Prediction of Potential Epitope within Segment 1 of Tilapia Lake Virus (TiLV) using In-Silico Immunoinformatic Approach

NurulJannah Binti Ariffin

F16A0218

Bachelor of Applied Science (Husbandry Science) (Hons)

Faculty of Agro-Based Industry

University Malaysia Kelantan

2021

Lampiran I

PERAKUAN

Saya akui karya ini adalah hasil kerja saya sendiri kecuali nukilan dan ringkasan yang setiap satunya telah saya jelaskan sumbernya.

_____

Tandatangan

Nama Pelajar : NurulJannah Binti Ariffin

No Matriks : F16A0218

Tarikh : 22/2/2022

Disahkan oleh :

_____

Tandatangan Penyelia

Nama Penyelia :

Cop :

Tarikh :

ACKNOWLEDGEMENT

This thesis could not have been done without participation and assistance of many people whose get directly involved and not directly involved during the preparation of this thesis. Their contributions in this thesis are sincerely appreciated and acknowledged.

Firstly, I am deeply grateful to Allah because I could do this thesis in a given time. I would like to express my gratitude to my parents, Mr. Ariffin bin Hassan and Mrs. Nasrina Binti Habib Mohd who are never stop supporting me to finish this thesis.

I wish to express my sincere thanks to my supervisor, Dr. Hazreen Nita Binti Mohd Khalid for her supervision, guidance and idea on how to improve the thesis to become a good thesis. It was a great honor for me to be under her supervision.

Prediction of Potential Epitope within Segment 1 of Tilapia Lake Virus (TiLV) using In-silico Immunoinformatics Approach

ABSTRACT

Tilapia Lake Virus (TiLV) is a new emerging disease that responsible for the mortality of farmed tilapia in many countries such as Israel, Egypt, Thailand and Malaysia. To date, there is still no commercial cure or vaccines that would treat the TiLV disease. Therefore, this study was conducted to predict B-cell and T-cell epitopes within Segment 1 protein of TiLV genome through immunoinformatics tools which later could be proposed as candidate target for development of peptide-based treatment or prevention tools. In this study, epitope is referring to part of antigen protein that can be recognized by B-cells or T-cells to initiate immune response. In addition, this study also determined the antigenicity and allergenicity profile of the selected epitopes to further ensure the potential of antigenic peptides. The sequence of amino acids of Segment 1 (Acc. No. KU751814) was obtained and downloaded from Genbank through National Centre International Biotechnology (NCBI) database in FASTA format. A series of immunoinformatics prediction was employed which involved prediction of B-cell epitopes through Kolaskar and Tangoankar tool, prediction of T-cell epitopes through Proped and Proped 1 server and prediction of binding affinity between epitopes and antigen-presenting cell of major histocompatibility complex (MHC) molecules through MHCPred 2.0.Further analysis was done to identify antigenicity and allergenicity profile of the epitope through Vaxijen 2.0 and AllerTop2.0 respectively. As a result, the present study has successfully identified 10 antigenic peptides that are potential to be proposed as candidate targets for development of peptide-based therapeutics such as vaccine. This study employed a computational approach to predict potential epitopes that are antigenic and safe, within Segment 1 of TiLV genome with advantages of speeding up the long and costly process of candidate target discovery for development of peptide-based therapeutic.

Keywords: Tilapia lake virus (TiLV), epitope, Immunoinformatics, computational study, allerginicity

Ramalan Potensi Epitop dalam Segmen 1 Virus Tasik Tilapia (TiLV)

Menggunakan Pendekatan In-Silico

ABSTRAK

Virus Tasik Tilapia (TiLV) adalah penyakit baru muncul yang bertanggungjawab terhadap kematian tilapia yang diternak di kebanyakan negara seperti Israel, Mesir, Thailand dan Malaysia. Sehingga kini, masih tiada penawar komersial atau vaksin untuk merawat penyakit TiLV. Oleh itu, kajian ini dijalankan untuk meramalkan epitop sel B dan sel T dalam protein Segmen 1 genom TiLV melalui kaedah imunoinformatik yang kemudiannya boleh dicadangkan sebagai sasaran calon untuk pembangunan kaedah rawatan atau pencegahan berasaskan peptida. Dalam kajian ini, epitope merujuk kepada sebahagian daripada protein antigen yang boleh dikenali oleh sel B atau sel T sebagai permulaan respon imunisasi. Selain itu, kajian ini juga menentukan profil antigen dan alergenik bagi epitop terpilih bagi mengenalpasti potensi peptide antigenik. Jujukan asid amino Segmen 1 (No Ak., KU751814) telah diperoleh dan dimuat turun daripada Genbank melalui pangkalan data National Center International Biotechnology (NCBI) dalam format FASTA. Satu siri kaedah ramalan imunoinformatik telah digunakan yang melibatkan ramalan epitop sel B melalui kaedah Kolaskar dan Tangoankar, ramalan epitop sel T melalui pelayan Proped dan Proped 1 dan ramalan pertalian pengikat antara epitop dan kompleks sel antigen pembentang, histokompatibiliti utama (MHC) melalui MHC Pred 2.0. Analisis lanjutan dilakukan untuk mengenal pasti profil antigen dan alergenisitas epitope melalui Vaxijen 2.0 dan AllerTop2.0 masing-masing. Hasilnya, kajian ini telah berjaya mengenal pasti 10 peptida antigenik yang berpotensi untuk menjadi sasaran calon bagi pembangunan terapeutik berasaskan peptida seperti vaksin. Kajian ini menggunakan pendekatan komputasional untuk meramalkan epitop berpotensi yang antigenik dan selamat, dalam Segmen 1 genom TiLV dengan kelebihan mempercepatkan proses penemuan sasaran calon yang panjang dan mahal untuk pembangunan terapeutik berasaskan peptida.

Kata kunci: Virus tasik tilapia, epitop, Immunoinformatik, kajian komputer, alergenik

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND SYMBOLS

TiLV	Tilapia Lake Virus
OIE	World Organization of Animal Health
PCR	Polymerase chain reaction
RNA	Ribonucleic acid
MHC	Major Histocompatibility Complex
NACA	Network of Aquaculture Centres in Asia-Pacific
HLA	Human Leukocyte Antigen

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Research Background

Aquaculture plays an important role to meet the need of foods for human. This sector provides the consumers with animal proteins which are highly nutrition. Economically, the food and income benefits provided in aquaculture industry can afford opportunities for empowering individuals where opportunities in other sectors are limited (Khay et al., 2018). Tilapia is the common name for several species of cichlid fish inhabiting freshwater streams, ponds, rivers and lakes and less commonly in brackish water. Although it is considered as an invasive species, tilapias had been reported as species that is now having increasing importance in aquaculture sector. Tilapia is also known as the second most farmed fish worldwide whereby its production has quadrupled over the past decade because of its suitability for aquaculture. Tilapia is

still preferred due to low trophic level feeder, fast rate of growth, low cost of production and highly demand in market.

Besides having many advantages, tilapia farming has a few flaws such as sensitivity to low temperature and low dress out percentage (James, 2019). Although Malaysia has started tilapia production a few years ago, the current production still not very promising due to lack of stock and seed production (Abdullah et al., 2018). Another issue that should be concerned is disease occurrence in tilapia such as fungi, bacterial and viral disease.

Disease is one of limiting factor in tilapia production because it could be an economic loss to the farmer due to high mortality in tilapia. Disease in tilapia could be affected by poor farming practices, conditions of water and lack of nutrients in tilapia feed.

Tilapia lake virus (TiLV) is an emerging disease which causes more than 90% mortality globally (Jansen et al., 2019). TiLV was first reported in Israel in 2013, although it has been found in the summer of 2009 and currently has been spread to a lot many of countries such as Egypt, Ecuador and including Malaysia. As the reports of mortality of tilapia have been increasing in worldwide, the source of virus is still remains unknown

1.2 Statement of the Problem

Early in 2017, Network of Aquaculture Centres in Asia-Pacific (NACA) and World Organization for Animal Health (OIE) had announced a new emerging disease in tilapia which designated as Tilapia Lake Virus (TiLV) that responsible for the mortality of farmed tilapia in many countries. TiLV is a negative-sense single-stranded RNA virus that belongs to family Amnoonviridae, under genus Tilapine virus and species of Tilapia tilapine virus. The first outbreak of TiLV occurrence was reported in Israel followed by other countries such as Ecuador, Colombia, Egypt, Thailand, Chinese Taipei, India, Malaysia, Bangladesh, Uganda, Tanzania, Peru, Mexico, Philippines, Indonesia, and USA. All life stages of tilapia fish are vulnerable to TiLV infection. Currently, there is no estimated number on socio-economic impact in national or global context has been published. However, TiLV infection could causes a significant amount of death in tilapia culture in most of countries reported.

TiLV in Malaysia was first detected in Timah Tasoh Lake in 2017, whereby sudden high mortality rate occurred among wild tilapia and wild river carps in the lake. A total of 10 black tilapia were caught and being observed for external clinical signs. The samples were tested for polymerase chain reaction (PCR) detection and found to be positive for TiLV. The results of the observation showed that the fish shown clinical

sign of watery, soft and haemorrhagic brain which was also known as signs of TiLV infection. Based on molecular identification of positive amplicon, sequencing results had shown 96% identity homology to TiLV Israel isolate. At the moment of this study, there is still no commercial cure or vaccines that would treat the TiLV disease.

