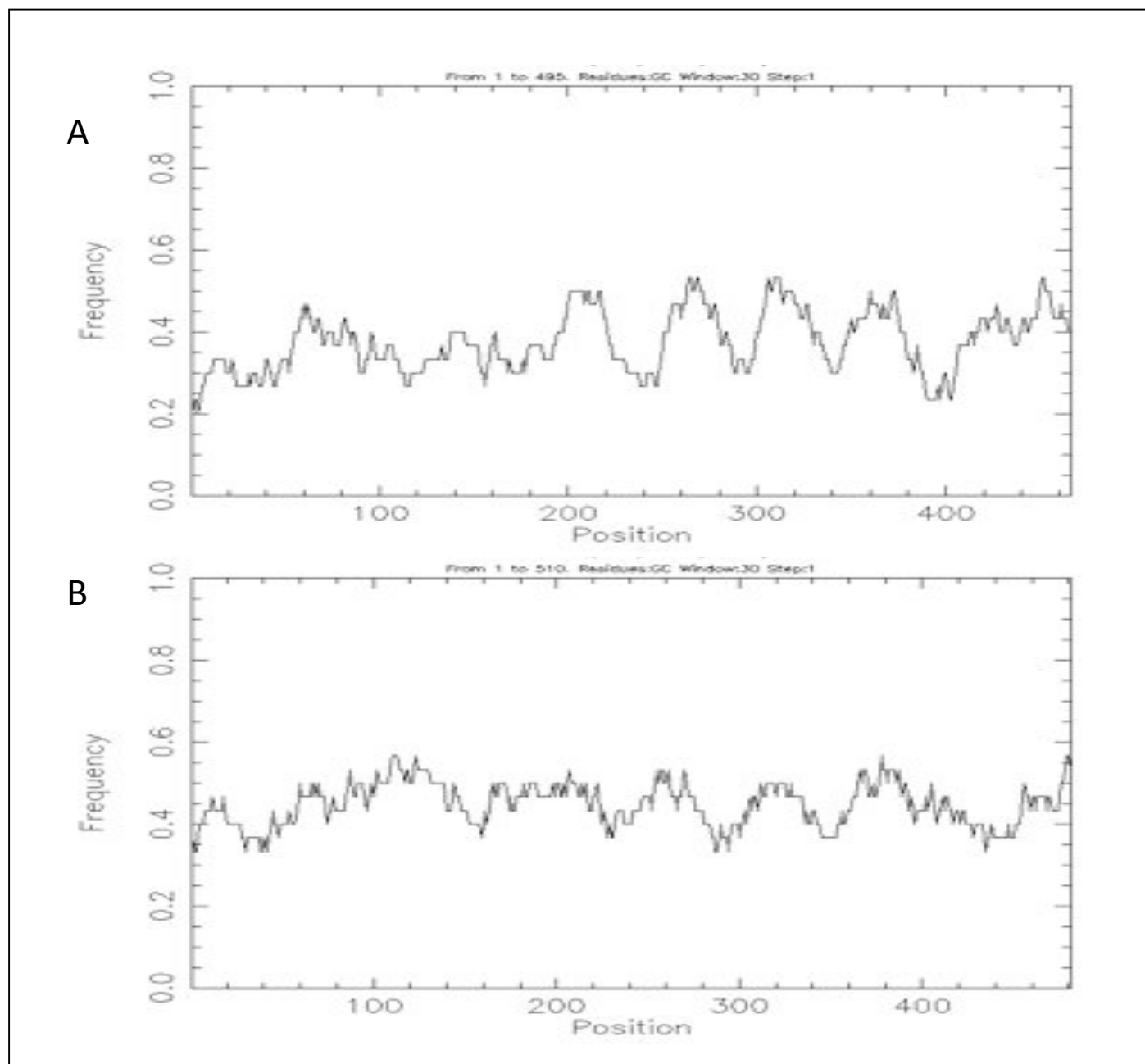


Figure 5: EMBOSS-freak plot results, for wild type (A) and synthetic T4L nucleotide sequence (B)

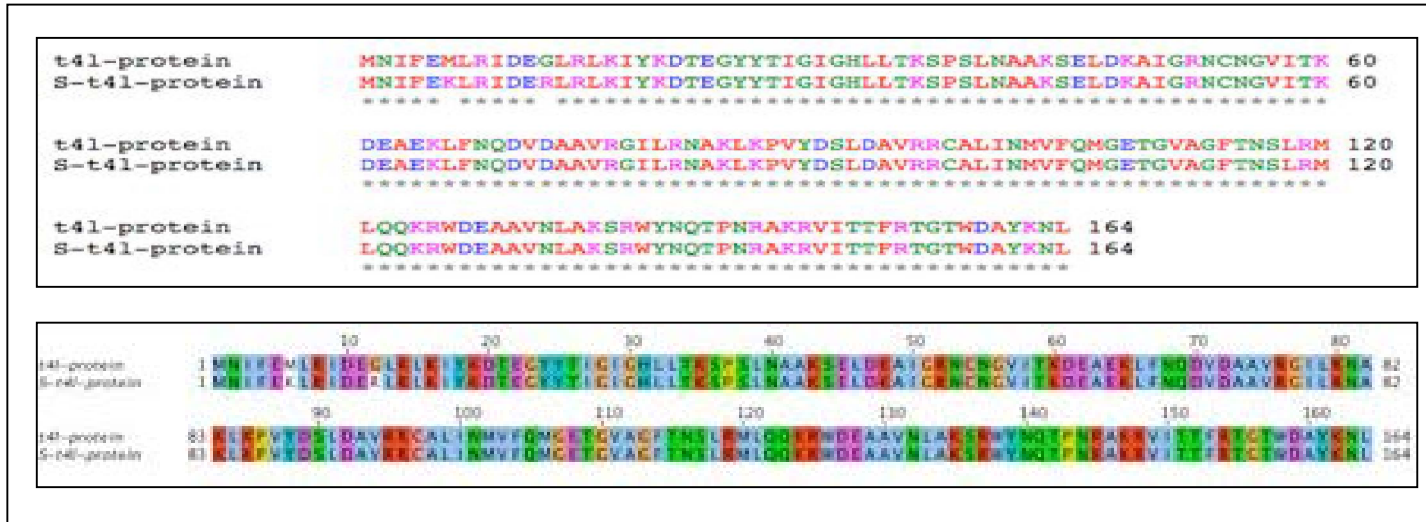


Figure 6: Amino acid sequence alignment of wild type (t4l-protein) and synthetic T4L (S-t4l-protein)

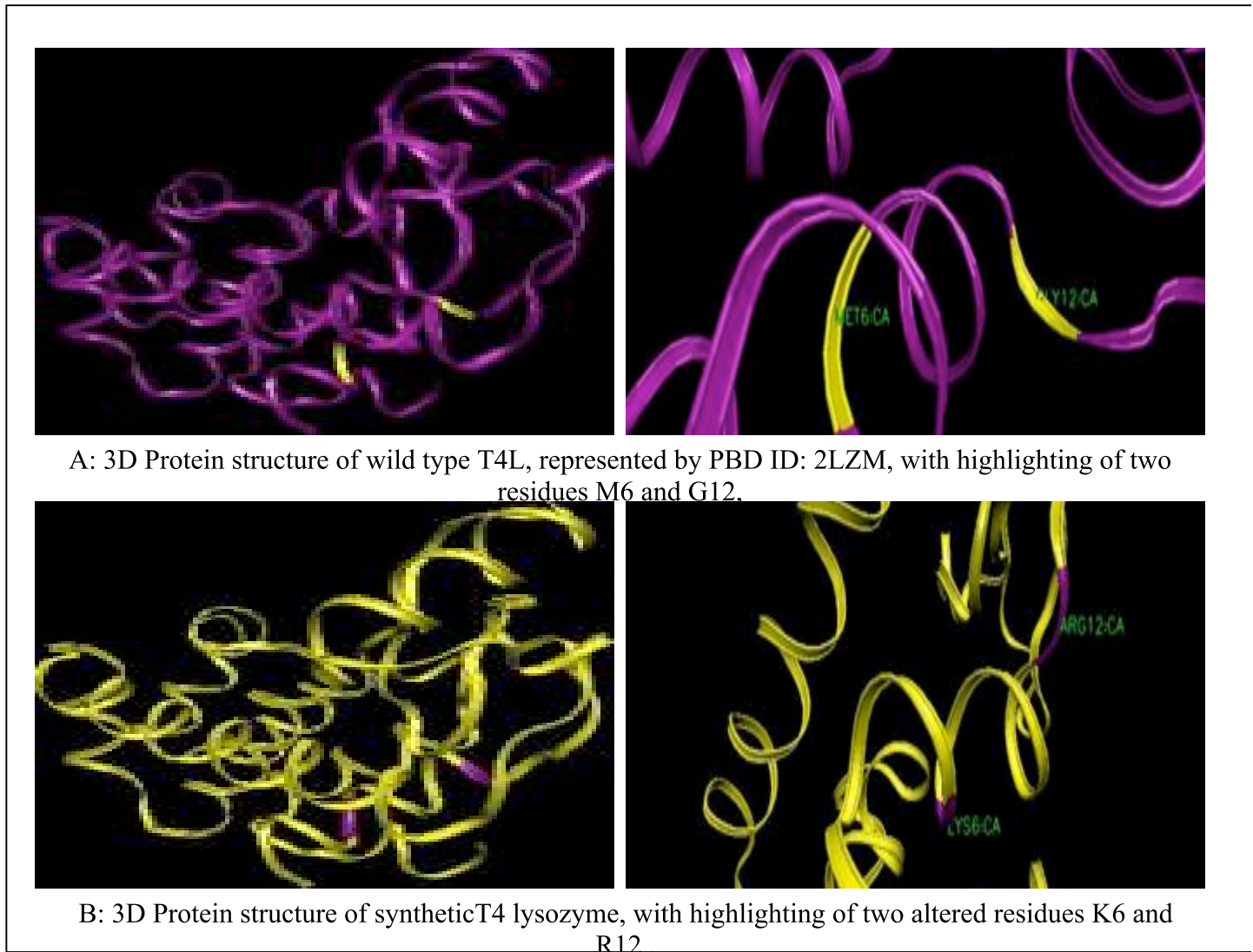


Figure 7: 3D protein structures for (A) wild type and (B) synthetic T4L with highlighting of the altered residues' positions on synthetic structure and the wild type for comparison

Table:**Table 1:** General Information about wild type *t4l*, from NCBI

Gene ID	1258585
Gene Description	Intracellular N-acetyl muramidase; lysis from within
Locus tag	T4p124
Gene type	Protein coding
RefSeq status	Provisional
Organism	Enterobacteria phage T4
Lineage	Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae; T4-like viruses

Table 2: General Information of gene represented synthetic T4L

Length	518
Score	342 bites (378)
E-value	2e-90
Identification	375 / 499 (76%)
Gap	0 / 499 (0%)
Accession	M18375

Table 3: Sequence produce significant alignment

No	Accession	Description	Total Score	Query coverage	E value
1	NP_049736.1	e Lysozyme murein hydrolase [Enterobacteria phage T4] >gb AAD42568.1 AF158101_155 e Lysozyme murein hydrolase [Enterobacteria phage T4] >emb CAA28212.1 unnamed protein product [Enterobacteria phage T4]	340	100%	4e-92
2	P00720.1	RecName: Full=Lysozyme; AltName: Full=Endolysin; AltName: Full=Lysis protein; AltName: Full=Muramidase Bacteriophage T4 Lysozyme	334	100%	2e-90
3	1L17_A	Chain A, Hydrophobic Stabilization In T4 Lysozyme Determined Directly By Multiple Substitutions Of Ile 3	333	100%	3e-90
4	1DYC_A	Chain A, Determination Of Alpha-Helix Propensity Within The Context Of A Folded Protein: Sites 44 And 131 In Bacteriophage T4 Lysozyme	333	100%	4e-90
5	1L14_A	Chain A, Contributions Of Hydrogen Bonds Of Thr 157 To The Thermodynamic Stability Of Phage T4 Lysozyme	333	100%	4e-90
6	1L52_A	Chain A, Structural And Thermodynamic Analysis Of The Packing Of Two Alpha-Helices In Bacteriophage T4 Lysozyme	333	100%	4e-90
7	3C7W_A	Chain A, Contributions Of All 20 Amino Acids At Site 96 To The Stability And Structure Of T4 Lysozyme	333	100%	5e-90
8	149L_A	Chain A, Conservation Of Solvent-Binding Sites In 10 Crystal Forms Of T4 Lysozyme	333	100%	5e-90
9	1DYF_A	Chain A, Determination Of Alpha-Helix Propensity Within The Context Of A Folded Protein: Sites 44 And 131 In Bacteriophage T4 Lysozyme	333	100%	6e-90
10	1DYD_A	Chain A, Determination Of Alpha-Helix Propensity Within The Context Of A Folded Protein: Sites 44 And 131 In Bacteriophage T4 Lysozyme	333	100%	6e-90
11	1L38_A	Chain A, Contributions Of Engineered Surface Salt Bridges To The Stability Of T4 Lysozyme Determined By Directed Mutagenesis	333	100%	6e-90

No	Accession	Description	Total Score	Query coverage	E value
1	P00720.1	RecName: Full=Lysozyme; AltName: Full=Endolysin; AltName: Full=Lysis protein; AltName: Full=Muramidase Bacteriophage T4 Lysozyme	340	100%	3e-92
2	1L17_A	Chain A, Hydrophobic Stabilization In T4 Lysozyme Determined Directly By Multiple Substitutions Of Ile 3	340	100%	4e-92
3	1DYC_A	Chain A, Determination Of Alpha-Helix Propensity Within The Context Of A Folded Protein: Sites 44 And 131 In Bacteriophage T4 Lysozyme	340	100%	4e-92
4	1L14_A	Chain A, Contributions Of Hydrogen Bonds Of Thr 157 To The Thermodynamic Stability Of Phage T4 Lysozyme	340	100%	5e-92
5	1L52_A	Chain A, Structural And Thermodynamic Analysis Of The Packing Of Two Alpha-Helices In Bacteriophage T4 Lysozyme	340	100%	5e-92
6	3C7W_A	Chain A, Contributions Of All 20 Amino Acids At Site 96 To The Stability And Structure Of T4 Lysozyme	339	100%	6e-92
7	1DYF_A	Chain A, Determination Of Alpha-Helix Propensity Within The Context Of A Folded Protein: Sites 44 And 131 In Bacteriophage T4 Lysozyme	339	100%	7e-92
8	149L_A	Chain A, Conservation Of Solvent-Binding Sites In 10 Crystal Forms Of T4 Lysozyme	339	100%	7e-92
9	1DYD_A	Chain A, Determination Of Alpha-Helix Propensity Within The Context Of A Folded Protein: Sites 44 And 131 In Bacteriophage T4 Lysozyme	339	100%	7e-92
10	1L33_A	Chain A, Contributions Of Left-Handed Helical Residues To The Structure And Stability Of Bacteriophage T4 Lysozyme	339	100%	8e-92

Table 4: Result of multiple sequence alignment (MSA) for wild type T4L amino acid sequences compared to its 10 selected amino acid similarity and result of MSA for synthetic T4L amino acid.