1.3 Research Objectives

1. To predict B-cell and T-cell epitopes within Segment 1 protein of TiLV through immunoinformatics tools.
2. To determine antigenicity and allergenicity profile of the selected epitopes using Vaxijen 2.0 and AllerTop 2.0 prediction tools.

1.4 Scope of Study

Scopes of the present study are animal health and biotechnology. The knowledge of animal health is importance in this study as it involves understanding on the concept of antigen peptide recognition by antigen-presenting cells of MHC molecule to trigger host immune system during infection of particular pathogen. The presentation of pathogen-derived proteins results in the elimination of the infected cell by the immune system.

Biotechnology knowledge is widely applied in various fields such as agriculture, human health and computer technology (bioinformatics). Biotechnology used the knowledge of biology to solve problems and make useful products. The most prominent area of biotechnology is the production of therapeutic proteins and other drugs through genetic engineering. Through immunoinformatics, candidate target for peptide-based treatment or prevention tools shall be discovered through computational approach without the need of laboratory experiments.

1.5 Signification of Study

Currently, there is no effective treatment or vaccine that available to treat TiLV and database for development of TiLV vaccine is still limited. TiLV is a RNA virus which belongs to family Amnoonviridae that have genome that consists of 10 protein segments. In this study, Segment 1 of TiLV that encodes for RNA-directed RNA-polymerase (RdRp) gene was used to predict epitopes or antigenic peptide that important for development of peptide-based therapeutic such as vaccine in prevention of the disease. The present study employed computational approach to predict potential epitopes within Segment 1 of TiLV. Computational methods used in the present study have advantages to speed up the long and costly process of candidate vaccine identification for development of peptide-based therapeutic. Based on B-cell, T-cell epitopes analysis, antigenicity selection and allergenicity profile prediction, potential antigenic peptides shall be selected and proposed as candidate target for development of peptide-based treatment or vaccine in future study. This study highlighted a preliminary database of TiLV Segment 1 antigenic peptides that shall be used and tested in near future.

1.6 Limitation of Study

Tilapia lake virus is a new emerging disease that have affected in many countries including Malaysia. The disease was first detected in Israel in 2009. Therefore, there is no effective treatment or vaccine that available to treat TiLV and database for development of TiLV vaccine is limited. *In-silico* immunoinformatics is a new approach that provides analysis to predict the potential epitopes that could be recognized by MHC molecules in host. However, the database of each prediction tools is limited, thus the most commonly found MHC molecules of classical human MHC (HLA) molecules were used in the prediction analysis. This approach surely can speed up the development of vaccine however the general panorama is limited because this analysis imply the interaction of many cells and molecules and, in many cases, the interactions are might not yet to be known.

CHAPTER 2

LITERATURE REVIEW

2.1 Tilapia

The name of tilapia refers to several species of mostly freshwater fish that belong to the Cichlidae family (order Perciformes). Tilapia is mainly freshwater fish inhabiting shallow streams, ponds, rivers, and lakes, and less commonly found living in brackish water. Tilapia represented by numerous, mostly freshwater species native to Africa. Tilapia production has quadrupled over the past decade due to its suitability for aquaculture, marketability and stable market prices. Tilapias are currently introduced into over 90 countries for aquaculture and fisheries (Prabu et al., 2019) . Tilapia continued its rapid increase in global production.

In Malaysia, tilapia (*Oreochromis sp.)* is the second highest harvested freshwater fish, with a total annual production of 33,437 tonnes. The estimated wholesale value in 2015 of this species was RM 259 million, which indicates the importance of tilapia farming in Malaysia (Amal et al., 2018). In 2018, red hybrid tilapia (Oreochromis spp.) and the black tilapia, including the genetically improved farmed tilapia (GIFT)

(*Oreochromis niloticus*) were collectively contributed to about 30.7 % of the total freshwater aquaculture production in Malaysia (Mohamad et al., 2021).

2.2 General Disease in Tilapia Fish

Fish health can be conflicted in two factors which are fish disease and the environmental pollution. Disease in tilapia fish might be due to parasites, mycotoxins (fungi), bacterial infections and virus outbreak which usually lead to high mortality rates in fish production. The parasites live in the fish are obtaining protection and food from the fish and this may cause a serious damage to the fish. Various parasites in fish can make pathological changes such as abnormal color, excessive mucus production, inflammation and hemorrhages. Trichodinasis is common to tilapia which affected the swimming behavior of tilapia with a sudden movement and their fins have an opaque color while their gills were observed pale. In addition, Oodiniasis in tilapia appears as a velvety layer on skin and gill tissues while pleistofloris caused the fish to lose weight due to lose in appetite (Valladão et al., 2016) .

In general, *Streptococcus spp.* is one of the most common diseases caused by bacteria in tilapia fish. It shows signs of granulomatous lesions in internal organs of the fish and the other signs are eye problems, skin hemorrhages, ascitess, menngoencephalitis, pericardis, skin darkening and erratic or circular swimming. The diagnosis of the disease is based on the identification of the pathogen using microscopy or PCR detection in the samples (Basri et al., 2020).

In addition, another cause of tilapia disease is mycotoxins, which are highly heat resistant toxic metabolites produced by different species fungi that were caused by contaminated feed (Gonçalves et al., 2020). Mycotoxins caused tilapias to lose in body weight, growth impairment and higher rates of mortality which will cause low in the aquaculture's production (Magouz et al., 2018). Other than that, viruses also give negative impacts to fish production and health management. It is clear that the research on the virus-caused diseases in Egypt is insufficient due to the shortage of the cell lines which are needed for viral isolation and identification (Nicholson et al., 2017). However, few viruses have been incriminated in fish mass mortalities across the Egyptian state. A new emerging viral disease in tilapia is known as Tilapia Lake Virus (TiLV).

2.3 Tilapia Lake Virus (TiLV)

TiLV is a segmented RNA virus that contains 10 RNA segments with size from 465 to 1641 bp (10.323 kb in the total genome). The virus has been taxonomically assigned to Tilapia tilapine virus in the family Amnoonviridae. The gene on Segment 1 is predicted to code for the PB1 protein, which shows a very low, but detectible, similarity to the polymerase proteins of influenza (Bacharach et al., 2016). Other segments, however, showed no detectible similarities to any known genes in the NCBI database, indicating that it is a new virus that is distantly related to all known viruses. Indeed, TiLV is currently classified as the only member of a new genus Tilapine virus (Walker et al., 2019), the sole genus in the family Amnoonviridae. Family Amnoonviridae is a recently established family in the order Articulavirales, which also contains the influenza virus (family Orthomyxoviridae) (Walker et al., 2019).

A number of genome sequences of TiLVs have been reported, and analyses of these sequences have led to better understanding of how the virus evolves and spreads across regions(Ahasan et al., 2020). A recent study has shown that reassortment is very common for TiLV (Chaput et al., 2020). The reassortment is a phenomenon in which multiple strains of viruses with segmented genomes co-infect the same host cell and exchange their genetic materials (Thawornwattana et al., 2021). This means that

analysis of individual genomic segments might not give a full picture of how TiLV evolves, and on the other hand, analysis of whole genomes without taking re-assortment into account might yield an erroneous history of the virus.

2.4 Impact of Viral Disease Outbreak

There is no specific scientific studies on the sosio-economic of TiLV has been done. In late 2000s, a large reduction in annual wild catch has been recorded in Israeli Sea of Galilee's main edible fish. At the year of 2009, a large amount on loss of farmed tilapia recorded throughout Israel (Jansen et al., 2019). Following this, many studies had reported identification of the virus from samples collected in various countries such as Ecuador (Al-Hussinee et al., 2019), Colombia (Contreras et al., 2021),Egypt (Nicholson et al., 2017), India (Behera et al., 2018), Indonesia (Reantaso, 2017), Thailand (Yamkasem et al., 2021) and Malaysia (Abdullah et al., 2018).

In Malaysia, the affected fish shown to have lethargy, loss of appetite, swimming near the pond edge, pale and isolated from schooling group. There were also skin redness and hemorrhages, particularly at the operculum area and at the base of

dorsal, caudal and anal fins found on the fish. Histopathological examinations revealed that swollen hepatocytes, hemorrhagic spleens and perivascular cuffing consisted of mononuclear cells in the brains. The PCR and sequence had been analyzed and confirmed the presence of TiLV and *A. veronii* in the affected fish. Phylogenetic tree revealed that Malaysian's TiLV strain was closely related to the virus isolated in Israel than in Egypt or Thailand, while Malaysian's *A. veronii* strain 5L was closely related with strains from China. Based on the findings, it is proven that TiLV infection is a significant threat to global tilapia industry (Abdullah et al., 2018).

2.5 Disease Management, Control and Prevention

Viral diseases cause high economic losses in aquaculture. There are a lot of challenges which limit the control and prevention of viral diseases in aquaculture, The major challenges include increase in the number of emerging viral diseases, wild reservoirs, migratory species, anthropogenic activities, limitations in diagnostic tools and expertise, transportation of virus contaminated water, and international trade. These problems can be solved by developing and implement biosecurity policies at global and national levels, implementation of biosecurity measures, speed up vaccine development, use of antiviral drugs and probiotics to combat viral infections, selective breeding of

disease-resistant fish, use of improved diagnostic tools, disease surveillance, as well as promoting the use of good husbandry and management practices. A multifaceted approach combining several control strategies would provide more effective long-term solutions to reduce viral infections in aquaculture than using a single disease control approach (Mugimba et al., 2021).