Amino Acid	Conserved Status with Position of residue(s)		
	Variable	Strongly	Fully
WT T4L	12, 137	96, 123, 131, 152, 157	(1-11), (13-95), (97-122), (124-130), (132-136), (138-151), (153-156), (158-164)
Synthetic T4L	6, 12	3, 96, 131, 152, 157	1, 2, 4, 5, (7-11), (13-95), (97-130), (132-151), (153-156), (158-164)

Table 5: Percentage of identity to wild type (a) and synthetic T4L (b) by Clustalx
(a)

No.	Sequence identify by organisms name	% Identity
1	Enterobacteria phage T4 (Bacteriophage T4)	100
2	Enterobacteria phage T4	98
3	Enterobacteria phage T4	98
4	Enterobacteria phage T4	98
5	Enterobacteria phage T4	98
6	Enterobacteria phage T4	98
7	Enterobacteria phage T4	98
8	Enterobacteria phage T4	98
9	Enterobacteria phage T4	98
10	Enterobacteria phage T4	98

(b)

No.	Sequence identify by organisms name	% Identity
1	Synthetic T4 Lysozyme (constructed gene)	100
2	Synthetic construct	98
3	Enterobacteria phage T4	98
4	Enterobacteria phage T4	98
5	Enterobacteria phage T4	98
6	Enterobacteria phage T4	98
7	Enterobacteria phage T4	98
8	Enterobacteria phage T4	98
9	Enterobacteria phage T4	98
10	Enterobacteria phage T4	98

Table 6: Sequences of seven motifs included in wild type T4L protein

Motifs	Motifs sequence
T4LYSOZYME1	FEMLRIDEGLRLKIYKDTEG
T4LYSOZYME2	YYTIGIGHLLTKSPSLNAA
T4LYSOZYME3	GRNCNGVITKDEAEKLFNQD
T4LYSOZYME4	DAAVRGILRNAKLKPVYDSL
T4LYSOZYME5	RRCALINMVFQMGETGVAGF
T4LYSOZYME6	LRMLQQQRWDEAAVNLAISR
T4LYSOZYME7	WYNQTPNRAKRVITTFRTGTWD

Table 7: Sequence of seven motifs determined in wild type and synthetic T4L protein with position, sequence, score, and description

Protein	Motif	Position	Sequence	Score	Description
Wild Type T4L	T4LYSOZYME1	4..23	FEMLRIDEGLRLKIYKDTEG	1280	Phage T4 lysozyme signature
	T4LYSOZYME2	24..42	YYTIGIGHLLTKSPSLNAA	1278	Phage T4 lysozyme signature
	T4LYSOZYME3	51..70	GRNCNGVITKDEAEKLFNQD	1301	Phage T4 lysozyme signature
	T4LYSOZYME4	72..92	DAAVRGILRNAKLKPVYDSL	1322	Phage T4 lysozyme signature
	T4LYSOZYME5	95..114	RRCALINMVFQMGETGVAGF	1301	Phage T4 lysozyme signature
	T4LYSOZYME6	118..137	LQQRWDEAAVNLAISR	1265	Phage T4 lysozyme signature
	T4LYSOZYME7	138..159	WYNQTPNRAKRVITTFRTGTWD	1342	Phage T4 lysozyme signature
Synthetic T4L	T4LYSOZYME1	4..23	FEKLRIDERLRLKIYKDTEG	1227	Phage T4 lysozyme signature
	T4LYSOZYME2	24..42	YYTIGIGHLLTKSPSLNAA	1278	Phage T4 lysozyme signature
	T4LYSOZYME3	51..70	GRNCNGVITKDEAEKLFNQD	1301	Phage T4 lysozyme signature
	T4LYSOZYME4	72..92	DAAVRGILRNAKLKPVYDSL	1322	Phage T4 lysozyme signature
	T4LYSOZYME5	95..114	RRCALINMVFQMGETGVAGF	1301	Phage T4 lysozyme signature
	T4LYSOZYME6	118..137	LRMLQQQRWDEAAVNLAISR	1265	Phage T4 lysozyme signature
	T4LYSOZYME7	138..159	WYNQTPNRAKRVITTFRTGTWD	1342	Phage T4 lysozyme signature

References

- Faber, H. R., & Matthews, B. W. (1990). A mutant T4 lysozyme displays five different crystal conformations. *Nature*, *348*(6298), 263-266. doi: 10.1038/348263a0
- Henrissat, B. (1991). A classification of glycosyl hydrolases based on amino acid sequence similarities. [Research Support, Non-U.S. Gov't]. *Biochem J*, *280* (Pt 2), 309-316.
- Henrissat, B. and Coutinho P.M. (1999) Carbohydrate-Active Enzymes server. Available at <http://www.cazy.org/>. Accessed on March 20, 2011.
- Henrissat, B., & Bairoch, A. (1993). New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. [Research Support, Non-U.S. Gov't]. *Biochem J*, *293* (Pt 3), 781-788.
- Henrissat, B., & Bairoch, A. (1996). Updating the sequence-based classification of glycosyl hydrolases. [Letter]. *Biochem J*, *316* (Pt 2), 695-696.
- Leiman, P. G., Kanamaru, S., Mesyanzhinov, V. V., Arisaka, F., & Rossmann, M. G. (2003). Structure and morphogenesis of bacteriophage T4. *Cell Mol Life Sci*, *60*(11), 2356-2370. doi: 10.1007/s00018-003-3072-1
- Luscombe, N. M., Greenbaum, D., & Gerstein, M. (2001). What is bioinformatics? A proposed definition and overview of the field. [Review]. *Methods Inf Med*, *40*(4), 346-358.
- Miller, E. S., Kutter, E., Mosig, G., Arisaka, F., Kunisawa, T., & Ruger, W. (2003). Bacteriophage T4 genome. *Microbiol Mol Biol Rev*, *67*(1), 86-156, table of contents.
- RealVNC Ltd. (n.d.). Retrieved from <http://www.realvnc.com/index.html>
- Sasa, H. (2012). Prototypical implementation and empirical evaluation of a combined communication and collaboration system, in the context of classroom teaching – Educator Collaborator (EC). *International Journal of Advanced Trends in Computer Science and Engineering*, *1*(1), 18 - 24.
- Stothard, P. (2012). An introduction to Linux for bioinformatics. University of Alberta. Retrieved from http://www.ualberta.ca/~stothard/downloads/linux_for_bioinformatics.pdf
- Susan, L. E., & Stansfield, W. D. (2010). *Genetics* (5th ed.). New York: McGraw-Hill Companies.
- Trun, N., & Trempy, J. (2004). *Fundamental Bacterial Genetics*. Malden: Blackwell Publishing.
- Weaver, L. H., & Matthews, B. W. (1987). Structure of bacteriophage T4 lysozyme refined at 1.7 Å resolution. *J Mol Biol*, *193*(1), 189-199.

Chapter III: Expression of a Synthetic T4 Lysozyme Gene in *Escherichia coli*

Abstract

Although T4L enzyme is similar to HEWL, T4L is considered the most active lysozyme in sterilization. In this study, developed T4L gene with codon-based modification has been cloned into three vectors with different promoters for expression in a select strain of *Escherichia coli*. Briefly, expression of synthetic *t4l* was done by transforming three separate plasmids-bearing-synthetic *t4l* into two strains of *E. coli*, JM109 and BL21(DE3)pLysS. Using the agar halo assay, the later was tested for its sensitivity and/or resistance to the lysis action of lysozyme enzyme before the transformation-took place. Following the optimization of culture conditions for transformants; conventional PCR was used to detect and amplify the synthetic *t4l* gene cloned into the plasmids. Also, SDS-PAGE assay demonstrated that the recombinant T4L was expressed in all three transformants of *E. coli*, which named as BST4, BLT4, and BPT4, after breaking bacterial cell wall by sonication. In conclusion, resistance of BL21(DE3)pLysS strain to lysozyme activity was confirmed with our halo assays, which show this strain is a good choice for expression of this lysis enzyme. Also, the results of PCR showed that the synthetic *t4l* was successfully transformed into selected *E. coli* strain. Furthermore, the expression of that gene was shown on SDS-PAGE assay with an apparent molecular mass of approximately 18 kDa. However, characterization and measurement the recombinant T4L production still requires more research.

Key words: *E. coli*, T4 Lysozyme, PCR, SDS-PAGE

3.1. Introduction

Bacteriophage is a bacterial virus, it invades and controls bacterial metabolism, then terminates its incubation through bacterial lysis, as is the case with lytic phages (Sulakvelidze *et. al.*, 2001). One type of bacteriophage is bacteriophage T4 (also it is called phage T4), which infects the bacterial cell. Moreover, upon infecting a bacterium and producing progeny phage T4 only undergoes lytic growth (Trun & Trempey, 2004). Unlike animal viruses, infection of host cells by tailed phage is highly efficient because only one phage T4 particle is adequate for initial infection process. Additionally, upon infection the phage

shuts down host-specific nucleic acid and protein syntheses, thus ensuring sole production of its own components in amounts sufficient to assemble up to 200 progeny virus particles per infected cell (Leiman *et. al.*, 2003).

In more detail, an insertion of T4 DNA into bacterium follows four steps; first phage T4 initiates infection of a susceptible bacterium by recognizing the lipopolysaccharide cell surface receptors with the distal ends of its long tail fibers (LTFs). Subsequently, the short tail fibers (STFs) unravel from underneath the baseplate and bind irreversibly to the lipopolysaccharide cell surface receptors, thus securely anchoring the baseplate to the cell membrane. Meanwhile, the baseplate changes its conformation from hexagonal to star-shaped, causing contraction of the tail sheath. Then the tail sheath contracts, driving the internal tail tube through the peptidoglycan-containing outer membrane. After, T4L is secreted to digest the cell wall allowing penetration of the tail tube to the inner membrane where the DNA is released from the phage head through the tail tube into the host cell. Once T4 DNA is in the cell cytoplasm, it initiates replication of its own DNA with a highly organized and coordinated program of gene expression, using both replication and recombination (Trun & Trempey, 2004; Leiman *et. al.*, 2003); thereby releasing new replicated phage particles. In addition, Phage T4 is one of the most structurally complex virus particles known, devoting more than 40% of their genetic information to the synthesis and assembly of the prolate icosahedral heads, tails with contractile sheaths, and six tail fibers that contribute to their very high efficiency of infection (Miller *et. al.*, 2003).

T4L, produced by phage T4, is structurally similar to HEWL in its active site and its site of attack on the peptidoglycan cell wall (Akinalp *et. al.*, 2007). Moreover, it is a well-characterized protein; its molecular weight is approximately 18 kDa. Specifically, it cleaves β -1,4 glycosidic bonds between N-acetylmuramic acid (NAM) and N-acetylglucosamine (NAG) in the bacterial peptidoglycan (Bower *et. al.*, 1998) (Figure 1). Therefore, T4L is said to destroy bacteria by its muramidase activity. However, another study demonstrates that the C-terminus of T4L mediates its bactericidal and fungistatic activities. For example, in heat-denatured T4L, the enzymatic activity was completely abolished, yet unexpectedly

the antimicrobial functions remain preserved. Comparable results were also obtained with HEWL (During *et. al.*, 1999 and Dong *et. al.*, 2008). This knowledge opens up many new opportunities for optimization of lysozymes as antimicrobial agents in various applications using protein engineering.

The present study aimed to express a synthetic T4L-encoding gene in select *Escherichia coli* (*E. coli*) host strains, through the transformation of three plasmids; each bearing-synthetic *t4l*. Additionally, culture conditions were optimized for transformants in order to enhance the production of recombinant T4L.

3.2. Materials and Methods

3.2.1. Synthetic T4 Lysozyme Gene

A synthetic (modified) T4L gene, homologous to wild-type T4L which originated from phage T4, was constructed; a generous gift from former labmate Dr. Xi Chen. There were three strategies Dr. Chen developed regarding construction of synthetic *t4l*, which are as follows; first, construction of recombinant plasmids where wild-type *t4l* is controlled by a strong promoter. Secondly, production of new *t4l* sequence using codon bias. Lastly, recombinant T4L production is enhanced through optimization of the transformant culture conditions. Specifically, the synthetic *t4l* is 510 base pair (bp) in size (Figure 2). Moreover, the synthetic gene was cloned into three different vectors along with different promoters. The three vectors bearing-synthetic *t4l* are named as follows: QS, QL, and T4L-pET-20b (+) (Figure 3). QS and QL vectors are 9,000 bp in size; plus, they contained *CbhI* promoter from *Trichoderma reesei* cellobiohydrolases I gene, which facilitates the transcription of synthetic *t4l* in *T. reesei*, to be discussed in the following study. However, the main difference between QS and QL vectors is in the signal peptide size; QS has 15 amino acids (a.a), while QL has 30 a.a. The third vector T4L -pET-20b(+), was 4,096 bp in size. It contains a T7 promoter that only binds T7 RNA polymerase and was designed with the pET expression system. Both promoters allow for manipulation of the quantity of expressed desired proteins and when the expression occurs.

3.2.2. Bacterial Strains and Media

Two strains of *Escherichia coli* (*E. coli*) were used, one for plasmid replication and the other for host expression of synthetic *t4l*, JM109 and BL21(DE3)pLysS, respectively. Also, *M. luteus* was used to test lysozyme activity. All strains were grown in Luria-Bertani (LB) broth (10 g/L peptone, 5 g/L yeast extract, 5 g/L NaCl), LB agar (same content of LB broth plus 12.8 g/L agar) and Super Optimal Broth (SOB) broth (20 g/L peptone, 5 g/L yeast extract, 0.58 g/L NaCl, 0.1875 g/L KCl, 2 g/L MgCl₂·6H₂O, 2.5 g/L MgSO₄·7H₂O, up to 1 L ddH₂O) (Maki *et. al.*, 2011; Hanahan, 1983). Additionally, the following antibiotics were used where appropriate: 100 µl/ml of ampicillin for growth of JM109, 50 µl/ml of ampicillin plus 34 µl/ml of chloramphenicol for growth of BL21(DE3)pLysS; and for selection of transformants. No antibiotics were used with the growth of *M. luteus*.

3.2.3. Halo Assay for Analysis of Host Strains

Due to the antimicrobial properties of lysozyme, it was necessary to determine whether *E. coli* BL21(DE3)pLysS strain would be a suitable host for expression; this was done by testing the strain against commercial lysozyme on agar plate via a method termed halo assay. Although halo assay is not quantitative, it is however, one of the most sensitive methods to detect enzymatic activity and is appropriate in the present context (Kuroki *et. al.*, 1999). The halo assay procedure was adapted from Pjura *et. al.*, (1993) (Pjura *et. al.*, 1993), with minor modifications. Specifically, in this assay a clear circle or halo on agar plate represents lysozyme activity during bacterial lawn growth, which varies from one species to another. Specifically, *M. luteus* and *E. coli* BL21(DE3)pLysS were tested against a commercial lysozyme (1 g of stock powder lysozyme, 50 ml of autoclaved 1X TE buffer). *E. coli* BL21(DE3)pLysS was selected because of its specific sensitivity against lysozyme and is considered the optimal substrate for studying lysozyme effects (Eggert-Kruse *et. al.*, 2000).

Briefly, a fresh single colony from *M. luteus* and *E. coli* BL21(DE3)pLysS LB agar plates was inoculated into a 250ml flask containing 20 ml LB broth individually, incubated at 37°C in shaker with 200 rpm, overnight. Next day, 200 µl of each bacterial culture was spread-plated on fresh LB agar with

appropriate antibiotics if necessary, using a sterile glass spreader. Prior to spread-plates, the optical density at 600 nm (O.D._{600nm}) for *M. luteus* and *E. coli* BL21(DE3)pLysS, were adjusted to 0.9 O.D._{600nm}.

Thereafter, plates were allowed to dry for approximately 30 min at room temperature. Then, on each plate four drops with varying volumes and concentrations of 20 µg/ml commercial lysozyme stock were added onto each quarter of the agar plate to optimize results, as follows: 2 µl, 5 µl of 0X; 5 µl of 10X; 5 µl of 100X. Once again, the plates were dried, then inverted and incubated at 37°C overnight. The following day, agar plates were visualized and photographed for halo analysis.

3.2.4. Electrotransformation of *E. coli* JM109 and BL21(DE3)pLysS

Electrocompetent cells of *E. coli* were prepared according to the method described by Majeed *et. al.*, (2011) (Majeed *et. al.*, 2011). Briefly, fresh single colony of *E. coli* JM109 and BL21(DE3)pLysS were individually inoculated into a flask containing 50 ml of LB broth and incubated at 37°C, shaking at 200 rpm overnight. Next day, 1 L of pre-warmed LB broth was inoculated by the overnight bacterial culture and once again incubated on a rotary shaker at the same conditions. Consequently, growth of bacterial cultures was measured at O.D._{600nm} every 30 min. Then, once the O.D._{600nm} of the cultures reached approximately 0.6 (+/- 0.2), the flasks were immediately transferred to an ice-water bath for 30 min, with occasional swirling. After, cultures were transferred to ice-cold centrifuge bottles and the cells were harvested by centrifugation at 6000 rpm for 30 min, at 4°C. Then, supernatant was discarded and sterile cold deionised and distilled H₂O (ddH₂O) was added to the same volume; followed by centrifugation at 5000 rpm for 10 min, at 4°C. The latter step was repeated twice to wash each strain. Similarly, the supernatant was decanted and the cell pellet resuspended first in 250 ml then in 10 ml of sterilized ice-cold 10% glycerol twice as previously described for sterile cold ddH₂O. Next, the supernatant was carefully decanted; then, the pellet was resuspended in 1 to 2 ml of sterile ice-cold 10% glycerol and aliquoted into 40 µl tubes.

Moreover, the three plasmids; QS, QL, and T4L-pET-20b(+) bearing synthetic *t4l*, were independently electrotransformed into competent cells of *E. coli* JM109 then BL21(DE3)pLysS strains by

electroporation technique according to Dower *et. al.*, (1988) (Dower *et. al.*, 1988). In brief, 40 µl of freshly made electrocompetent cells were pre-chilled, along with 0.2 cm gap electroporation cuvettes on ice. Also, 1 – 2 µl of each plasmid DNA, ranging in concentration from 2 to 7 µg, was pre-mixed with chilled electrocompetent cells and incubated on ice for 30 – 60 seconds. Rapidly, DNA/cell mixture was pipetted into a chilled electroporation cuvettes. Next, a pulse of electricity was delivered to the cells using Gene Pulser Xcell™ microbial system (Bio-Rad Laboratories, Richmond, CA), with 2 mm / 2.5 kV. Immediately thereafter, the electroporation cuvette was removed and 1 ml of SOB medium added at room temperature; then transferred to 1.5 ml microfuge tube and incubated 1 hour at 37°C, shaking at 50 rpm. Following, 200 µl of electroporated cells were spread plated onto the surface of LB agar plates, including appropriate antibiotics (100 µl/ml of ampicillin for JM109, 50 µl/ml of ampicillin plus 34 µl/ml of chloramphenicol for BL21(DE3)pLysS). All plates were incubated at 37°C; transformed colonies appeared in ~ 18 hours of the incubation period and were then streaked out on separate LB plates to ensure the purity (Figure 4).

Furthermore, the plasmid DNA was extracted from transformants using GeneJET™ Plasmid Miniprep Kit (Fermentas), according to the supplier's protocol. The DNA samples were run on 1% agarose gel; gels were then examined under UV illumination to observe extracted plasmid DNA.

3.2.5. PCR Amplification of Synthetic *t4l*

Detection and PCR amplification was performed to confirm the presence of synthetic *t4l* in the extracted plasmid DNA from transformants *E. coli* JM109 and BL21(DE3)pLysS (Snounou *et. al.*, 1993). The following primers were used to amplify a PCR product of 510 b.p. from synthetic *t4l* gene in the vectors: “T4L-Primer 5” (Forward (FW): 5'-TCC CAT ATG AAC ATC TTC GAG AAG TTG AGA-3') and “T4L-Primer 6” (Reverse (REV) 5'-ATA CTC GAG CAA GTT CTT GTA AGC GTC CCA-3'). Oligonucleotides were designed on the basis of the known DNA sequences of the synthetic *t4l* using DNAMAN software version 7.212 (Lynnon Corporation, Canada) and custom-synthesized by Eurofins MWG Operon, Canada. The PCR reaction mixtures was performed using 20 µl in 0.2 ml of PCR tubes

consisting of 6.6 μ l of Master-mix, which included 4 μ l 5X Phusion HF buffer (BioLab), 0.4 μ l 10 mM dNTPs, 1 μ l of each 10 pmol forward and reverse primers, and 0.2 μ l Phusion High-Fidelity DNA polymerase (BioLab), 1.5 μ l of (10 ng) extracted plasmid DNA as template and 5.3 μ l of sterile UV-treated ddH₂O. The PCR reaction was run in MyCycler thermal cycler (BIO-RAD) according to the following protocol: primary denaturation 5 minutes at 95°C, followed by 30 amplification cycles consisting of denaturing at 94°C for 1 minute, annealing for 1 minute at 60°C, and extension at 72°C for 1 minute, upon completion of 30 amplification cycles a final extension step was done at 72°C for 10 minutes. Finally the reaction was held at 4°C. The PCR products were visualized by 1% agarose gel electrophoresis stained with ethidium bromide for visual detection on an UV transillumination. Gels were then examined under UV illumination and photographed to confirm size, quantity and purity.

3.2.6. Expression of T4 Lysozyme from Transformants and SDS-PAGE

To produce recombinant T4L, transformants BL21(DE3)pLysS carrying three plasmids bearing synthetic-*t4l*-gene were induced by the addition of isopropyl-B-D-thiogalactoside (IPTG) according to the method described by Langen *et. al.*, (2000) (Langen *et. al.*, 2000). In the beginning, one fresh single colony of each transformants was inoculated in LB broth (5ml) supplemented by 50 μ g/ml Ampicilin and 34 μ g/ml Chloramphenicol at 35°C with 180 rpm in shaker, for overnight. Then, the cultured bacteria (500 μ l) diluted in 50 ml of LB broth. The O.D₆₀₀ measurements followed until reached 0.5, at that time, 500 μ l of 0.1 M IPTG (final concentration 1.0 mM IPTG) added to the culture. After incubation for 2-3 hours more at 25°C instead of 37°C, cells were collected by centrifugation for 10 min at 3,5000 rpm. Subsequently, 1 ml of the supernatant collected in microcentrifuge tube and kept in ice, while the rest of supernatant was discarded. The pellet resuspended with 0.01 M PBS buffer (pH 7.5) and centrifuge with same condition mentioned above for washing. Adding 500 μ l of same buffer was before sonicated the cells at 70 for 5s and cool down for 2 min intervals, for total of 6 times. Follow that, samples centrifuged for 2 min at 15,000 xg, then the supernatant was collected also. Using Sodium Dodecyl Sulfate

Polyacrylamide Gel Electrophoresis (SDS-PAGE) with these two collected samples set was for detection of the production of recombinant T4L.

SDS-PAGE has been used to separate proteins according to their electrophoretic mobility. Definitely, adjusting the gel concentration based on the molecular weight of protein, in our case, 12% resolving gel was used for denaturing SDS-PAGE, while for stacking gel 5%. Also, using tetramethylethylenediamine (TEMED) with both gels was essential. TEMED was used with ammonium persulfate (APS) to catalyze the polymerization of acrylamide when making polyacrylamide gels. Therefore, after it was added to the solution, gel casted right away. Both collected samples set, before and after sonication, was prepared as follow: 24 μ l of sample mixed with 2 μ l of reducing buffer and 10 μ l of dye; except the sonicated sample has been diluted 1:1 with 0.01 M PBS buffer pH 7.5. After mixing the samples, they were boiled for 5 min at 100°C, before loading 10 μ l in the gel. Next, the samples were loaded in the gel followed by attaching the power supply, and 80 V was used first for 10 – 20 min, after that, the voltage increased up to 110 V, for approximately two hours and half.

After running an SDS-PAGE gel, gel was washed with dH₂O several times gently. Definitely, there are several kinds of staining that have been used for staining the SDS-PAGE gel such as, PageBlue, Coomassie, and Silver staining. Here, coomassie staining was used by putting the gel in container for staining by coomassie blue staining solution and kept in staining solution for 30 min with gentle shaking (60 rpm). Then, stained gel was washed with H₂O and putted in a clean try and coomassie blue destaining solution was added to remove the stain from the gel for overnight. Next day, the gel was washed with H₂O and gel photographed. The expect band for T4L should has a molecular weight of approximately 18 kilodalton (kDa).

3.3. Results and Discussions

3.3.1. Halo Assay

Growth inhibition of *M. luteus* was visualized by the appearance of a clear zone after treatment with varying concentrations of commercial lysozyme (Figure 5.A). In contrast, there was no lysis activity

observed towards *E. coli* BL21(DE3)pLys growth when the same volume and concentration of commercial lysozyme was used (Figure 5.B). Therefore, the results demonstrate that BL21(DE3)pLysS strain is a good potential candidate for the expression of recombinant T4L such as synthetic *t4l*. The exact mechanism of lysozyme resistance is not fully understood and may vary according to the bacterial strain or species; however, various possibilities have been discussed in other studies such as: O-acetylation of the NAM of PG as described for 11 Gram positive and Gram negative species, attachment of other polymers. For example, the attachment of polysaccharides to the cell wall, as described for streptococci, a high degree of peptide cross-linking, or N-non-substituted glucosamine residues in the peptidoglycan, as in *Bacillus* spp. and *Streptococcus pneumoniae*. However, researchers suggested the resistance of BL21(DE3)pLysS is related to *OatA* gene encoding O-acetyltransferase A.

3.3.2. Synthetic *t4l* amplification

Two primers: T4L-Primer 5 and T4L-Primer 6, were used in combination with PCR to successfully amplify the synthetic *t4l* fragment from extracted plasmid DNA of transformants. A confirmation of the amplified synthetic *t4l* fragment was validated by band on a 1% agarose gel with an approximate size of 510 b.p., this was the exact size of the gene of interest: synthetic *t4l*. (Figure 6).

3.3.3. Detection of Recombinant T4 Lysozyme

After PCR confirmation of the three plasmids into *E. coli* BL21(DE3)pLysS, the expression of synthetic *t4l* was examined via the detection of *t4l* in the medium. Thus, supernatant samples of the transformants BL21(DE3)pLysS containing *t4l* were collected and ran on SDS-PAGE gel. The first set of collected samples, without sonication treatment; displayed no significant bands for the presence of recombinant T4L. However, the experimental samples, collected after sonication, displayed very clear bands for all transformants; while, the negative control (native BL21(DE3)pLysS) did not show a band. Based on the ladder, the bands had a molecular weight equivalent to approximately 18 kDa, which is the exact molecular weight of T4L (Figure 7). This was an excellent indication of successful expression of the synthetic *t4l* in *E. coli* system.

Conclusions

The results of this study show *E. coli* BL21(DE3)pLysS may be good potential candidate for the production of T4L as a bioproduct. Moreover, it has been considered for several years, that *M. luteus* is the most lysozyme-action sensitive; thus, it was used in the current study as a test strain. Future work will be done on the transformants to optimize T4L external production and evaluate individual T4L activities to look for novel recombinant T4Ls. As well, the production of this natural antimicrobial agent in the *E. coli* system is important in industry to help overcome costly hurdles in the sterilization process. In this study, all transformants of BL21(DE3)pLysS were found to produce T4Ls and their T4Ls will be further studied. These transformants represent important T4L-producing hosts and may be an integral part of future work to develop efficient T4L producing system which can be used for industry. Finally, the expression of synthetic *t4l* in *E. coli* system, may also contribute to expression of the same gene in other microorganism, which are known for producing heterologous proteins.