Currently, there were no effective protocols and medication to reduce the impact of the virus outbreak (Nilav Aicha, 2020). Genetic selection of TiLV resistant tilapia brood stock and development of vaccines and appropriate biosecurity protocols might be offer a long-term in health management of the fish. If TiLV is detected in the farm, it is recommended to perform disinfection prior to World Organisation for Animal Health measures (Nilav Aicha, 2020). Any dead tilapia should be rejected because it is not suitable for human consumption and should not be fed to other animals.

Fish farmers should purchase fingerlings only from tested TiLV-free sources and to work with the local aquaculture extension agent and fish health professional to tailor an appropriate health management and biosecurity plan. Fish farmers also need to implement a one-month quarantine period of any imported stocks held at 25°C with daily monitoring for any signs of TiLV disease which includes bulging eyes, lethargy, loss of appetite, darkened or ulcerated skin, and decreased schooling behavior. Quarantine systems should be located in a separate building from resident fish.

Quarantine place should be separated and dedicated equipment is required and be handled only at the end of the day or only by employees who will not return to other farm areas that day. If there were an increase in mortality, A restriction movement of live tilapia, equipment, and employees should be implemented if there were a sudden increase in mortality to prevent transmission of TiLV to other uninfected parts of the premises (Surachetpong et al., 2020).

In addition to quarantine, the production team should practice standard biosecurity measures in all aspects of husbandry and farm/facility maintenance to minimize the introduction and spread of pathogens. Physical stress plays an important role in triggering outbreaks among cultured tilapia by decreasing their immune response to pathogens (Richard Kabuusu, 2017). Therefore, stress of the fish should be minimized by providing good quality water and nutrition, exercising care when handling fish, preventing overcrowding, and reducing or eliminating other unnecessary stressor.

2.6 Vaccination in Tilapia

Fish vaccination has been carried out for over 50 years and accepted as an effective method to prevent a lot of viral and bacterial diseases (Jansen et al., 2019). Vaccination gives positive impacts to environmental, social and economic sustainability in global aquaculture. Live vaccines are more effective because they mimic natural pathogen infection and generate a strong antibody response, therefore having a greater potential to be administered via oral or immersion routes. Modern vaccine technology has targeted specific pathogen components, and vaccines developed using such approaches may include subunit, or recombinant, DNA/RNA particle vaccines.

Although the fish vaccination has been carried out over 50 years, there are still no effective vaccine has to be found on market to control TiLV (Jansen et al., 2019). Researchers are still studying on how to treat the virus through vaccines. A study has been done by weiwei et.al. has identified that TiLV segment 8 encoded a protein (VP20) with immunogenicity sufficient for use as a vaccine antigen. In the study, the immune responses and protective efficacy elicited by VP20 with different prime-boost vaccination regimens (DNA only, protein only, and DNA plus protein) in tilapia was evaluated. As a result, both pV-opti VP20 plasmid and rVP20 protein induced humoral and cellular immune responses (Zeng et al., 2021). Tilapia in

the DNA prime-protein boost group developed significantly higher levels of antibody response compared with those immunized with either the DNA or the protein alone which indicates that the vaccine could elicit both humoral and cellular immune responses and significantly protected the fish against infectious challenge by TiLV. The DNA prime-protein boost immunization strategy showed the potential advantage over a solo vaccination (Mai et al., 2022).

2.7 Computational Study in Immunoinformatics

Immunoinformatics is a science which helps to create significant immunological information using bioinformatics software. The application of immunoinformatics is the prediction of a variety of specific epitopes for B-cell recognition and - cell through MHC class I and II molecules. This method reduces costs and time compared to laboratory tests (Raoufi et al., 2020). Immunoinformatics has paved the way for a better understanding of some infectious disease pathogenesis, diagnosis, immune system response and computational vaccinology. The importance of immunoinformatics in the study of infectious diseases is diverse in terms of computational approaches used. However, it is united by common qualities related to host–pathogen relationship. Bioinformatics methods are also used to assign functions to uncharacterized genes which can be targeted as a candidate in vaccine design and can be a better approach toward the inclusion of women that are pregnant into vaccine trials and programs (Oli et al., 2020).

Computational study had played an important role in vaccination study. There were a lot of researchers has used this approach in this study. A study about analysis of epitope-based vaccine against tuberculosis has been conducted using experimentally confirmed *Mycobacterium tuberculosis*, Mtb antigens, including Rv2608, Rv2684, Rv3804c (Ag85A), and Rv0125 (Mtb32A) which resulting known multi-epitope vaccine may activate humoral and cellular immune responses and maybe a possible tuberculosis vaccine candidate (Bibi et al., 2021). Another study conducted was to propose vaccine study on African animal trypanosomiasis caused by vector-transmitted parasites of the genus *Trypanosoma. T. congolense* and *T. brucei brucei* are predominant in Africa; *T. evansi* and *T. vivax* in America and Asia (Michel-Todó et al., 2020).

2.8 Antigenicity and Allegernicity in Vaccination Study Through Bioinformatics

Antigenicity reflects the ability of antibody responses in which the host develops against a virus and capacity of viruses to bind to specific antibody molecules (Zhang & Tao, 2015) meanwhile allergenicity refers to cause an abnormal immune response which leads to physiological function disorder or tissue damage. In allergenicity, computational approaches has been introduced and applied to identify T- cell epitopes to target-allergen T-cells to improve the safety of immunotherapy (Soria-Guerra et al., 2015). A study has been carried out to improve vaccine design against *Toxoplasma gondii* through bioinformatics approaches which analyzed different ROP8 protein (Foroutan et al., 2018). It resulting a several potential B and T-cell epitopes were identified and antigenicity and allergenicity evaluation remarked as immunogenic and non-allergen.

CHAPTER 3

METHODOLOGY

In this chapter 3, the epitopes within Segment 1 protein of Tilapia Lake Virus (TiLV) was determined using immunoinformatic approach described in Figure 3.1.

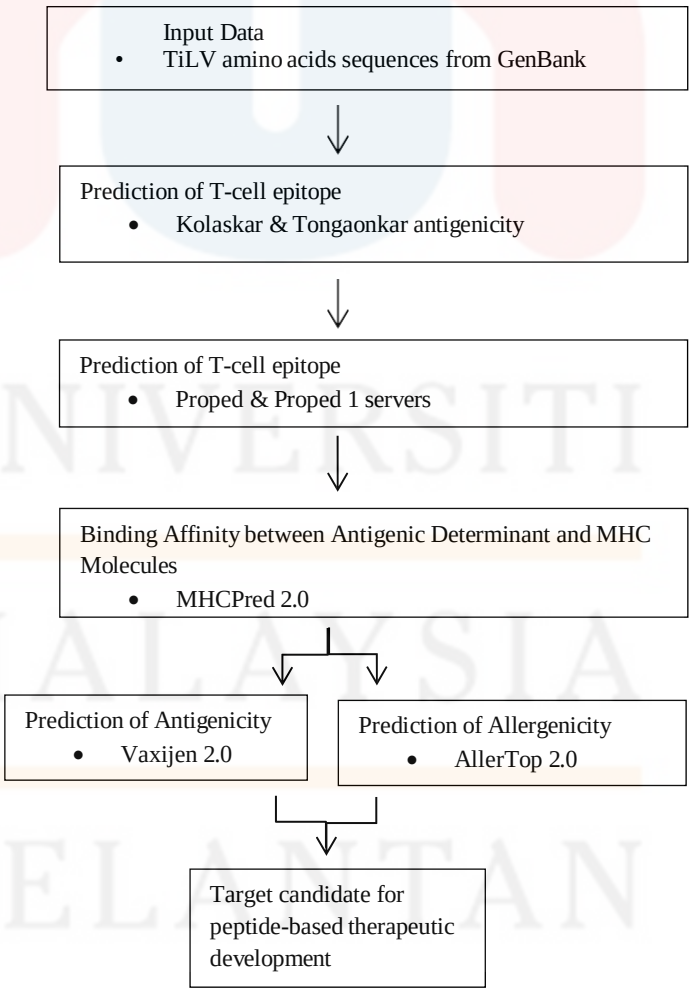


Figure 3.1: Flowchart of determination of epitopes within Segment 1 protein of TiLV using immunoinformatic approach.

3.1 The Amino Acids Sequences Retrieve from GenBank

The sequence of Segment 1 protein (Acc. No. KU751814) of TiLV were downloaded from GenBank through National Centre International Biotechnology (NCBI) server (https://www.ncbi.nlm.nih.gov/protein/). The amino acids sequences of Segment 1, TiLV were presented in FASTA format. The sequence was then subjected to a series of immunoinformatic tools to predict the epitopes as well as identifying their antigenicity and allergenicity profile which could have potential criteria as target candidate for development of peptide-based treatment or prevention tools.