Figure legend

Figure:

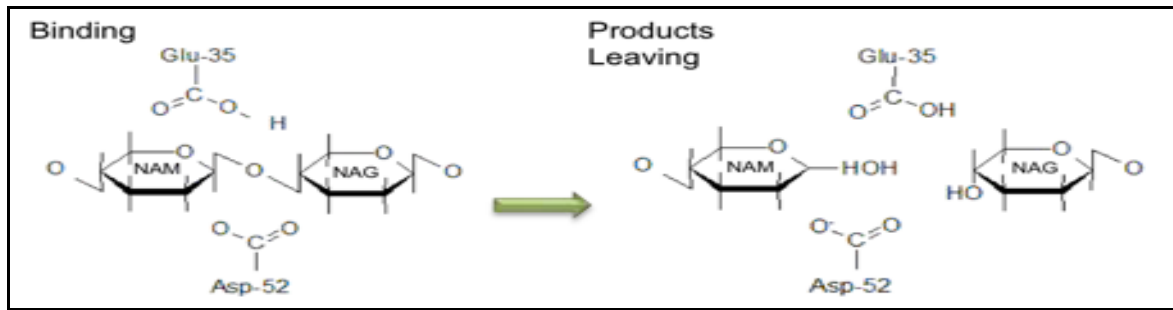


Figure 1: The catalytic mechanism of lysozyme. The substitution on the sugar rings have been omitted to improve the clarity of the diagram. The diagram drew by using ACD/ChemSketch Freeware software, which is a free comprehensive chemical drawing package (<http://www.acdlabs.com>). It was originally adapted from (Smith and Wood, 1991).

```

1      GAATTCATGA ACATCTTCGA GAAGTTGAGA ATCGACGAGA GATTGAGATT GAAGATTTAC
61     AAGGACACTG AGGGTTACTA CACTATCGGT ATCGGTCACT TGTTGACTAA GTCCCCATCC
121    TTGAACGCTG CTAAGTCCGA GTTGGACAAG GCTATCGGTA GAAACTGTAA CGGTGTTATC
181    ACTAAGGACG AGGCTGAGAA GTTGTTC AAC CAAGACGTTG ACGCTGCTGT TAGAGGTATC
241    TTGAGAAACG CTAAGTTGAA GCCAGTTTAC GACTCCTTGG ACGCTGTTAG AAGATGTGCT
301    TTGATCAACA TGGTTTTCCA AATGGGTGAG ACTGGTGTG CTGGTTTAC TAATCCTTG
361    AGAATGTTGC AACAAAAGAG ATGGGACGAG GCTGCTGTTA ACTTGGCTAA GTCCAGATGG
421    TACAACCAAA CTCCAACAG AGCTAAGAGA GTTATCACTA CTTTCAGAAC TGGTACTTGG
481    GACGCTTACA AGAAGTTGTA AGCGGCCGCA

```

Figure 2: Synthetic *t4l* gene sequence in FASTA format

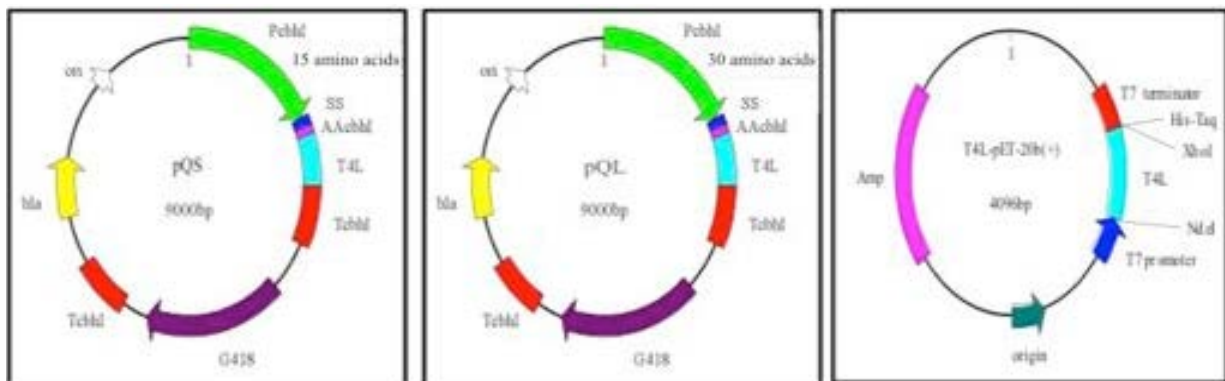


Figure 3: Plasmids bearing synthetic *t4l* gene; QS with 9000 b.p. and signal sequence peptide size (SS) 15 a.a., QL with 9000 b.p. and SS 30 a.as., and T4L-pET-20b (+) with 4096 b.p.

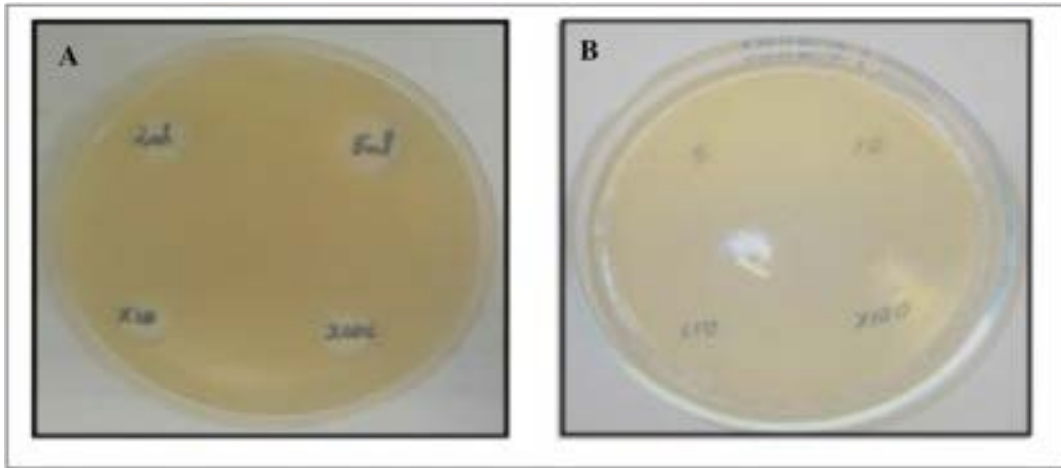


Figure 5: Halo Assay, *M. luteus* (A), BL21(DE3)pLysS (B),

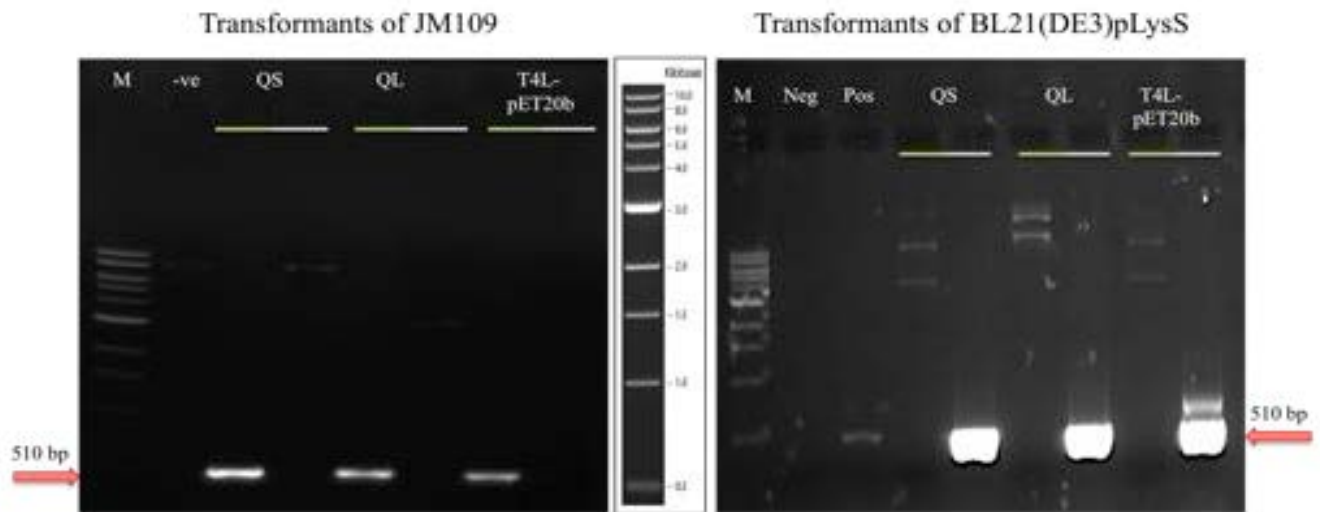


Figure 6: PCR products of transformants JM109 (on the left) and BL21(DE3)pLysS (on the right)

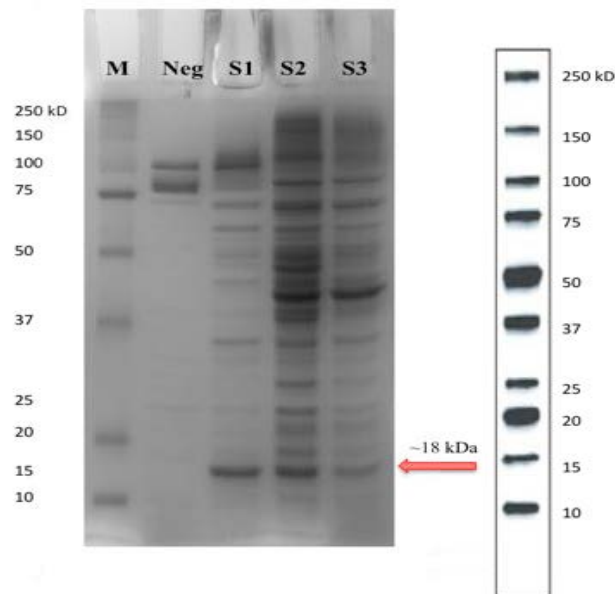


Figure 7: SDS-PAGE M; Lader 250 kDa, Neg; BL21(DE3)pLysS, S1; BST4, S2; BLT4, S3; BPT4

References

- Akinalp, A.S., M. Asan and N. Ozcan. (2007). Expression of T4 lysozyme gene (gene e) in *Streptococcus salivarius* subsp. *thermophilus*. *Afr. J. Biotechnol.*, 6: 963-966.
- Bower, C. K., Xu, Q., & McGuire, J. (1998). Activity losses among T4 lysozyme variants after adsorption to colloidal silica. *Biotechnol Bioeng*, 58(6), 658-662.
- Dong, S., Shew, H. D., Tredway, L. P., Lu, J., Sivamani, E., Miller, E. S., & Qu, R. (2008). Expression of the bacteriophage T4 lysozyme gene in tall fescue confers resistance to gray leaf spot and brown patch diseases. *Transgenic Res*, 17(1), 47-57. doi: 10.1007/s11248-007-9073-3
- Dower, W. J., Miller, J. F., & Ragsdale, C. W. (1988). High efficiency transformation of *E. coli* by high voltage electroporation. *Nucleic Acids Res*, 16(13), 6127-6145.
- During, K., Porsch, P., Mahn, A., Brinkmann, O., & Gieffers, W. (1999). The non-enzymatic microbicidal activity of lysozymes. *FEBS Lett*, 449(2-3), 93-100.
- Eggert-Kruse, W., Botz, I., Pohl, S., Rohr, G., & Strowitzki, T. (2000). Antimicrobial activity of human cervical mucus. *Hum Reprod*, 15(4), 778-784.
- Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol*, 166(4), 557-580.
- Henrissat B., A classification of glycosyl hydrolases based on amino-acid sequence similarities. *Biochem. J.* 280:309-316(1991).
- KIM, J.W. and YOEE, S.M. (2008). Cloning and prokaryotic expression of C-type lysozyme gene from *Agrius convolvuli*. *Animal Cells and Systems*, vol. 12, no. 3, p. 149-155.
- Kuroki, R., Weaver, L. H., & Matthews, B. W. (1999). Structural basis of the conversion of T4 lysozyme into a transglycosidase by reengineering the active site. *Proc Natl Acad Sci U S A*, 96(16), 8949-8954.
- Langen, R., Oh, K. J., Cascio, D., & Hubbell, W. L. (2000). Crystal structures of spin labeled T4 lysozyme mutants: implications for the interpretation of EPR spectra in terms of structure. *Biochemistry*, 39(29), 8396-8405.
- Leboffe, M. J., & Burton E. Pierce, 2005, A Photographic Atlas for the Microbiology Laboratory, 3rd Edition, Morton Publishing Company
- Leiman, P. G., Kanamaru, S., Mesyanzhinov, V. V., Arisaka, F., & Rossmann, M. G. (2003). Structure and morphogenesis of bacteriophage T4. *Cell Mol Life Sci*, 60(11), 2356-2370. doi: 10.1007/s00018-003-3072-1
- Majeed M. T., Hussain K., Noreen M., and Izhar M. (2011). Optimum range of plasmid supercoiled DNA for preparation of competent Top 10 *E. coli*. Research Article, Archives of Pharmacy Practice, Vol. 2, Issue 2, 2011, ISSN 2045-080X
- Maki, M. L., Broere, M., Leung, K. T., & Qin, W. (2011). Characterization of some efficient cellulase producing bacteria isolated from paper mill sludges and organic fertilizers. *Int J Biochem Mol Biol*, 2(2), 146-154.
- Miller, E. S., Kutter, E., Mosig, G., Arisaka, F., Kunisawa, T., & Ruger, W. (2003). Bacteriophage T4 genome. *Microbiol Mol Biol Rev*, 67(1), 86-156, table of contents.
- Murray, B. E., Singh, K. V., Heath, J. D., Sharma, B. R., & Weinstock, G. M. (1990). Comparison of genomic DNAs of different enterococcal isolates using restriction endonucleases with infrequent recognition sites. *J Clin Microbiol*, 28(9), 2059-2063.

- Pjura, P., Matsumura, M., Baase, W. A., & Matthews, B. W. (1993). Development of an in vivo method to identify mutants of phage T4 lysozyme of enhanced thermostability. *Protein Sci*, 2(12), 2217-2225. doi: 10.1002/pro.5560021221
- Snounou, G., Viriyakosol, S., Zhu, X.P., Jarra, W., Pinheiro, L., do Rosario, VE., Thaithong, S., Brown, K.N. (1993). High sensitivity of detection of human malaria parasites by the use of nested polymerase chain reaction. *Mol Biochem Parasitol.*, 61(2): 315-320.
- Sulakvelidze, A., Alavidze, Z., & Morris, J. G., Jr. (2001). Bacteriophage therapy. *Antimicrob Agents Chemother*, 45(3), 649-659. doi: 10.1128/AAC.45.3.649-659.2001
- Towbin, H., Staehelin, T., & Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc Natl Acad Sci U S A*, 76(9), 4350-4354.
- Trun, N., & Trempey, J. (2004). *Fundamental Bacterial Genetics*. UK: Blackwell Publishing.
- van de Guchte, M., van der Wal, F. J., Kok, J., & Venema, G. (1992). Lysozyme expression in *Lactococcus lactis*. *Appl Microbiol Biotechnol*, 37(2), 216-224.

Chapter IV: Expression of T4 Lysozyme gene in *Trichoderma reesei*

Abstract

In this study, *Trichoderma reesei* QM6a was developed as a host to produce recombinant T4L using protoplast-based transformation method. The biological output of T4L in *T. reesei* strain was done using synthetic *t4l*, which was previously studied through bioinformatics and expression in *E. coli* cells. However, here the exogenous gene of T4L was modified for control in *T. reesei* expression, using *T. reesei* cellobiohydrolase I (*CbhI*) gene promoter for induction mode. Thus, specific culture and growth conditions could improve the biological output of *t4l* and its extraction.

The operon of *t4l* was cloned into three different plasmids and was successfully integrated in *T. reesei* QM6a strain, to form recombinant strains, termed: ST4L1, ST4L2, and ST4L3. Moreover, the synthetic *t4l* was successfully expressed in two of our recombinant strains, as confirmed by SDS-PAGE. Therefore, I developed a novel system by which a foreign gene can be successfully inserted into the chromosome of *T. reesei*, and where transcription can be driven by *cbhI* promoter, with the addition of the inducer carbon source. T4L production by this system was stable and the system may be useful in biotechnological fermentation processes other recombinant proteins.

Key words: *T. reesei*, T4 Lysozyme, PCR, Sequencing, SDS-PAGE

4.1. Introduction

Fungi are nonmotile eukaryotes and their cell wall is usually made of the polysaccharide chitin. Unlike animals that ingest then digest their food, fungi are absorptive heterotrophs. That is, they secrete extracellular enzymes into the environment, then absorb the digested nutrients. Additionally, fungi are informally divided into unicellular yeast and filamentous molds based on their reproductive structures; the majority of fungi are filamentous; one of them *Trichoderma reesei* (*T. reesei*) species, can be found hidden underground or within decaying matter (Leboffe & Pierce, 2005). *T. reesei*, also known as *Hypocrea jecorina*, is widely used in industry for cellulase production (Lorito *et. al.*, 2010). Originally, *T. reesei* QM6a strain was isolated from the Solomon islands in World War II because of its degradation of

canvases and garments of the US army (Seidl *et. al.*, 2009). More specifically, the *T. reesei* haploid genome is estimated at 33 Mb, organized in 7 chromosomes, comprised of 9,120 genes targeting 9,115 proteins (Chambergo FS *et. al.*, 2002; Martinez D *et. al.*, 2008). This fungus represents an ideal model system to study the regulation of plant cell wall degrading enzymes, which play an important role in the conversion of cellulosic biomass into glucose, which is then fermented to bioethanol by yeast (Seidl *et. al.*, 2009). *T. reesei* produces a variety of extracellular enzymes, like cellulases and hemicellulases that hydrolyse plant-derived polysaccharides into monomeric sugars, which in turn are used as a source of carbon and energy by the fungal cells. Also, the cellulolytic system of the fungus secretes several enzymes along with cellobiohydrolase I (*CbhI*) gene, which comprise the major part of the total extracellular protein produced by the fungus (Pakula *et. al.*, 2005).

For almost a century, *Trichoderma spp.*, have been widely applied as biocontrol fungi due to their beneficial effects on plants and soil. Thus, they have been extensively studied using a variety of research approaches, including genomics, transcriptomics, proteomics, and metabolomics, etc.. A successful case of translational research, in which omics-generated novel understanding is directly translated into new or improved crop treatments and management methods (Lorito *et. al.*, 2010).

Moreover, with the systematic improvement of production, it was reported *T. reesei* could produce over 40 g/liter of extracellular protein. Also, due to its excellent secretion capacity and cheap easy cultivation, *T. reesei* is also a potential host for the large-scale production of heterologous proteins (Chen *et. al.*, 2010; Nakari-Setala & Penttila, 1995). Particularly, filamentous fungi, as cellulolytic organisms, are of great interest to industrialists because they have the capacity to produce exceptionally large amounts of extracellular enzymes (Ferreira *et. al.*, 2011). For example, *T. reesei* is primarily used for the production of its native cellulolytic and hemicellulolytic enzymes. However, during the last two decades, *T. reesei* has been used as a host to express numerous recombinant proteins. Early examples of this work include two mammalian proteins: bovine chymosin and antibody Fab fragments. The Fab production was greatly improved when the heavy chain gene was fused with the native cellobiohydrolase *CbhI* serving as a

secretion carrier to the Fab. Some studies suggest that there are spatial restrictions in secretion of foreign proteins whereas native proteins with native *CbhI* can greatly enhance its secretion out of the cells (Saloheimo & Pakula, 2012).

Furthermore, advancement in biotechnology especially in the area of genetics, protein engineering, developments in bioinformatics, and the availability of sequence data have opened a new era of enzyme applications in many industrial processes (Dahiya *et. al.*, 2006). In addition, genetic transformation is an essential part of modern fungal research requiring the production of protoplasts (i.e. cells are stripped of their walls by enzymatic digestion) (Turgeon *et. al.*, 2010). However, the recovery of protoplasts might be problematic, especially in slow growing species and the results may vary considerably between different enzyme batches resulting in inconsistent transformation efficiencies (Minz and Sharon, 2010). Strategies, such as the use of specific lytic enzymes, to remove the cell wall along with optimum timing can improve protoplast production; the choice of enzyme for digesting the cell walls being a key factor. Also, protoplasts are osmotically stabilized with sodium chloride, magnesium sulphate, mannitol, and sorbitol or sucrose. On the other hand, mutant strains transformed less efficiently for reasons that have not been properly investigated or understood (Ferreira *et. al.*, 2011).

In the last century, with the development of biotechnologies, the importance has extended beyond agriculture, into enzyme production, food industry, paper and pulp treatment, bioremediation, etc.; *T. reesei* continues being studied extensively. Almost 80 years ago, it was discovered that *Trichoderma spp.* have the ability to attack and control plant pathogenic fungi. Thus, a complex system for fungal prey detection was developed through the study of antagonistic mechanisms in *Trichoderma*, which demonstrated the involvement of many hydrolytic enzymes, also capable of acting synergistically with highly fungitoxic antibiotics. More so, *Trichoderma* enzyme-encoding genes have been used in studies to improve plant resistance to pathogens and salt stress, yet none of the resulting transgenic cultivars have been commercialized (Lorito *et. al.*, 2010). Structural and functional genomic investigations are making

an important impact on the current understanding and application of microbial agents used for plant disease control.

In this study, we present one of the latest discoveries on the *Trichoderma* expressome which may have demands in pharmaceutical and industrial applications; the production of economically competitive recombinant T4L from genetically engineered *T. reesei*. Specifically, I aimed to improve *T. reesei* strain genetically, for the production of a novel strain hosting a synthetic *t4l*. Successful evaluation of recombinant T4L production through this genetically engineered system is challengeable

4.2. Materials and Methods

4.2.1. Strains, Media, and Cultivation Conditions

In this study, two *T. reesei* strains were used; QM6a and its mutant QM9414, for the possibility of expressing a synthetic *t4l* after transformation. Fungal strains, were grown on the surface of two kinds of agar plates; Potatoes Dextrose Agar (PDA) (15.0 g/L potato dextrose broth, 20.0 g/L D-glucose, and 18.0 g/L agar) and/or Yeast extract Glucose (YG) agar (5.0 g/L yeast extract, 20.0 g/L D-glucose, and 14.0 g/L agar) (Benko *et. al.*, 2007); with no antibiotic supplementation. The former were used for spore-forming growth, while the latter for mycelium-forming growth. Plates were incubated at 30°C for 24 h and were then kept upside-down for approximately 10 days at same temperature. Of note, plasmids bearing-synthetic-*t4l*, which were QS, QL, and T4L-pET20b(+), contain hygromycin B phosphotransferase (*hph*) resistance gene; therefore, 25 µl/ml hygromycin B (BioShop, Canada) was used as the selection marker to maintain selection of positive clones.

4.2.2. Antifungal Plate Assay

The commercial lysozyme was examined towards hyphal growth of *T. reesei* on agar plates, as described previously (Mizuno *et. al.*, 2008). *T. reesei* QM6a and QM9414 were used as test fungi. Antifungal plate assay is basic test to determine whether lysozyme has hydrolysis activity on agar plate growth of *T. reesei*. Therefore, this assay was used to determine if any difficulties will occur during expression of a synthetic *t4l* in these organisms. Briefly, seven sterile paper discs were placed on fresh

PDA plates away from the center of the plate by 15 mm each. Next, grown spores were collected from fresh agar plates by adding 3 – 5 ml of 0.9% NaCl on the top of agar with use of a glass spreader to collect most of the grown spores. Then, an autoclaved funnel containing 12 lens papers (Fisher Scientific, Canada) were used to separate spores in the filtrate, from mycelium. After, the spore concentration was adjusted to $\sim 1 \times 10^6$ spores/ml, and 200 μ l of spores were added to the paper disc in the center of PDA agar plate. Subsequently, the plates were incubated for 24 h at 30°C; different volumes of commercial lysozyme (Sigma-Aldrich, Canada) with final concentration 20 μ g/ml was applied to the paper discs; as follows: 1 μ l, 5 μ l, 10 μ l, 15 μ l, 20 μ l, and 30 μ l, while the negative control included 10 μ l of autoclaved ddH₂O. Finally, agar plates were re-incubated at 30°C for a final 24 h; then placed upside down. Visualization of hydrolysis of *T. reesei* growth was followed until growth of spores reached the edge of the plate, and then pictures were taken.

4.2.3. Protoplasts Preparation and Genetic Transformation

T. reesei QM6a protoplasts were prepared using a new method developed through the modification of Szewczyk *et. al.*, (2006) protocol (Szewczyk *et. al.*, 2006). To begin, spores were collected from *T. reesei* QM6a, as previously described using 3 – 5 ml of 0.9 % NaCl and a glass spreader; then, 12 lens papers were used for separation the spores from the mycelium. Following, one drop of the filtrate was used to count spores with a Petroff-Hausser cell counter (American Optical, USA) under a microscope and spore concentration was adjusted to $\sim 1 \times 10^8$ spores/ml. Thereafter, 1 ml of the spores was added to 100 ml YG broth with 0.6 M KCl, including 40 μ l of trace elements. Glass flasks were then incubated at 30°C, shaking at 200rpm for ~ 16 h; the grown hyphae was harvested by 12 filter papers, then mycelium collected and inoculated in 40 ml YG broth with 16 μ l of trace element, plus, 40 ml of 2X protoplast solution (5.12g Novozyme mixed with 40 ml KCl-citric acid, filtered through 6 filter papers, using a syringe). This mixture was incubated at 30°C, shaking at 75 rpm, for 1.5 - 2 h. After the incubation period, mycelium was removed using 8 filter papers with the funnel standing against the flask wall to obtain good protoplast yield. Next, the mixture containing protoplasts was centrifuged for 10 min at 1,800 g and

supernatant discard, then, 15 ml of 0.6 M KCl was added, followed by centrifugation for 10 min at 1,800 g and the supernatant discarded once again. The pellet was then resuspended in 2 ml of 0.6 M KCl and transferred to two sterile 1.5 ml microcentrifuge tubes in 1 ml aliquots. Again, pellet the protoplasts were centrifuged at 2,400 g for 3 min, then supernatant discarded and the pellet resuspend in 1 ml of 0.6 M KCl; this step was repeated twice. Thereafter, the pellet was resuspended in 1 ml (0.5 ml in each tubes) 0.6 M KCl, 50 mM CaCl₂, and centrifuged at 2,400 g for 3 min.

For transformation, 10 - 15 µl of QS, QL, and T4L-pET20b(+) plasmids were independently mixed with 100 µl of protoplasts by vortexing 6 - 8 times for 1 s at maximum speed. Subsequently, 50 µl of PEG solution was added, then vortexed 4 - 5 times for 1 s at maximum speed; all tubes were chilled on ice for 15 min before 0.5 - 1 ml of PEG solution was also added. Next, tubes were incubated at room temperature for 25 min, and 100 ml of pre-warmed YG-sucrose agar, containing 40 µl of trace element and 25 µg/ml of hygromycin, was added and poured into small petri dishes. Finally, all plates were incubated at 30°C, and monitored for 12 – 24 h, and then transformants were screened on PDA plates containing 25 µg/mL hygromycin B as the selection marker.

4.2.4. Confirmation of Positive Clones

4.2.4.1. Genomic DNA Isolation and *st4l* Amplification

Firstly, *T. reesei* QM6a transformants were grown after single spore-separation. Then, the genomic DNA was extracted using the UltraClean Microbial DNA Isolation kit (MO BIO Laboratories Inc., Carlsbad, CA) following the provided protocol by supplier. Following, the genomic DNA extracts were used as a template in PCR to amplify the approximate 510 bp sequence of synthetic t4l using the following primers designed in conserved region: T4L-Primer 5 Forward (FW 5`-TCC CAT ATG AAC ATC TTC GAG AAG TTG AGA-3`) and T4L-Primer 6 Reverse (REV 5`-ATA CTC GAG CAA GTT CTT GTA AGC GTC CCA-3`). The PCR reaction contained 10 ng of genomic DNA individually from each extracted genomic DNA, 10 pmol of forward and reserve primers, 10x Taq buffer with KCL 25 mmol l⁻¹ MgCl₂, 0.2 mmol deoxynuclotide triphosphate, and 5 U Taq DNA polymerase per 50 µl reaction.

The PCR program was as follows: primary denaturation 5 minutes at 95°C, followed by 30 amplification cycles consisting of denaturing at 94°C for 1 minute, annealing for 1 minute at 60°C, and extension at 72°C for 1 minute, upon completion of 30 amplification cycles at final extension step was done at 72°C for 10 minutes and finally the reaction was held at 4°C. Ultimately, the PCR products were then viewed on a 1% agarose gel to confirm size and photograph under UV illumination.

4.2.4.2. Transformants Identification and Sequencing

Previously extracted genomic DNA of transformants was sequenced using standard run modules on the ABI 3730xl automatic sequencer (Eurofins MWG Operon, Canada). Sequencing results were individually inputted online into the nucleotide BLAST tool through the NCBI database (<http://blast.ncbi.nlm.nih.gov/>) to identify the homology of synthetic *t4l* sequence to the general nucleotide database. The percentage of inquiry coverage was recorded.

4.2.5. Detection of Recombinant T4 Lysozyme

Three transformants were selected in T4L production experiments, along with two different strains of *T. reesei* as negative controls, QM6a and its mutant QM9414. Firstly, all strains were grown and maintained on PDA agar plates and after 10 days incubation at 30°C, spores were suspended in ~ 5 ml 0.9% NaCl. Spores were then separated from the mycelium as previously described and counted using a Petroff-Hausser cell counter (American Optical, USA). Next, isolated spores were added to a final concentration of $\sim 1 \times 10^6$ spores/ml to 250 ml flasks containing 50 ml Mandel and Andreotti (MA) medium, as described by Mandel and Andreotti (1978) (Mandel & Andreotti, 1978) with 1% (w/v) glycerol as a sole carbon source (the media was autoclaved at 121°C for 15 min); cultures were then incubated on a rotary shaker (200 rpm) at 30°C. Pre-grown mycelia were washed thrice with MA medium containing no carbon source to remove any residual glycerol. Then, mycelia were transferred into 250 ml flasks containing 50 ml MA medium with 1% glycerol (w/v). Subsequently, to determine recombinant T4L production, a time course trial was conducted similar to that of Cianchetta *et. al.*, (2010) (Cianchetta *et. al.*, 2010). Specifically, the flasks were incubated at 30°C, shaking at 200 rpm for a total of 60 hours.

Finally, samples of 500 μ l were taken from each flask every 24 h in a biosafety cabinet to maintain sterility. These 500 μ l samples were centrifuged at 13000 rpm for 5 min, and the supernatant was used as the source of enzyme (Benko *et. al.*, 2007).

SDS-PAGE has been used to separate proteins according to their electrophoretic mobility. Definitely, adjusting the gel concentration based on the molecular weight of protein, in our case, 12% resolving gel was used for denaturing SDS-PAGE, while for stacking gel 5%. Also, using tetramethylethylenediamine (TEMED) with both gels was essential. TEMED was used with ammonium persulfate (APS) to catalyze the polymerization of acrylamide when making polyacrylamide gels. Therefore, after it was added to the solution, gel casted right away. All collected samples set were prepared as follow: 24 μ l of sample mixed with 2 μ l of reducing buffer and 10 μ l of dye. After mixing the samples, they were boiled for 5 min at 100°C, before loading 10 μ l in the gel. Next, the samples were loaded in the gel followed by attaching the power supply, and 80 V was used first for 10 – 20 min, after that, the voltage increased up to 110 V, for approximately two hours.