3.2 Prediction of B-Cell Epitope

The B-cell epitope is part of antigen that antibodies bind to. In this study, Kolaskar and Tongaonkar antigenicity scale tool (http://tools.iedb.org/bcell/) was used to predict antigenic determinants on the antigen protein. A semi-empirical method has been used in this prediction which involved physicochemical properties of amino acid residues and the frequencies of occurrence in experimentally known as epitopes. The application of this method to a large number of proteins can be predicted 75% accuracy of antigenic determinants which is better than most known methods.

3.3 Prediction of T-Cell Epitopes

The result in subsection 3.2 were then subjected to MHC recognition site using ProPed-I (https://webs.iiitd.edu.in/raghava/propred1/index.html) and Proped (https://webs.iiitd.edu.in/cgibin/propred/index.html) servers to predict MHC class I and MHC class II binding regions towards TiLV antigen epitope. The threshold of PrePed-I was set at 4% while the immunoproteasome and proteasome filter were set at threshold of 5% respectively. This server is used to predict the MHC class-II binding regions in an antigen sequence by using quantitative matrices which will insist in locating promiscuous binding regions that are important in selecting vaccine candidates. The threshold is set at 3% as an optimum threshold to avoid lower number of antigenic

3.4 Prediction of Binding Affinity between Selected Epitopes and MHC Proteins by MHCPred Version 2.0

The prediction of binding affinity of the selected antigenic obtained from the previous section towards MHC proteins was conducted through MHCPred 2.0 (http://www.ddg-pharmfac.net/mhcpred/MHCPred/). The model database of

MHCPRED is a partial least squares-based multivariate approach to the quantitative prediction of peptide binding to major histocompatability complexes (MHC), the key checkpoint on the antigen presentation pathway within adaptive cellular immunity. MHCPred 2.0 in this study implemented robust statistical models for both for MHC Class I (HLA-A 0201, HLA-A 1101 and HLA-B 3501) and MHC Class II (HLA DRB 0101).

HLA-A 1101 is a human leukocyte serotype. Serotype is determined by the antibody recognition of α1 subset of HLA-A α-chains. For A11, the alpha "A" chain are encoded by the HLA-A11 allele group and the β-chain are encoded by B2M locus. The serotype of HLA-A 0201 is determined by the antibody recognition of the α2 domain of the HLA-A α-chain. For A 02, the α chain is encoded by the HLA-A 02 gene and the β chain is encoded by the B2M locus HLA-A 0201 is a vital restriction element for peptide presentation to T cells in disease and cancer. Mutation studies and analyses using cytotoxic T-lymphocytes have shown the functional relevance of subtype-specific differences in HLA-A2 molecules for peptide binding and T-cell receptor recognition. HLA-B35 (B35) is an HLA-B serotype which identified the more common HLA-B 35 gene products. B35 is one of the largest B serotype groups and currently has 97 known nucleotide variants and 86 polypeptide isoforms and it particularly susceptible to HIV infection.

DR1 is associated with seronegative-rheumatoid arthritis, penicillamine-induced myasthenia and schizophrenia. DR1 is increased in patients with systemic sclerosis andarthritis and in ulcerative colitis with patients that have articular manifestations. In this study, only HLA DRB 0101 was only being applied. DRB 0101 is associated with rheumatoid arthritis, in anti-Jk mediated hemolytic transfusion reactions, foliaceous pemphigus, HTLV-1-associated myelopathy/tropical spastic paraparesis, and lichen planus. In lyme disease arthritis, 0101 appears to play a role in presentation of triggering microbial antigens(Ali et al., 2010).

3.5 Prediction of Antigenicity and Allergenicity of the Potential Antigenic Peptide

The antigen determinants that have great binding affinity towards MHC molecules were subjected to antigenicity analysis through Vaxijen 2.0 server (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html) which uses internal leave-one-out cross validation (LOO-CV) on its training sets and external validation test sets with accuracy range between 70% and 89%. In this study, this server was used to predict the protective viral at the threshold of 0.4 with 70% accuracy and 89% sensitivity. The analysis allows antigen classification based on the physiochemical properties of the peptide irrespective of the sequence length. This model had shown remarkable stability that known as reliable and consistent tool for protective antigen

prediction. This analysis will determine whether the peptides chosen were antigen or non-antigen. The non-antigen was not included in this study.

The final server that was used in this study is Allertop version 2.0 (https://www.ddg-pharmfac.net/AllerTOP/index.html) which is use to asses allergenicity towards human interactions. The purpose of the prediction is to select non-allergen antigenic determinant that could be a potential candidate for TiLV vaccine development.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 The Amino Acids Sequences in Segment 1 of TiLV

The amino sequenced in Segment 1 of TiLV was retrieved from GenBank with accession number ID of KU751814.1 in FASTA format as shown in Figure 4.1. There are 519 amino acids in the protein sequence which then was subjected to prediction analysis of B-cells and T-cells epitopes.


```
>AMR44593.1 hypothetical protein [Tilapia lake virus]
MWAFQEGVCKGNLLSGPTSMKAPDSAARESLDRASEIMTGKSYNAVHTGDLSKLPNQGESPLRIVDSDLY
SERSCCWVIEKEGRVVCKSTTLTRGMTGLLNTTRCSSPSELICKVLTVESLSEKIGDTSVEELLSHGRYF
KCALRDQERGKPKSRAIFLSHPFFRLLSSVVETHARSVLSKVSAVYTATASAEQRAMMAAQVVESRKHVL
NGDCTKYNEAIDADTLLKVWDAIGMGSIGVMLAYMVRRKCVLIKDTLVECPGGMLMGMFNATATLALQGT
TDRFLSFSDDFITSFNSPAELREIEDLLFASCHNLSLKKSYISVASLEINSCTLTRDGDLATGLGCTAGV
PFRGPLVTLKQTAAMLSGAVDSGVMPFHSAERLFQIKQQECAYRYNNPTYTTRNEDFLPTCLGGKTVISF
QSLLTWDCHPFWYQVHPDGPDTIDQKVLSVLASKTRRRRTRLEALSDLDPLVPHRLLVSESDVSKIRAAR
QAHLKSLGLEQPTNFNYAIYKAVQPTAGC
```

Figure 4.1 The amino acid sequence of Segment 1 of TiLV retrieved from GenBank. [Acc.No. KU751814.1]

Recently, Pulido et al., 2019 proposed two genetic clades of TiLV which are Israeli and Thai clades. The proposal was made based on multilocus sequence phylogenetic analysis (MLSA) of a total of 8305 nucleotides from five TiLV genomes. Due to limited database of complete TiLV genomes available in the GenBank, another previous was conducted to investigate genetic diversity of TiLV isolates based on the open reading frame (ORF) of segment 1 PB1 gene (1560 nucleotides) obtained from 21 TiLV isolates. These isolates were obtained from Israel, Peru, Ecuador, and Thailand. The study had revealed comparative findings on nucleotide sequences in Segment 1of TiLV, which relatively conserved with 95–99.9% identity among all isolates used. The identiy of the sequence also is alike Orthomyxoviruses, a virus RNA polymerase PB1 (97.6–99.8%) and Influenza C virus polymerase PB1 (97.1–99.3%). Therefore, the decision to utilize Segment 1 of TiLV in the present study shall be a promising gene for prediction of target candidate for development of peptide-based therapeutic of TilV disease.

4.2 The B-Cells Epitope within Segment 1 of TiLV

Kolaskar and Tongaonkar antigenicity scale tool was used in this study to predict B-cells epitope based on 519 amino acids sequence retrieved from the Genbank. From the result, a total of 17 sequences of linear B-cells epitopes with varying lengths between 6 – 36 amino acids were identified (Table 4.1). The epitope regions prediction through Kolaskar and Tonganonkar antigenicity tools were presented in Figure 4.2.

Table 4.1 Results on B Cell Epitope Using Kolaskar and Tongaonkar

No.	Start	End	Peptide	Length
1	5	12	QEGVCKGN	8
2	61	79	PLRIVDSDLYSERSCCWVI	19
3	83	90	GRVVCKST	8
4	105	123	CSSPSELICKVLTVESLSE	19
5	130	145	VEELLSHGRYFKCALR	16
6	154	189	SRAIFLSHPFFRLLSSVVETHARSVLSKVSAVYTAT	36
7	199	213	AAQVVESRKHVLNGD	15
8	223	232	ADTLLKVWDA	10
9	239	263	GVMLAYMVRRKCVLIKDTLVECPGG	25
10	272	278	TATLALQ	7

11	284	289	FLSFSD	6
12	304	335	IEDLLFASCHNLSLKKSYISVASLEINSCTLT	32
13	341	394	ATGLGCTAGVPFRGPLVTLKQTAAMLSGAVDSGVM PFHSAERLFQIKQQECAYR	54
14	408	438	LPTCLGGKTVISFQSLLTWDCHPFWYQVHPD	31
15	444	453	DQKVLSVLAS	10
16	462	486	LEALSDLDPLVPHRLLVSESDVSKI	25
17	489	500	ARQAHLKSLGLE	12

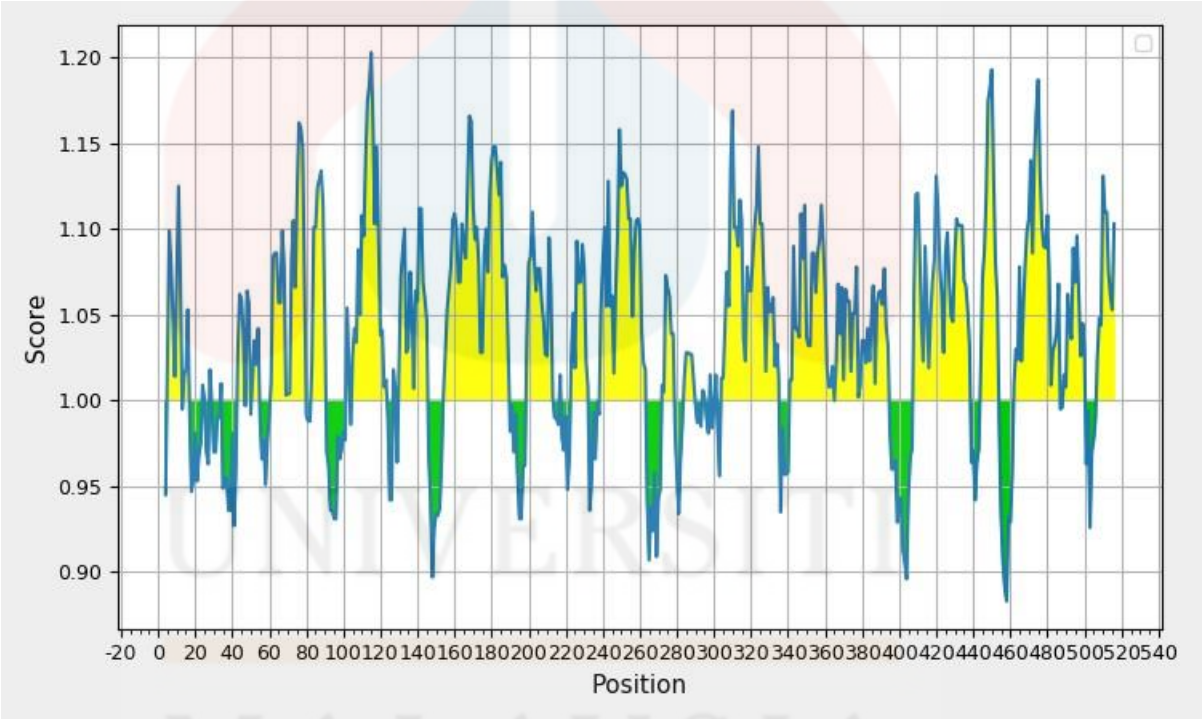


Figure 4.2 Epitope regions using Kolaskar and Tangoankar antigenicity prediction method. The yellow-colored peaks represent the epitope regions (threshold score >1), while the green-colored slopes represent the non-epitope regions.

In the present study, B cell epitopes were predicted from the Segment 1 of TiLV which encodes for RdRP gene of virus genome. From the analysis, the potential epitopes that would be recognized by B-lymphocytes (B-cells) can be identified which important to initiate immune response during an infection. As target candidate of a vaccine, this epitope is used to be recognized by B-cells to enhance antibodies production prior infection and this epitopes must be linear and usually located on the surface of the antigen protein to be easily accessible by B-cells receptors. In addition, the candidate target of vaccine should demonstrate greater antigenicity to elicit antibodies production. Therefore, Kolaskar and Tongaonkar antigenicity scale from IEDB analysis resources were used to identify antigenic region of B-cells epitopes within Segment 1 of TiLV.

4.3 The T-Cells Epitopes within Segment 1 of TiLV.

The epitope regions of B-cells obtained in subsection 4.2 were further modulated through Proped and Proped I server simultaneously with the purpose to predict T-cells epitopes. From Proped and Proped I analysis, 17 epitopes obtained earlier were further analysed for T-cell epitope which involve prediction of epitope recognition by MHC II and MHC I binding region that crucial for stimulating the immune response towards

antigen epitopes. From the analysis, there are 8 epitopes were recognized by MHC class I while 29 epitopes able to interacted with MHC class II binding allelle (Table 4.2).

Table 4.2: MHC binding region in Propred server.

Propred-1 server (Binding allele of MHC I)			**Propred server (Binding allele of MHC II)**		
Epitope	Percentage (%)	Amino acid Position	Epitope	Percentage (%)	Amino acid position
ARSVLSKVS	4.26	175-182	LRIVDSDLY	27.45	62-70
VETHARSVL	19.15	171-179	LYSERSCCW	23.92	69-77
KSYISVASL	44.68	319-327	ICKVLTVES	11.76	112-120
CHNLSLKKS	2.13	312-320	VLTVESLSE	7.84	115-123
PLVTLKQTA	2.13	355-363	LICKVLTVE	1.96	111-119
GPLVTLKQT	10.64	354-362	LLSHGRYFK	19.61	133-141
PFRGPLVTL	4.26	351-359	IFLSHPFFR	15.69	157-165
VPHRLLVSE	6.38	472-480	FLSHPFFRL	5.88	158-166
			FFRLLSSVV	21.57	163-171
			FRLLSSVVE	58.82	164-172
			LLSSVVETH	13.73	166-174
			VVETHARSV	1.96	170-178
			VLSKVSAVY	7.84	178-186
			VVESRKHVL	3.92	202-210
			MVRRKCVLI	47.06	245-253
			VRRKCVLIK	72.55	246-254
			YMVRRKCVL	27.45	244-252
			LIKDTLVEC	15.69	252-260
			MLAYMVRRK	33.33	241-249
			LAYMVRRKC	37.25	242-250
			VLIKDTLVE	52.94	251-259
			FQSLLTWDC	49.02	420-428
			VISFQSLLT	19.61	417-425
			LLTWDCHPF	7.84	423-429
			LGGKTVISF	21.57	412-420
			LLVSESDVS	15.69	476-484
			LVPHRLLVS	72.55	471-479
			LVSESDVSK	1.96	477-485
			VPHRLLVSE	1.96	472-480

The main purpose to conduct this analysis is to identify epitope that could be recognized by both MHC I and MHC II molecules. However, only one of the epitope is predicted to be recognized by both MHC I and II which is “VPHRLLVSE”. Regardless to its promising identity, all epitopes listed in Table 4.2 were selected for further analysis to predict their binding affinity towards MHC molecules.

4.4 The Peptide Binding Affinity towards MHC I and II Protein.

To strengthen the antigenic protein prediction, selected epitopes were analysed through MHC Pred 2.0 using the most common MHC models which are MHC Class I (HLA-A 020, HLA-A 1101 and HLA-B 3501) and MHC Class II (HLA DRB 0101). This analysis is important to predict the binding affinity between the selected epitopes to ensure that these epitope would have strong or intermediate interaction with MHC molecule. The epitope which recognized as non-binder will be excluded from further analysis.

Table 4.3: Binding affinity between epitope and HLA Class I and Class II Allele

No.	Epitope	HLA Class I Allele			HLA Class II
		HLA-A 0201	**HLA A -A 1101**	**HLA - B 3501**	**HLA DRB 0101**
1.	ARSVLSKVS	Non - binder	30.55	448.75	46.34
2.	VETHARSVL	3845.92	88.10	520.00	66.22
3.	KSYISVASL	148.59	63.53	358.10	77.80
4.	CHNLSLKKS	4236.43	42.66	452.90	86.90
5.	PLVTLKQTA	905.73	39.54	534.56	12.05
6.	GPLVTLKQT	Non - binder	41.78	696.63	123.88
7.	PFRGPLVTL	746.45	38.82	755.09	59.98
8.	VPHRLLVSE	Non - binder	39.45	563.64	208.93
9.	LRIVDSDLY	1499.68	164.06	462.38	46.34
10.	LYSERSCCW	Non - binder	140.28	398.11	66.22
11.	ICKVLTVES	576.77	65.92	654.64	77.80
12.	VLTVESLSE	638.26	130.62	683.91	86.90
13.	LICKVLTVE	1250.26	110.66	1013.91	12.05
14.	LLSHGRYFK	941.89	173.78	843.33	123.88
15.	IFLSHPFFR	107.89	46.99	521.19	59.98
16.	FLSHPFFRL	165.96	32.66	132.74	208.93
17.	FFRLLSSVV	824.14	35.73	239.33	46.34
18.	FRLLSSVVE	1013.91	57.68	171.40	66.22
19.	LLSSVVETH	248.31	136.46	2182.73	77.80

20.	VVETHARSV	1076.47	40.64	1000.00	86.90
21.	VLSKVSAVY	703.07	123.31	827.94	12.05
22.	VVESRKHVL	Non - binder	60.67	602.56	123.88
23.	MVRRKCVLI	1253.14	15.60	781.63	59.98
24.	VRRKCVLIK	1432.19	40.27	829.85	208.93
25.	YMVRRKCVL	3083.19	61.09	328.10	46.34
26.	LIKDTLVEC	424.62	91.20	792.50	66.22
27.	MLAYMVRRK	699.84	66.22	548.28	77.80
28.	LAYMVRRKC	2570.40	123.31	682.34	86.90
29.	VLIKDTLVE	92.68	136.14	877.00	12.05
30.	FQSLLTWDC	Non - binder	48.98	190.99	12.05
31.	VISFQSLLT	Non - binder	21.53	1028.02	123.88
32.	LLTWDCHPF	242.66	287.08	373.25	59.98
33.	LGGKTVISF	Non - binder	109.65	701.46	208.93
34.	LLVSESDVS	3404.08	127.35	796.16	46.34
35.	LVPHRLLVS	Non - binder	179.89	477.53	66.22
36.	LVSESDVSK	2133.04	29.04	623.73	77.80

From the findings shown in Table 4.3, there are nine epitopes that were identified as non-binder to HLA-A 0201. Although these epitopes can be recognized by B-cells or T-cells, the weak interaction that was predicted through their binding affinity suggested that they are not a suitable candidate of antigenic peptide. Thus, these epitopes will be excluded form the final analysis of antigenicity and allerginicity.