After running an SDS-PAGE gel, gel was washed with dH₂O several times gently. Definitely, there are several kinds of staining that have been used for staining the SDS-PAGE gel such as, PageBlue, Coomassie, and Silver staining. Here, coomassie staining was used by putting the gel in container for staining by coomassie blue staining solution and kept in staining solution for 30 min with gentle shaking (60 rpm). Then, stained gel was washed with H₂O and putted in a clean tray and coomassie blue destaining solution was added to remove the stain from the gel for overnight. Next day, the gel was washed with H₂O and photographed. The expect band for T4L should has a molecular weight of approximately 18 kilodalton (kDa).

4.3. Results and Discussions

4.3.1. Strain optimization and Halo Assay

The concentration of hygromycin B was critical; low concentration resulted in high background causing an inability to distinguish transformants from the agar plate, whereas at high concentrations there

are no colonies. Therefore, 25 µg/ml hygromycin B was the optimal concentration which would allow the development of low background after 24 – 48 h, and has been used as a selection marker for transformation of *T. reesei*.

Additionally, halo assay results showed that commercial lysozyme has the ability to inhibit the hyphal-growth on agar plate assay (Figure 1). In other words, lysozyme was considered as an antifungal agent, based on the results. The inhibition of hyphal extension has been explained by the presence of chitin as an important constituent of fungal cell wall. Chitin, a β -(1,4)-linked polymer of two *NAG*, is widely distributed in nature, particularly as a structural polysaccharide in fungal cell walls (Dahiya *et. al.*, 2006; Okazaki *et. al.*, 2004). Several reports are available regarding the presence of chitin in fungal cell walls. In filamentous fungi, it comprises 16% of the dry weight of the organism. In addition, chitin has the same β (1-4) glucosidic bonds as the bacterial PG, except that the bond links two *NAG* residues instead of an *NAG* and *NAM* (Dahiya *et. al.*, 2006). On the other hand, chitinases, which hydrolyze chitin, occur in a wide range of organisms including viruses, bacteria, fungi, insects, higher plants, and animals (Okazaki *et. al.*, 2004, Watanabe *et. al.*, 1999). The roles of chitinases in these organisms are diverse. For example, chitinases have been implicated in plant resistance against fungal pathogens because of their inducible nature and antifungal activities in vitro. Additionally, chitinase in fungi is thought to have autolytic, nutritional, and morphogenetic roles. Moreover, chitinases can also be used in agriculture to control plant pathogens, as shown in immense potential for increasing the production of several useful products economically (Dahiya *et. al.*, 2006). Therefore, lysozyme could be considered as a good chitinase based on our results; because, lysozyme had the ability to lyse chitin of fungal cell wall as was shown in the halo assays (Figure 2). Also, lysozyme possesses chitinase activity, however there is no obvious amino acid sequence similarity found between these two types of enzymes.

4.3.2. Production of Protoplast and Genetic Transformation

The procedure for producing transformable protoplasts from grown hyphae was described here in detail. This procedure used Novozyme, as a selected enzyme, and could be easily scaled to produce larger

quantities of protoplasts. Figure 3 shows transformable protoplasts produced from *T. reesei* QM6a. Similarly, a transformation procedure was described here and transformations were successful with initial observations of growth on agar plates. Consequently, wild-type *T. reesei* QM6a did not grow well on agar plates when 25 µg/ml hygromycin was used; therefore, the observational growth of transformants was a strong signal for positive colonies. Most of the positive colonies appeared after 2 days and were transferred to fresh PDA only when a clear colony was produced, which usually takes 3 days. Furthermore, additional colonies appeared during several days, these were also stored. Three mitotic stable hygromycin-resistant transformants were selected by single spore isolation to insure pure cultures. The three transformants (ST4L1, ST4L2, and ST4L3) were cultured for more than 8 generations.

4.3.3. Confirmation of Positive Clones

4.3.3.1. Characterization of T4 Lysozyme Transformants

The integration of synthetic *t4l* gene into the genome of *T. reesei* QM6a was confirmed by PCR and sequencing using primers designed to target full size of synthetic *t4l*, with an expected size of 510 bp (Figure 4). However, the migration of PCR products' bands on agarose gel determined that both transformants and wild-type had similar bands sizes on the gel. In another words, there were bands with same molecular weight and size of synthetic *t4l*, from both transformants and QM6a as well QM9414; when the above mentioned primers were used. Therefore, new primers were designed; before that, both primers were aligned to the genomic DNA sequence of *T. reesei* QM6a, which was available onto NCBI under accession number AAIL00000000, to identify region of the *T. reesei* genome had similarity with the existing primers. This was a defined gene: bifunctional dethiobiotin synthetase/adenosylmethionine-8-amino-7-oxonanoate aminotransferase (*bioDA*), which is available in NCBI under accession number BK007868, and its size is 2736 b.p.. Then, the nucleotide sequence of *bioDA* gene was recorded and aligned to synthetic *t4l* using MultAlin software (multiple sequence alignment by Florence Corpet) (Corpet, 1988). After analysis, I explored the possibility of designing new primers to target a small of portion of synthetic *t4l* only; there were however, high conserved residues between both nucleotide

sequences as presented in Figure 5. Thus, a newly designed set of primers would be unspecific to synthetic *t4l*.

4.3.3.2. Positive clone Identification and Relevancy

Sequencing results were obtained for all positive clones, and sequencing analysis was performed twice for each sample (i.e. extracted genomic DNA of transformant); one sequencing result with forward primer and the other result with reverse primer. Each resulting sequence was then inputted to the nucleotide BLAST feature of the NCBI database to obtain possible identities based on homology to database, synthetic *t4l*, genomic DNA of *T. reesei* QM6a, and *bioDA* gene. The nucleotide BLAST results are shown in Table 1. From BLAST search results, confirmation of ST4L1 and ST4L3 was determined based on 97-98% homology compared to other synthetic T4L genes in the database, with the exception of low (69% homology) for ST4L2 to the same gene mentioned above. However, the similarity of sequencing results to the nucleotide sequence of synthetic *t4l* yielded 98-100% homology with ST4L1, 34-70% with ST4L2, and 97% with ST4L3. Additionally, all positive clones displayed no significant similarity to both *bioDA* and genomic DNA of *T. reesei*, while ST4L2 was found to have low percentage of similarity (3% homology) to *bioDA*. In conclusion, these results showed the insertion of synthetic *t4l* into the genomic DNA of *T. reesei* QM6a was successful with the three plasmids bearing synthetic-*t4l*.

4.3.4. Detection of Recombinant T4 Lysozyme Production

After confirmation of the successful integration of synthetic *t4l* gene into the genome of *T. reesei*, via both PCR targets and sequencing, I explored how synthetic *t4l* could be expressed in the media. Supernatant samples were collected from transformants and control every 12 h of incubation for a total 72 h and ran on SDS-PAGE gel. However, samples collected with 12 h, 24 h, 36 h, and 72 h incubation did not display significant bands; thus, indicating a lack in presence of recombinant T4L (gel image not provided). Nonetheless, the set of samples collected after 48 h incubation, had very clear bands for two transformants, ST4L2 and ST4L3, which was not shown in the negative control (native *T. reesei* QM6a). In addition, there was clear induced band present in ST4L3 sample collected after 60 h incubation. These

bands had molecular weight of approximately 15 kDa, which is not the exact molecular weight of the native T4L (approximately 18 kDa) (Figure 6). The slight molecular weight reduction of the recombinant protein could be due to RNA processing, including pre-mRNA and mRNA, which are followed by mRNA translation to protein in the fungus; as it is eukaryotic organism. In other words, a gene is first transcribed into a pre-mRNA, which is a copy of the genomic DNA, containing intronic regions to be removed during pre-mRNA processing (called RNA splicing) as well as exonic sequences that are retained within the mature mRNA. In conclusion, the results of SDS-PAGE were an excellent indication of successful expression of the synthetic *t4l* in *T. reesei* system. The SDS-PAGE indicated that it is very likely the T4L protein was produced, but a western blot verification is necessary.

Conclusions

The development of genetic transformation systems has been an essential factor in the advancement of recombinant DNA techniques. However, the development of genetic transformation systems for species in which none exists can present significant challenges; thus, it is pertinent to collect methods which have been successfully applied to other species and determine through troubleshooting how best to apply them to the target organism. Transformation is an essential part of modern fungal research. Moreover, the number of filamentous fungi which have been genetically transformed has increased, thereby opening the door to more detailed molecular projects. This study represents one of the cutting-edge fungal biotechnology works illustrating the development of protoplast-based transformation of *T. reesei*, for the purpose of expression and production of antimicrobial enzyme; T4L enzyme. The use of genetically improved *T. reesei* strain is allows limited microbiological contamination during an industrial process. Especially, *T. reesei* is of great interest to industrialists because they have the capacity to produce extracellular enzymes in very large amounts.

According to this study, the method of protoplast-based transformation presented here was efficient, allowing for the production of heterologous proteins such as here recombinant T4L, as an example. Moreover, I was able to extract a heterologous protein from the media easily using *Cbhl* promoter.

According to this study, fungal protoplasts have been used as an effective experimental tool in strain improvement for biotechnological applications. We have successfully modified the genomic DNA of three transformants: ST4L1, ST4L2, and ST4L3, with potential for industrial use in limiting the contamination through the production of this antimicrobial agent. These transformants may lay the foundation for the current exploitation of this enzyme by further investigation. Also, these transformants may have great potential for developing fungal biotechnology in the near future to enhance the sterilization and help overcome costly hurdles being faced in the industrial production of antimicrobial agents.

Figure and table legend

Figure:

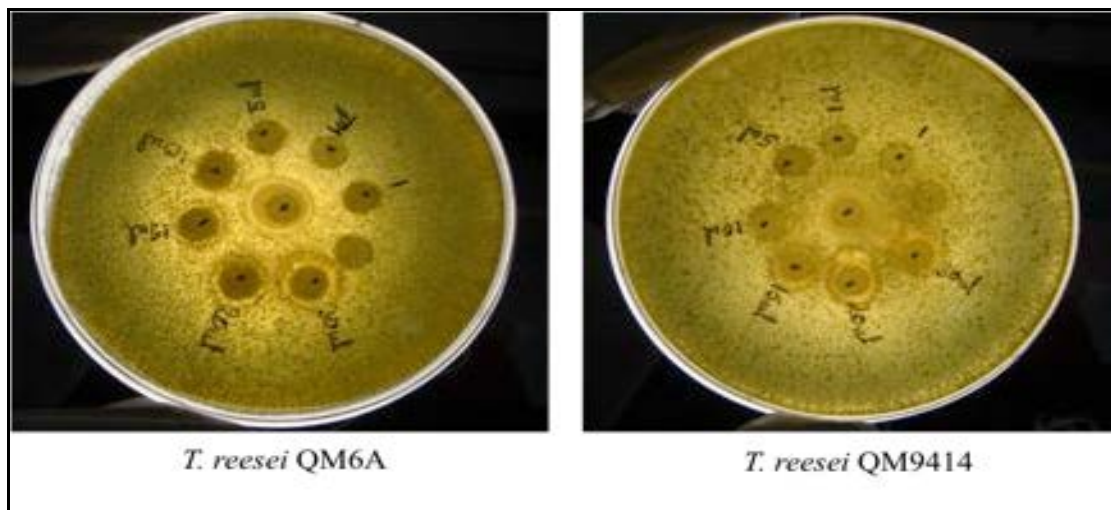


Figure 1: Halo assay – antifungal activity: a suspension of fungal spores was placed on the central disc on a PDA plate and different volumes of commercial lysozyme added to perimeter discs. After incubation period hyphal extension-inhibition generated a clear zone by lysozyme action.

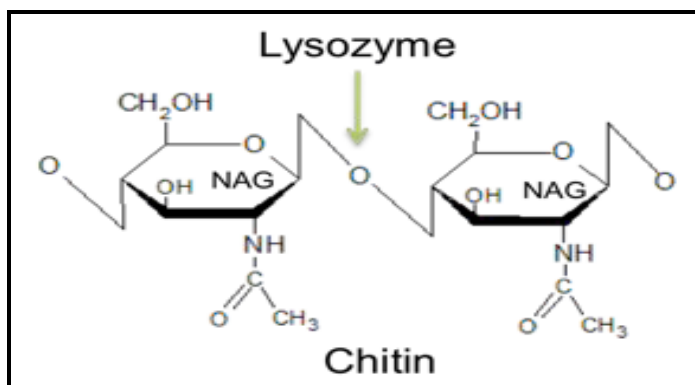


Figure 2: Structure of fungal cell wall (chitin), the two NAG residues linked by $\beta(1-4)$ glucosidic bonds as the bacterial PG.

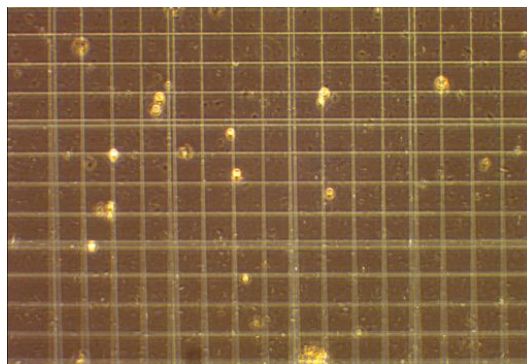


Figure 3: Transformable protoplasts produced from *T. reesei* QM6a strain under the microscope.