4.5 Antigenicity and Allergenicity Epitopes Prediction.

Further analysis was performed to determine the epitopes' antigenicity and allergenicity profile. Antigenicity test was conducted using Vaxijen 2.0 with threshold of 0.4. Therefore, any epitope that showed antigenic score above or equal to 0.4 is considered as antigen and most likely a good antigenic peptide within the Segment 1 of TiLV. From the antigenicity analysis, 15 peptides out of 28 peptides were identified as antigen (Table 4.4).

Table 4.4: Allergenicity and antigenicity of epitope

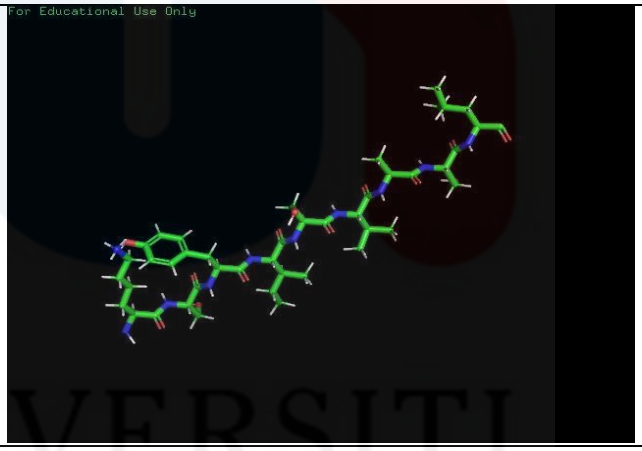
No.	Epitope	Antigen / Non - Antigen	Allergen / Non - Allergen	Nearest Allergen
1.	VETHARSVL	Non - antigen	Non - allergen	
2.	KSYISVASL	Antigen	Non - allergen	
3.	CHNLSLKKS	Antigen	Allergen	*Oryza sativa Japonica Group* (Japanese rice)
4.	PLVTLKQTA	Antigen	Allergen	*Uquus cabbalus* (horse)
5.	PFRGPLVTL	Non - antigen	Non - allergen	
6.	LRIVDSDLY	Antigen	Allergen	*Mus musculus* (house mouse)

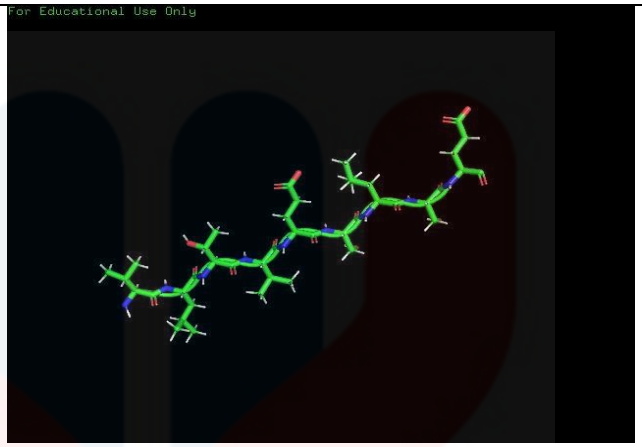
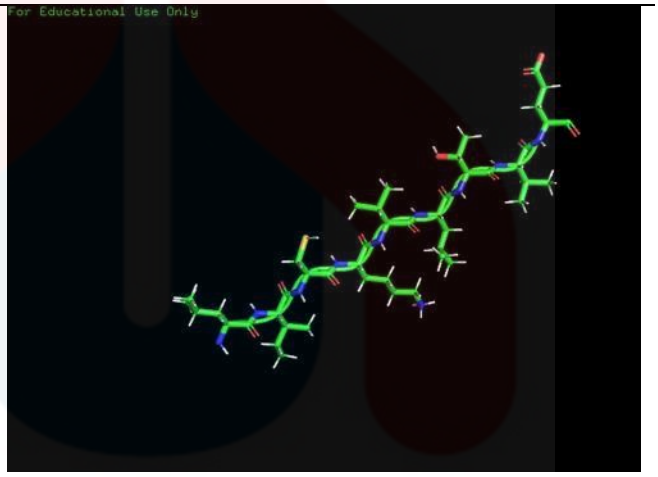
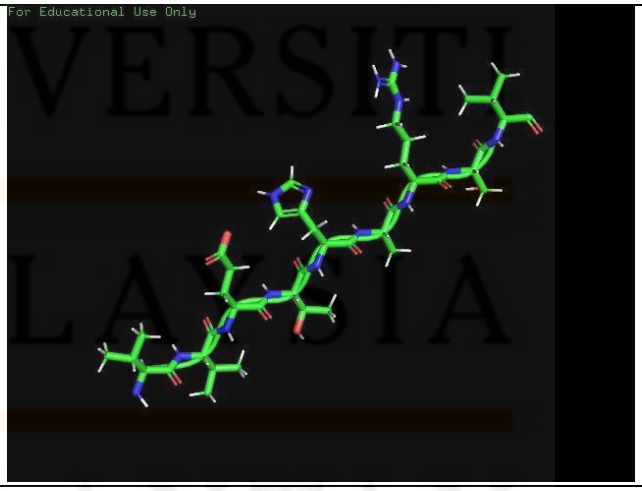
7.	ICKVLTVES	Non - antigen	Allergen	*Dermatophagoides pteronyssinus* (European house dust mite)
8.	VLTVESLSE	Antigen	Non - allergen	
9.	LICKVLTVE	Antigen	Non - allergen	
10.	LLSHGRYFK	Non - antigen	Non - allergen	
11.	IFLSHPFFR	Non - antigen	Non - allergen	
12.	FLSHPFFRL	Non - antigen	Non - allergen	
13.	FFRLLSSVV	Non - antigen	Non - allergen	
14.	FRLLSSVVE	Non - antigen	Non - allergen	
15.	LLSSVVETH	Non - antigen	Allergen	*Felis catus* (domestic cat)
16.	VVETHARSV	Antigen	Non - allergen	
17.	VLSKVSAVY	Antigen	Allergen	*Chironomus thummi thummi* (Midge)
18.	MVRRKCVLI	Antigen	Non - allergen	
19.	VRRKCVLIK	Antigen	Non - allergen	
20.	YMVRRKCVL	Antigen	Non - allergen	
21.	LIKDTLVEC	Non - antigen	Non - allergen	
22.	MLAYMVRRK	Antigen	Non - allergen	
23.	LAYMVRRKC	Antigen	Non - allergen	
24.	VLIKDTLVE	Non - antigen	Non - allergen	
25.	LLTWDCHPF	Antigen	Allergen	*Triticum aestivum* (Wheat)
26.	LLVSESDVS	Antigen	Non - allergen	
27.	LVSESDVSK	Non - antigen	Non - allergen	
28.	VPHRLLVSE	Non - antigen	Allergen	*Actinidia deliciosa* (Kiwi)

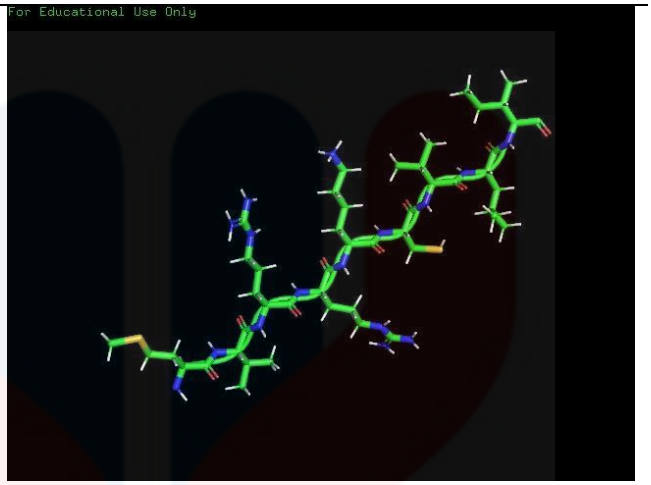
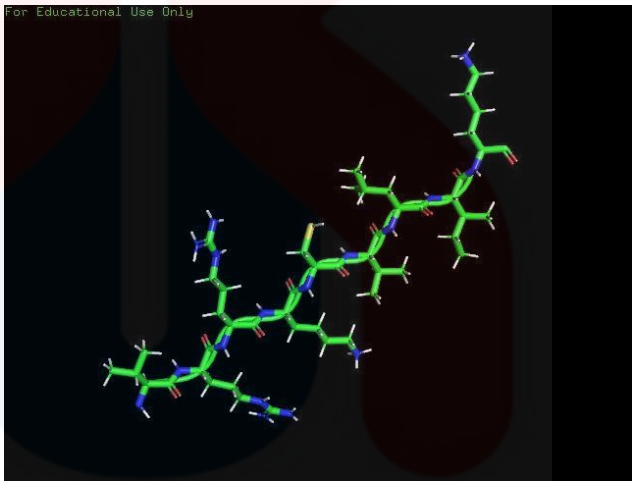
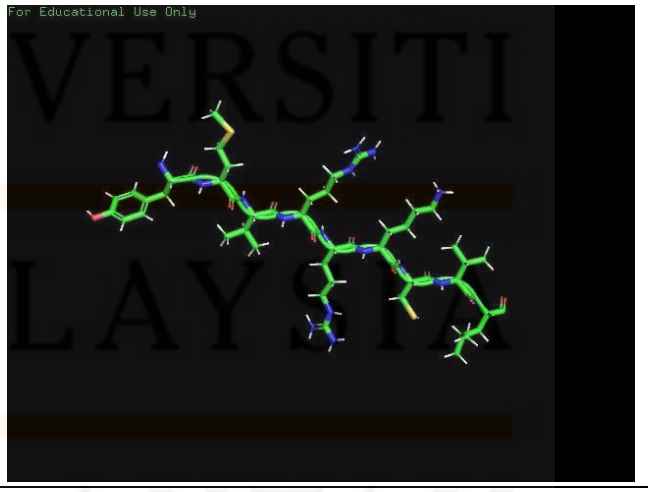
Meanwhile, allergenicity test was also being performed to identify non-allergen epitopes through AllerTop Test. Finally, a non-allergenic peptides that shown to be antigenicity are selected as candidate peptides which is the more likeable to choose to be also considered as safe and efficient candidate of vaccines against TiLV. In this