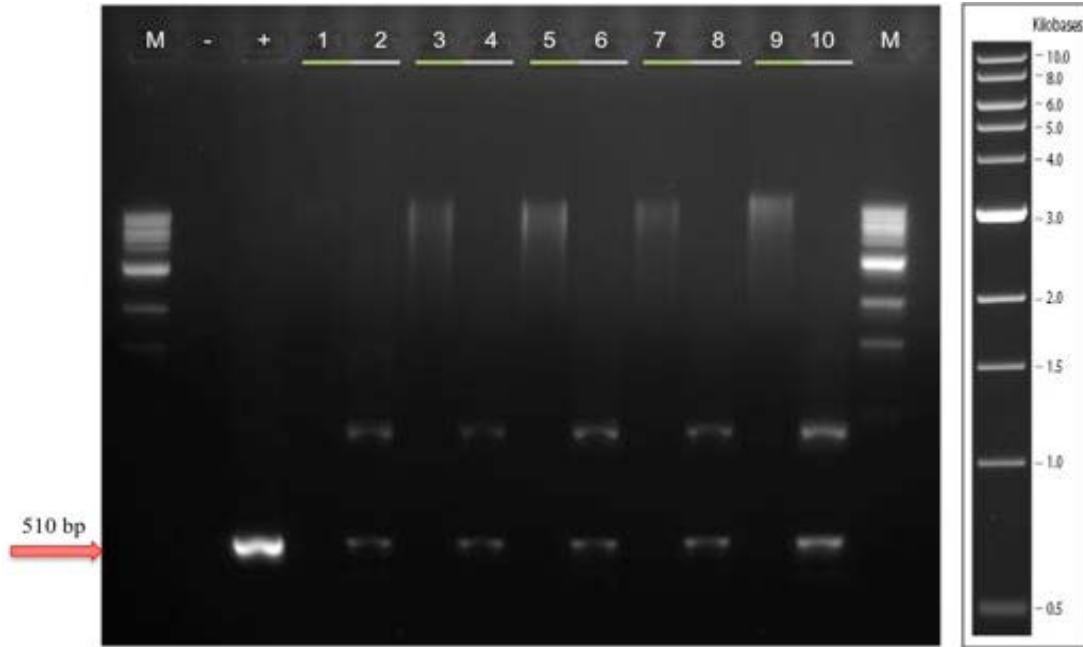


Figure 4: PCR Products on agarose gel electrophoresis, **M**; 1 kb, **Neg**; **POS**; **1**; QM6a gDNA, **2**; QM6a gDNA + PCR Products, **3**; ST4L1 gDNA, **4**; ST4L1 gDNA PCR Products, **5**; ST4L2 gDNA, **6**; ST4L2 gDNA PCR Products, **7**; ST4L3 gDNA, **8**; ST4L3 gDNA + PCR Products, **9**; QM9414 gDNA, **10**; QM9414 gDNA + PCR Products, **M**; 1 kb. The gene in *reesei* genome that produces significant alignment with the used primers was called bifunctional dethiobiotin synthetase/adenosylmethionine-8-amino-7-oxononanoate aminotransferase (*bioDA*) gene, its size 2736 b.p.

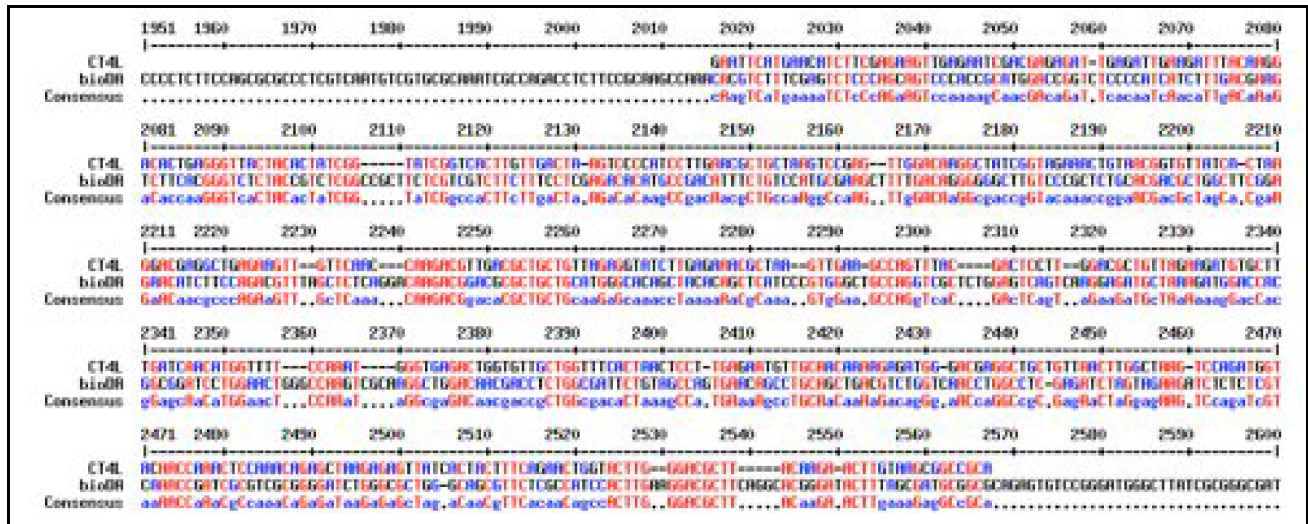


Figure 5: Multiple Sequence Alignment (MSA) by using MultAlin. **Red**; indicates HIGH consensus residues, **Blue**; for LOW consensus residues, and **Black**; NEUTRAL consensus residues,

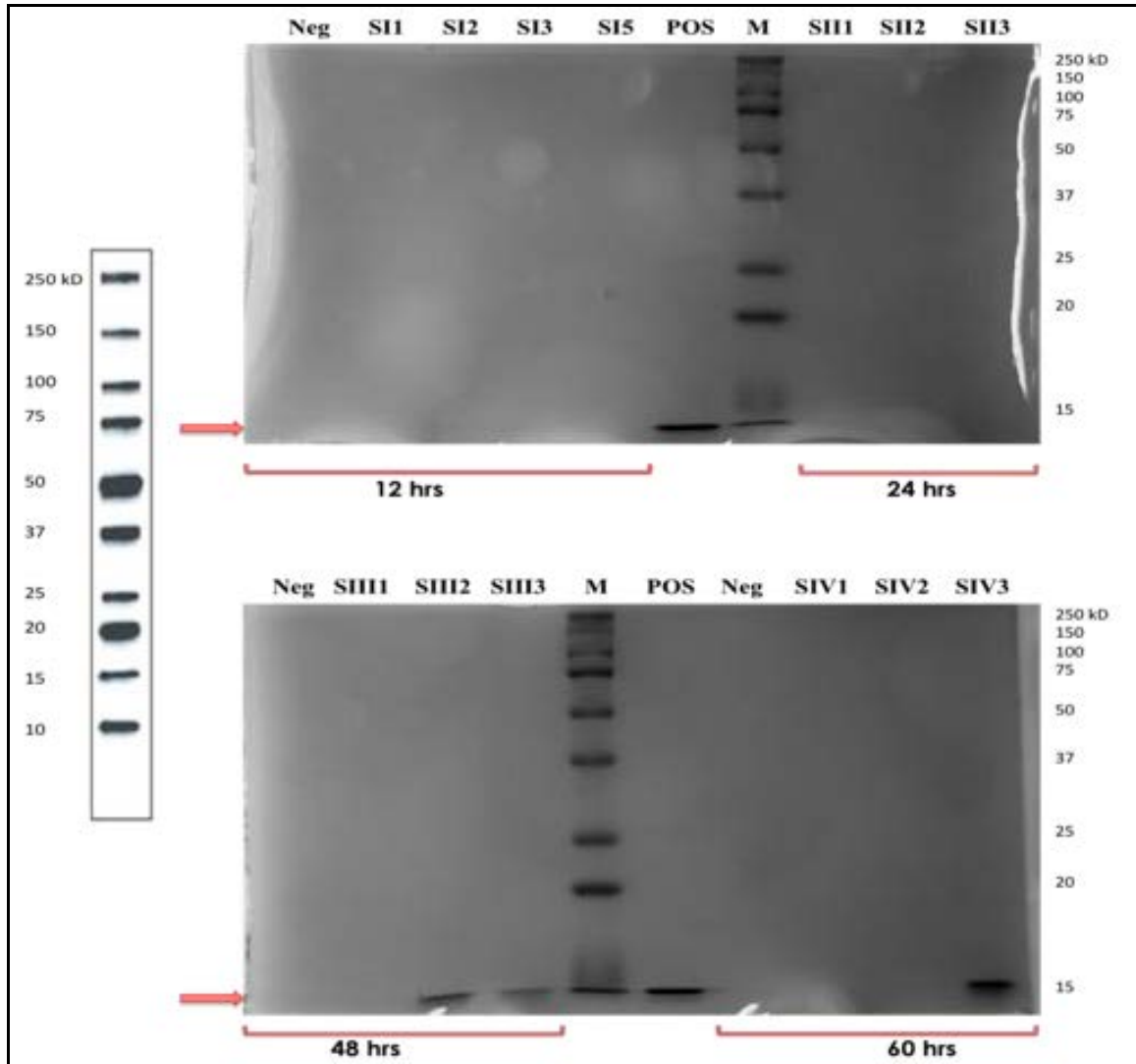


Figure 6: SDS-PAGE, for gel picture on the top: Neg; QM6a, SI1; ST4L1, SI2; ST4L2, SI3; ST4L3, POS; mol wt ~15 kDa, M; Lader 250 kDa, SII1; ST4L1, SII2; ST4L2, SII3; ST4L3. For gel picture on the bottom: Neg; QM6a, SIII1; ST4L1, SIII2; ST4L2, SIII3; ST4L3, M; Lader 250 kDa, POS; mol wt ~15 kDa, Neg; QM6a, SIV1; ST4L1, SIV2; ST4L2, SIV3; ST4L3.

Table:**Table 1:** Sequencing results along with BLAST search results

Transformant	Sequencing Results	Primer Used	Alignment to	Similarity (%)	Description
ST4L1	ST4L1-6	FWD	Database	98%	Synthetic T4 lysozyme gene, complete cds
			Synthetic <i>t4l</i>	100%	
			<i>T. reesei</i> QM6a gDNA	0	
			<i>bioDA</i> gene	0	
	ST4L1-5	REV	Database	71%	Synthetic T4 lysozyme gene, complete cds
			Synthetic <i>t4l</i>	98%	
			<i>T. reesei</i> QM6a gDNA	0	
			<i>bioDA</i> gene	0	
ST4L2	ST4L2-6	FWD	Database	0	
			Synthetic <i>t4l</i>	34%	
			<i>T. reesei</i> QM6a gDNA	0	
			<i>bioDA</i> gene	0	
	ST4L1-5	REV	Database	69%	Synthetic T4 lysozyme gene, complete cds
			Synthetic <i>t4l</i>	70%	
			<i>T. reesei</i> QM6a gDNA	3%	<i>bioDA</i> gene
			<i>bioDA</i> gene	3%	<i>bioDA</i> gene
ST4L3	ST4L3-6	FWD	Database	0	
			Synthetic <i>t4l</i>	0	
			<i>T. reesei</i> QM6a gDNA	0	
			<i>bioDA</i> gene	0	
	ST4L1-5	REV	Database	97%	Synthetic T4 lysozyme gene, complete cds
			Synthetic <i>t4l</i>	97%	
			<i>T. reesei</i> QM6a gDNA	0	
			<i>bioDA</i> gene	0	

References

- Benko Z, Drahos E, Szengyel Z, Puranen T, Vehmaanpera J and Reczey K. Thermoascus aurantiacus CBHI/Cel7A production in Trichoderma reesei on alternative carbon sources. *Appl Biochem Biotechnol* 2007; 137-140: 195-204.
- Chambergó, F. S., Bonaccorsi, E. D., Ferreira, A. J., Ramos, A. S., Ferreira Junior, J. R., Abrahao-Neto, J., El-Dorry, H. (2002). Elucidation of the metabolic fate of glucose in the filamentous fungus *Trichoderma reesei* using expressed sequence tag (EST) analysis and cDNA microarrays. *J Biol Chem*, 277(16), 13983-13988. doi: 10.1074/jbc.M107651200
- Chen, X., Liang, Y., Hua, J., Tao, L., Qin, W., & Chen, S. (2010). Overexpression of bacterial ethylene-forming enzyme gene in *Trichoderma reesei* enhanced the production of ethylene. *Int J Biol Sci*, 6(1), 96-106.
- Cianchetta S, Galletti S, Burzi PL and Cerato C. (2010). A novel microplate-based screening strategy to assess the cellulolytic potential of *Trichoderma* strains. *Biotechnol Bioeng*; 107: 461-468.
- Corpet F. (1988). Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res*, 16(22):10881–10890.
- Dahiya, N., Tewari, R., & Hoondal, G. S. (2006). Biotechnological aspects of chitinolytic enzymes: a review. *Appl Microbiol Biotechnol*, 71(6), 773-782. doi: 10.1007/s00253-005-0183-7
- Ferreira, N. L., and Margeot, A. (2011). Patent US 2011/0008862 A1, United States, Patent Application Publication.
- Henrissat B., (1991). A classification of glycosyl hydrolases based on amino-acid sequence similarities. *Biochem. J.* 280:309-316(1991).
- Leboffe M. J. & Pierce B. E., (2005). *A Photographic Atlas for the Microbiology Laboratory*, 3rd Edition, Morton Publishing Company
- Mandels, M. M., and Andreotti R.E. (1978). The cellulose to cellulase fermentation. *Proc. Biochem*; 13: 6-13.
- Martinez, D., Berka, R. M., Henrissat, B., Saloheimo, M., Arvas, M., Baker, S. E., Brettin, T. S. (2008). Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei* (syn. *Hypocrea jecorina*). *Nat Biotechnol*, 26(5), 553-560. doi: 10.1038/nbt1403
- Michael J. Leboffe & Burton E. Pierce, 2005, *A Photographic Atlas for the Microbiology Laboratory*, 3rd Edition, Morton Publishing Company, ISBN-10: 0895826569
- Minz, A., & Sharon, A. (2010). Electroporation and Agrobacterium-mediated spore transformation. *Methods Mol Biol*, 638, 21-32. doi: 10.1007/978-1-60761-611-5_2
- Mizuno, R., Itoh, Y., Nishizawa, Y., Kezuka, Y., Suzuki, K., Nonaka, T., & Watanabe, T. (2008). Purification and characterization of a rice class I chitinase, OsChia1b, produced in *Escherichia coli*. *Biosci Biotechnol Biochem*, 72(3), 893-895.
- Nakari-Setälä, T., & Penttilä, M. (1995). Production of *Trichoderma reesei* cellulases on glucose-containing media. *Appl Environ Microbiol*, 61(10), 3650-3655.
- Okazaki, K., Yamashita, Y., Noda, M., Sueyoshi, N., Kameshita, I., & Hayakawa, S. (2004). Molecular cloning and expression of the gene encoding family 19 chitinase from *Streptomyces* sp. J-13-3. *Biosci Biotechnol Biochem*, 68(2), 341-351.
- Pakula, T. M., Salonen, K., Uusitalo, J., & Penttilä, M. (2005). The effect of specific growth rate on protein synthesis and secretion in the filamentous fungus *Trichoderma reesei*. *Microbiology*,

- 151(Pt 1), 135-143. doi: 10.1099/mic.0.27458-0
- Saloheimo, M., & Pakula, T. M. (2012). The cargo and the transport system: secreted proteins and protein secretion in *Trichoderma reesei* (Hypocrea jecorina). [Review]. *Microbiology*, 158(Pt 1), 46-57. doi: 10.1099/mic.0.053132-0
- Seidl, V., Seibel, C., Kubicek, C. P., & Schmoll, M. (2009). Sexual development in the industrial workhorse *Trichoderma reesei*. *Proc Natl Acad Sci U S A*, 106(33), 13909-13914. doi: 10.1073/pnas.0904936106
- Szewczyk E, Nayak T, Oakley CE, Edgerton H, Xiong Y, Taheri-Talesh N, Osmani SA, Oakley BR. (2006). Fusion PCR and gene targeting in *Aspergillus nidulans*. *Nat Protoc*, 1(6):3111–3120.]
- Lorito, M., Woo, S. L., Harman, G. E., & Monte, E. (2010). Translational research on *Trichoderma*: from 'omics to the field. *Annu Rev Phytopathol*, 48, 395-417. doi: 10.1146/annurev-phyto-073009-114314
- Trun, N., & Trempy, J. (2004). *Fundamental Bacterial Genetics*. UK: Blackwell Publishing.
- Turgeon, B. G., Condon, B., Liu, J., & Zhang, N. (2010). Protoplast transformation of filamentous fungi. *Methods Mol Biol*, 638, 3-19. doi: 10.1007/978-1-60761-611-5_1
- Watanabe, T., Kanai, R., Kawase, T., Tanabe, T., Mitsutomi, M., Sakuda, S., & Miyashita, K. (1999). Family 19 chitinases of *Streptomyces* species: characterization and distribution. *Microbiology*, 145 (Pt 12), 3353-3363.