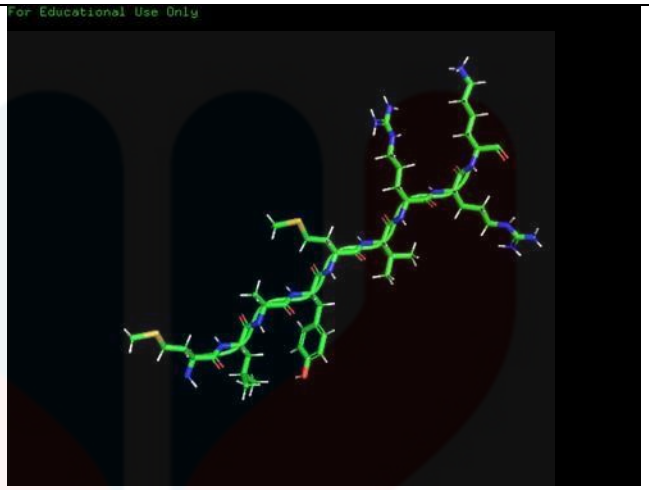
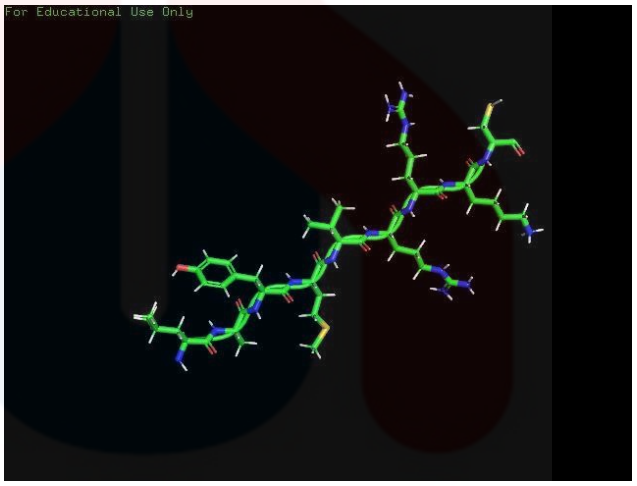
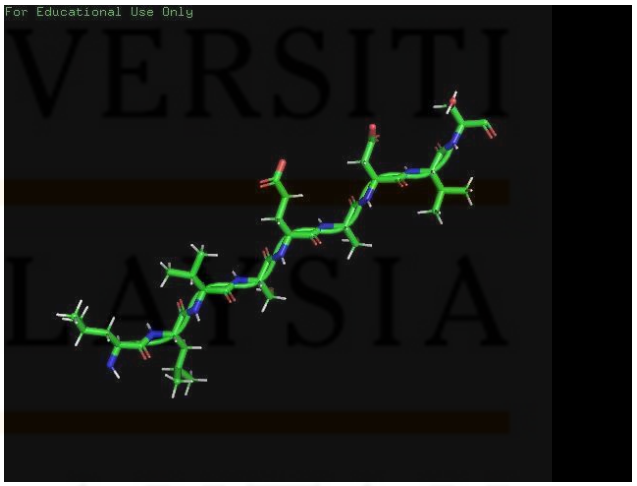
study, there were a total of 15 antigenic peptides that has been identified. These antigenic peptides were subjected to allergenicity test which revealed 10 out of 15 peptides were having non-allergen profile. These epitopes were potentionally safe for further studies. The present study has successfully identified 10 antigenic and non-allergen peptides as potential target candidate that can be used in development of peptide-based therapeutics of TiLV. These peptides were visualized in Pymol software and listed in Table 4.5.

Table 4.5: Visualization of potential target candidate of antigenic peptides

No.	Epitope	Visualization of Potential Epitope
1.	KSYISVASL	

2.	VLTVESLSE	
3.	LICKVLTVE	
4.	VVETHARSV	

5.	MVRRKCVLI	
6.	VRRKCVLIK	
7.	YMVRRKCVL	

8.	MLAYMVRRK	
9.	LAYMVRRKC	
10.	LLVSESDVS	

CHAPTER 6

CONCLUSION AND RECCOMENDATION

As conclusion, through a series of *in-silico* immunoinformatics approach, the study has successfully predicted and identified 36 potential epitopes of B-cells and T-cells. However, by assessment of their binding affinity with MHC molecules, only 25 epitopes displayed as having good and moderate interaction with MHC molecules. From here, further analysis was done to evaluate the epitopes antigenicity and allegenicity profile. A good target candidate that would be used as candidate vaccine or other application of peptide-based approach must be identified as antigen and having a non-allergen profile for safety handling by human. Based on the prediction, only 10 antigenic peptides were determined as target candidate of Segment 1 of TiLV which shall be tested in near future for its effectiveness through experimental approach. This study highlights the advantages of computational approach that shall provide rapid discovery at low cost of target antigen candidate.

As a recommendation, this study needs to proceed to predict the molecular docking to reveal the stability of vaccine candidate. Researchers also need to conduct laboratory test to ensure the safety of the vaccine to be consumed especially the effects of this vaccine to human. The effectiveness of vaccine to lower the risk of affected by TiLV in tilapia also needs to be monitored because virus can mutate.

REFERENCES

Abdullah, A., Ramly, R., Ridzwan, M. M., Sudirwan, F., Abas, A., Ahmad, K., Murni, M., & Kua, B. C. (2018). First detection of tilapia lake virus (TiLV) in wild river carp (Barbonymus schwanenfeldii) at Timah Tasoh Lake, Malaysia. *J. Fish Dis, 41*(9), 1459-1462.

Ahasan, M. S., Keleher, W., Giray, C., Perry, B., Surachetpong, W., Nicholson, P., Al-Hussinee, L., Subramaniam, K., & Waltzek, T. B. (2020). Genomic characterization of tilapia lake virus isolates recovered frommoribund Nile tilapia (Oreochromis niloticus) on a farm in the United States. *Microbiology resource announcements, 9*(4), e01368-01319.

Al-Hussinee, L., Subramaniam, K., Surachetpong, W., Popov, V., Hartman, K., Starzel, K., Yanong, R., Watson, C., Ferguson, H., & Frasca Jr, S. (2019). Tilapia Lake Virus (TiLV): a Globally Emerging Threat to Tilapia Aquaculture: FA213, 4/2019. *EDIS, 2019*(2).

Amal, M., Koh, C., Nurliyana, M., Suhaiba, M., Nor-Amalina, Z., Santha, S., Diyana-Nadhirah, K., Yusof, M., Ina-Salwany, M., & Zamri-Saad, M. (2018). A case of natural co-infection of Tilapia Lake Virus and Aeromonas veronii in a Malaysian red hybrid tilapia (Oreochromis niloticus× O. mossambicus) farm experiencing high mortality. *Aquaculture, 485*, 12-16.

Bacharach, E., Mishra, N., Briese, T., Zody, M. C., Kembou Tsofack, J. E., Zamostiano, R., Berkowitz, A., Ng, J., Nitido, A., & Corvelo, A. (2016). Characterization of a novel orthomyxo-like virus causing mass die-offs of tilapia. *MBio, 7*(2), e00431-00416.

Basri, L., Nor, R. M., Salleh, A., Saad, M. Z., Barkham, T., & Amal, M. N. A. (2020). Co-Infections of Tilapia Lake Virus, Aeromonas hydrophila and Streptococcus agalactiae in Farmed Red Hybrid Tilapia. *Animals, 10*(11), 2141.

Behera, B., Pradhan, P., Swaminathan, T., Sood, N., Paria, P., Das, A., Verma, D., Kumar, R., Yadav, M., & Dev, A. (2018). Emergence of tilapia lake virus associated with mortalities of farmed Nile tilapia Oreochromis niloticus (Linnaeus 1758) in India. *Aquaculture, 484*, 168-174.