Chapter V: General Discussion and Future Directions

Enzymes that are able to hydrolyze cell membrane polymers generally play important roles for protecting plants and animals against infection by insects and pathogens. All of the work presented here in this thesis approaches the challenge of expressing a codon-modified gene encoding T4L protein in different microorganisms with a main focus on the expression of that gene. The research presented in this thesis led to the discovery of the possibility of T4L-producing bacteria as well as fungi with potential in future studies for the characterization and exploitation of that enzyme for protecting plants and animals against infection by insects and pathogens. Additionally, this work lends knowledge to the future study and use of *T. reesei* in simultaneous production of antimicrobial agents. However, until now there has been no global industrialization of T4L. Currently, there are several challenges in producing this enzyme.

Firstly, my work focused on studying the synthetic *t4l* bioinformatically. With that work, the synthetic *t4l* has been compared to its native nucleotide sequence, and then the 3D protein structures of both proteins were analysed. The findings allowed us to determine the variation between these T4Ls in computational biology to demonstrate the results of codon bias modification on the nucleotide sequence onto the amino acid sequence. The two altered amino acids found in this study have shown the average distance between atoms when the backbone coordinate was calculated by RMSD. However, these results gave us a quantitative comparison between the structure of a partially folded protein and the structure of the native state. This shows us how well a submitted structure matches the native state, which raises predictions on whether the produced enzyme will be in an active or inactive state after its expression in a microorganism. Further, if it is in an inactivate state, the question will be whether this is related to the codon-based modifications or due other biological approach matter. Whatever the end result, future work is required for understanding the T4L from a bioinformatics approach. This includes studying the genetic variation of other mutants in the database and/or synthetic T4L genes that have been used with other studies. Because the genetic variation of this enzyme has increased the demand to uncover it via a bioinformatics approach, there will be an advantage of effective mutations and/or modifications occurring

in specific genetic regions. Application of such an approach will be useful to unclassified mutations as well as provide strong support for those interested in working on this protein at advanced levels or in the literature. Moreover, predicting the effects of amino acid substitutions on protein function by using computational biology will be very useful for further research in terms of time and cost. Some genetic variations are neutral, but single base changes in and around a gene can affect its expression or the function of its protein products. Nonsynonymous single nucleotide polymorphisms (nsSNPs) are coding variants that introduce amino acid changes in their corresponding proteins. However, nsSNPs can affect protein function. Therefore, it is important to distinguish those nsSNPs that affect protein function from those that are neutral, super-expressed, and more stable functionally. In addition, amino acid substitution (AAS) prediction methods could be applied to this benchmark, which use sequences and/or structure to predict the effect of an AAS on protein function.

In my second study, I used a well-known molecular biology technique in the incorporation and expression of exogenous genetic material—synthetic *t4l*—in a bacterial cell. There were several reasons for working on a bacterial strain at the beginning. First, bacteria can be easily cultured and manipulated with the transformation. Second, bacteria have long been used in the field of molecular biology. These attributes lend to the potential exploitation of producing recombinant enzymes by identifying a bacterial strain that is commonly used in laboratory settings for genetic alteration and that could also produce certain enzymes with antimicrobial properties. After confirming the resistance of the *E. coli* BL21(DE3)pLysS strain towards the commercial lysozyme, our transformants successfully carried out our plasmids and manipulated them. In addition, our transformants displayed a successful expression of our synthetic *t4l*. Thus, our transformants BL21(DE3)pLysS with plasmids bearing *t4l* successfully produced T4L. However, the T4L enzyme produced needs to be measured and its activity characterized. This can be accomplished with two methods: plate bioassay and turbidity assay. These two assays will complement each other to determine and measure the recombinant T4L production. Plate bioassay is based on plate growth inhibition by visualizing the clear zone of the agar plate. Growth inhibition will be indicated by a

clear zone on the agar plate, whose diameter will be proportional to the catalytic activities of the enzyme. *M. luteus* is a good candidate to be used as a test strain, and in this case commercial lysozyme could be used as a positive control. Thus, plate bioassay is simple and a good choice for determining T4L activity. On the other hand, measuring the production of recombinant T4L could be performed by turbidity assay. A spectrophotometer could be used to find the measurement of turbidity in liquid, with *M. luteus* again used as a test strain. The activity of recombinant T4L will be determined by measuring the decrease in optical density of a tested strain. Also, the thermostability and pH effects of the enzyme could be tested in the same manner. Not surprisingly, researches have produced a recombinant T4L; however, it appears to be inactive, probably due to the absence of correctly formed disulphide bonds. For instance, incorrect folding of T4L resulting in an enzymatically inactive protein has been reported in *E. coli* under conditions of overproduction (van de Guchte *et al.*, 1992).

In addition, over the past few decades a number of reports have highlighted the growing importance of electrochemical biosensing devices in both clinical and environmental analyses. This may not seem as surprising when one considers that electrochemical transduction processes have inherent advantages, such as low cost, high sensitivity, independence from solution turbidity, simple miniaturization that is well-suited to microfabrication, low power requirements, and relatively straightforward associated instrumentation. These characteristics make electrochemical-biosensing methods highly attractive in myriad applications, including the detection of cancer, infectious diseases, and biological compounds (for example T4L), to name a few. The aims of the biosensor work will include the development of a detection method for T4L using Surface Plasmon Resonance (SPR) technology, for example, which will produce a detection system with a greater throughput and a greater limit of detection. An anti-lysozyme aptamer will be used as a receptor for T4L in a buffered sample. Part of the T4L biosensor method will include the regeneration of the sensor chip for further lysozyme measurements in general. Once this system has been established and optimized, then samples will be tested. A comparison of the T4L content will be measured by ELISA, and the biosensor method will be carried out as a confirmatory test.

The remaining focus of my studies was to utilize molecular biology techniques to further improve a known extracellular enzyme-producing fungal system. I chose the well-known filamentous *T. reesei*, which is used for cellulase and other industrial applications related to biorefining. There were many advantages for choosing this strain over *E. coli*. *T. reesei* is eukaryotic, and thus a foreign gene can be integrated into the filamentous fungus genome and stable transformants obtained. Also, it is well known that this fungus secretes proteins extracellularly along with inducer regulation by a strong promoter, and post-translational modification and well-established fermentation technology has been applied to this strain in other studies. Additionally, *T. reesei* is non-toxic to humans and animals, resulting in a strain that has no side effects and is cost effective during the industrial production of the antimicrobial agent T4L. In addition, genetic transformation is an essential part of modern fungal research. Primarily, protoplasts' production and recovery might be problematic, especially in slow growing species. It is also known that results vary considerably between different enzyme batches, resulting in inconsistent transformation efficiencies (Minz and Sharon, 2010).

Thus, my next project focused on the development of a genetically engineered strain of *T. reesei* QM6a, which is capable of producing an antimicrobial agent. The addition of exogenous T4L gene originated from bacteriophage T4 proved to be a substantial challenge in producing recombinant T4L with such a system. Thus, I chose to express our synthetic *t4l* in *T. reesei* QM6a. Due to the difficulty in genetic modification of *T. reesei* and due to time constraints, I chose a number of methods that generated genetically modified strains by using the fungal system in general. Then, I tried to match the most similar steps in those protocols and considered the major different steps. From a historical perspective, the first eukaryotic organism to be transformed was *Saccharomyces cerevisiae* in 1978 (Hinnen *et al.*, 1978), and the first filamentous fungus transformed was the saprobe *Neurospora crassa* (Case *et al.*, 1979). These achievements were followed by the transformation of the saprobe *Aspergillus nidulans* in 1984 (Yelton *et al.*, 1984). The efficiency of transformation, i.e., the fraction of cells that take up DNA, is low (Turgeon *et al.*, 2010); therefore, a strategy was used here to identify those cells that have been transformed, which

was the introduction of a selectable marker that transforms the original strain from drug sensitive to drug resistant. That strategy was much easier to work with during the transformation process. In our plasmids, the hygromycin phosphotransferase (hph) gene was used, which confers resistance to hygromycin B as a selectable marker. A number of choices for drugs are available and sensitivity of protoplasts to the drug should be experimentally determined on the solid medium for the fungal strain used for transformation. In the study, a range of hygromycin B concentrations were tested on the growth of *T. reesei* QM6a on PDA agar plates, and 25 µg/ml was the optimal concentration for transformants. During this study, I was able to design a modified protoplast-based modification to be used in future studies for *T. reesei* genetic modification. I also expressed synthetic *t4l* from QS, QL, T4L-pET20b(+) plasmids and saw a significant production with SDS-PAGE.

Therefore, in my last study, I investigate the development of the stable transformant of *T. reesei* QM6a carrying our synthetic *t4l*. The developed system here lends knowledge for future improvements. Our method of producing transformable protoplast could be applied by using other genes of interest with the QM6a strain, while a synthetic *t4l* into the genome of *T. reesei* QM6a strain was successfully engineered and the enzyme produced extracellularly. I propose that development of a novel strain that produces both chitinase and lysozyme, as both are key enzymes in the plant defense system against fungal infection, will have huge value in the industry. However, during this study I note some important factors that could affect application of the discussed protocol in this thesis, which lends knowledge towards future studies. Firstly, the lytic enzyme used can alter the overall transformation rate. Novozyme is widely used; however, it does not always work well (Turgeon *et al.*, 2010). In some cases, it may be difficult to digest cell walls if grown too long, so fresh sample is always recommended. Also, adjusting the time from adding lytic enzyme to generating protoplasts is critical. Losing produced protoplasts in this step happens with different laboratory conditions. Secondly, a lower concentration of osmotic stabilizer is utilized in the medium on which the transformants are plated. Subsequent chilling of the protoplasts in the presence of divalent cations such as Ca^{2+} (in CaCl_2) renders protoplasts permeable to DNA and thus “competent” to

take up DNA (Turgeon *et al.*, 2010). Thirdly, strains carrying mutations could transform less efficiently for reasons that have not been investigated in detail and are not yet understood. Of note, quality of protoplasts deteriorates upon storage. We do not recommend using frozen protoplasts for important experiments, since the entire transformation procedure is labor intensive.

Although it is beyond the scope of this protocol to discuss the handling of transformation and assessment of transformants in detail, it is absolutely necessary to verify that transformants have the correct insertion. Gene replacements could be analyzed by quantitative Real-Time PCR (qRT-PCR). Spores are streaked on selective media to allow growth of single colonies. DNA is extracted from these colonies and tested. This will show whether an insertion of the right size is present in the right place in the genome; it will also verify that a single correct insertion has occurred and that no insertions have occurred elsewhere in the genome. The copy number of the synthetic *t4l* integrated into the *T. reesei* genome of the transformants will be known by this analysis. If transformants obtain more copy of the synthetic *t4l*, their enzyme productions will be compared to those with fewer copy numbers.

Additionally, if there is time, my continuing work will be related to the effects of growth rate on protein synthesis and secretion in the filamentous fungus *T. reesei*. Very little information is available on the cellular responses to protein production. *T. reesei* has the potential to produce extracellular proteins in very large quantities, which sets a demand for the cells to adjust the capacity of protein synthesis and transport to the level required, and may also provoke stress responses within the cells. To obtain information on protein production and factors affecting the processes at different growth rates of the organism, carbon-limited chemostat cultures of the strain QM6a, both wild-type and our transformants, must be analysed in detail under different physiological conditions and under the influence of different factors affecting protein production by the fungus. Specifically, the capacity of the cells to synthesize and secrete proteins could be studied using metabolic labeling of the proteins after setting up a methodology for analysis of protein synthesis and secretion in a defined culture condition. In addition, the expressions of the major cellulose gene *cbhI*, as well as genes involved in protein folding and transport, have to be

analysed under these conditions. This study could offer new insight into the protein production characteristics of the industrially important host organism *T. reesei* (Pakula *et al.*, 2005). Also, genetic modification in fungal organisms continues to advance, rendering greater opportunities for the development of economically feasible greener technologies to produce enzymes for biorefining applications.

In conclusion, the synthetic *t4l* has been successfully studied and expressed in both biological systems; *E. coli* and *T. reesei*, and T4L-producing transformants have potential for further study. This thesis project still has huge potential. If more time was allowed, several additional experiments would have been included to both extend work executed in all chapters and also to explore other avenues in the hope of producing better tests and results. In brief, all of the transformants that are known to be T4L-producing transformants must have their activities measured towards the known bacterial sensitivity and/or resistance to lysis activity of the enzyme to determine how similar our modified enzyme is to the native T4L. This will allow us to determine the practicality of these transformants in an industrial setting while addressing industrial, environmental, technical, and economic challenges.

References

- Case, M. E., Schweizer, M., Kushner, S. R., & Giles, N. H. (1979). Efficient transformation of *Neurospora crassa* by utilizing hybrid plasmid DNA. *Proc Natl Acad Sci U S A*, 76(10), 5259-5263.
- Henrissat B. (1991). A classification of glycosyl hydrolases based on amino-acid sequence similarities. *Biochem. J.* 280:309-316.
- Hinnen, A., Hicks, J. B., & Fink, G. R. (1978). Transformation of yeast. *Proc Natl Acad Sci U S A*, 75(4), 1929-1933.
- Minz, A., & Sharon, A. (2010). Electroporation and Agrobacterium-mediated spore transformation. *Methods Mol Biol*, 638, 21-32. doi: 10.1007/978-1-60761-611-5_2
- Pakula, T. M., Salonen, K., Uusitalo, J., & Penttila, M. (2005). The effect of specific growth rate on protein synthesis and secretion in the filamentous fungus *Trichoderma reesei*. *Microbiology*, 151(Pt 1), 135-143. doi: 10.1099/mic.0.27458-0
- Yelton, M. M., Hamer, J. E., & Timberlake, W. E. (1984). Transformation of *Aspergillus nidulans* by using a *trpC* plasmid. *Proc Natl Acad Sci U S A*, 81(5), 1470-1474.