Bibi, S., Ullah, I., Zhu, B., Adnan, M., Liaqat, R., Kong, W.-B., & Niu, S. (2021). In silico analysis of epitope-based vaccine candidate against tuberculosis using reverse vaccinology. *Scientific Reports, 11*(1), 1-16.

Chaput, D. L., Bass, D., Alam, M., Al Hasan, N., Stentiford, G. D., Van Aerle, R., Moore, K., Bignell, J. P., Haque, M. M., & Tyler, C. R. (2020). The segment matters: Probable reassortment of tilapia lake virus (TiLV) complicates phylogenetic analysis and inference of geographical origin of new isolate from Bangladesh. *Viruses, 12*(3), 258.

Contreras, H., Vallejo, A., Mattar, S., Ruiz, L., Guzmán, C., & Calderón, A. (2021). First report of tilapia lake virus emergence in fish farms in the department of Córdoba, Colombia. *Veterinary World, 14*(4), 865.

Foroutan, M., Ghaffarifar, F., Sharifi, Z., Dalimi, A., & Pirestani, M. (2018). Bioinformatics analysis of ROP8 protein to improve vaccine design against Toxoplasma gondii. *Infection, Genetics and Evolution, 62*, 193-204.

Gonçalves, R. A., Schatzmayr, D., Albalat, A., & Mackenzie, S. (2020). Mycotoxins in aquaculture: feed and food. *Reviews in Aquaculture, 12*(1), 145-175.

James, N. (2019). Mass tilapia farming: real potential in the right areas. *Farmer's Weekly, 2019*(19011), 29-29.

Jansen, M. D., Dong, H. T., & Mohan, C. V. (2019). Tilapia lake virus: a threat to the global tilapia industry? *Reviews in Aquaculture, 11*(3), 725-739.

Khay, S., Joshi, R. C., & Sastroutomo, S. S. (2018). Invasive apple snails: integrated management in lowland ricefields of Cambodia and probing their utilization in aquaculture. *Fish for the People, 16*(3), 34-37.

Magouz, F., Salem, M., & Hashad, M. (2018). Effect of some mycotoxin on growth performance and feed utilization of Nile tilapia (Oreochromis niloticus). *Iraqi Journal of Veterinary Sciences, 32*(1), 99-108.

Mai, T. T., Kayansamruaj, P., Soontara, C., Kerddee, P., Nguyen, D.-H., Senapin, S., Costa, J. Z., del-Pozo, J., Thompson, K. D., & Rodkhum, C. (2022). Immunization of Nile tilapia (Oreochromis niloticus) broodstock with tilapia lake virus (TiLV) inactivated vaccines elicits protective antibody and passive maternal antibody transfer. *Vaccines, 10*(2), 167.

Michel-Todó, L., Bigey, P., Reche, P. A., Pinazo, M.-J., Gascón, J., & Alonso-Padilla,

J. (2020). Design of an epitope-based vaccine ensemble for animal trypanosomiasis by computational methods. *Vaccines*, *8*(1), 130.

Mohamad, S. N., Noordin, W. N. M., Ismail, N. F., & Hamzah, A. (2021). Red Hybrid Tilapia (*Oreochromis spp*.) Broodstock Development Programme in Malaysia: Status, Challenges and Prospects for Future Development. *Asian Fisheries Science*, *34*, 73-81.

Mugimba, K. K., Byarugaba, D. K., Mutoloki, S., Evensen, Ø., & Munang'andu, H. M. (2021). Challenges and solutions to viral diseases of finfish in marine aquaculture. *Pathogens*, *10*(6), 673.

Nicholson, P., Fathi, M., Fischer, A., Mohan, C., Schieck, E., Mishra, N., Heinimann, A., Frey, J., Wieland, B., & Jores, J. (2017). Detection of Tilapia Lake Virus in Egyptian fish farms experiencing high mortalities in 2015.

Nilav Aicha, A. P., Tanmoy Gon Choudhuryc, Himadri Saha. (2020). Tilapia Lake Virus (TiLV) disease: Current status of understanding *Aquacultures and Fisheries, Volume 6*.

Oli, A. N., Obialor, W. O., Ifeanyichukwu, M. O., Odimegwu, D. C., Okoyeh, J. N., Emechebe, G. O., Adejumo, S. A., & Ibeanu, G. C. (2020). Immunoinformatics and vaccine development: an overview. *ImmunoTargets and therapy*, *9*, 13.

Prabu, E., Rajagopalsamy, C., Ahilan, B., Jeevagan, I. J. M. A., & Renuhadevi, M. (2019). Tilapia–an excellent candidate species for world aquaculture: a review. *Annual Research & Review in Biology*, 1-14.

Raoufi, E., Hemmati, M., Eftekhari, S., Khaksaran, K., Mahmodi, Z., Farajollahi, M. M., & Mohsenzadegan, M. (2020). Epitope prediction by novel immunoinformatics approach: a state-of-the-art review. *International Journal of Peptide Research and Therapeutics*, *26*(2), 1155-1163.

Reantaso, M. B. (2017). Tilapia lake virus (TiLV) caused by an orthomyxo-like virus (Family Orthomyxoviridae) threatening cultured and wild stocks of tilapia. *FAO Aquaculture Newsletter*(57), 9-11.

Richard Kabuusu, A. A., Stroup DF,Macpherson CNL and Ferguson HW. (2017). Production-level risk factors for syncytial hepatitis in farmed tilapia (Oreochromis niloticus L). *Journal of Fishes Diseases, Volume 6*.

Soria-Guerra, R. E., Nieto-Gomez, R., Govea-Alonso, D. O., & Rosales-Mendoza, S. (2015). An overview of bioinformatics tools for epitope prediction: implications

on vaccine development. *Journal of biomedical informatics, 53*, 405-414.

Surachetpong, W., Roy, S. R. K., & Nicholson, P. (2020). Tilapia lake virus: The story so far. *Journal of Fish Diseases, 43*(10), 1115-1132.

Thawornwattana, Y., Dong, H. T., Phiwsaiya, K., Sangsuriya, P., Senapin, S., & Aiewsakun, P. (2021). Tilapia lake virus (TiLV): Genomic epidemiology and its early origin. *Transboundary and Emerging Diseases, 68*(2), 435-444.

Valladão, G., Alves, L., & Pilarski, F. (2016). Trichodiniasis in Nile tilapia hatcheries: diagnosis, parasite: host-stage relationship and treatment. *Aquaculture, 451*, 444-450.

Walker, P. J., Siddell, S. G., Lefkowitz, E. J., Mushegian, A. R., Dempsey, D. M., Dutilh, B. E., Harrach, B., Harrison, R. L., Hendrickson, R. C., & Junglen, S. (2019). Changes to virus taxonomy and the International Code of Virus Classification and Nomenclature ratified by the International Committee on Taxonomy of Viruses (2019). *Archives of virology, 164*(9), 2417-2429.

Yamkasem, J., Tattiyapong, P., Gorgoglione, B., & Surachetpong, W. (2021). Uncovering the first occurrence of Tilapia parvovirus in Thailand in tilapia during co-infection with Tilapia tilapinevirus. *Transboundary and Emerging Diseases*.

Zeng, W., Wang, Y., Chen, X., Wang, Q., Bergmann, S. M., Yang, Y., Wang, Y., Li, B., Lv, Y., & Li, H. (2021). Potency and efficacy of VP20-based vaccine against tilapia lake virus using different prime-boost vaccination regimens in tilapia. *Aquaculture, 539*, 736654.

Zhang, J., & Tao, A. (2015). Antigenicity, Immunogenicity, Allergenicity. In *Allergy Bioinformatics* (pp. 175-186). Springer.

APPENDICES

Result of the calculation of MHC I and II

	MHC Class I		MHC Class II	
	Peptide	Percentage (%)	Peptide	Percentage (%)
QEGVCKGN	Too small		Too small	
PLRIVDSDLY SERSCCWVI	Too small		LRIVDSDLY	14/51x100 = 27.45
			LYSERSCCW	2/51 x 100= 3.92
GRVVCKST	0		Too small	
CSSPSELICK VLTVESLSE	0		ICKVLTVES	6/5 1x 100 = 11.76
			VLTVESLSE	4/51 x 100 = 7.84
			LICKVLTVE	1/51 x 100 = 1.96
VEELLSHGR YFKCALR	0		LLSHGRYFK	10/51 x 100 = 19.61
SRAIFLSHPFF RLLSSVVETH	ARSVLSKVS	2/47 x 100= 4.26	IFLSHPFFR	8/51 x 100 = 15.69
ARSVLSKVS AVYTAT			FLSHPFFRL	3/51 x 100 = 5.88
	VETHARSVL	9/47x100 =19.15	FFRLLSSVV	11/51 x 100 = 21.57
			FRLLSSVVE	30/51 x 100 = 58.82
			LLSSVVETH	7/51 x 100 = 13.73
			VVETHARSV	1/51 x 100 = 1.96
			VLSKVSAVY	4/51 x 100 = 7.84
ADTLLKVWDA	0		0	

GVMLAYMVRRKCVLIKDTLVECPGG	0		MVRRKCVLI	24/51 x 100 = 47.06
			VRRKCVLIK	37/51 x 100 = 72.55
			YMVRRKCVL	14/51 x 100 = 27.45
			LIKDTLVEC	8/51 x 100 = 15.69
			MLAYMVRRK	17/ 51 x 100 = 33.33
			LAYMVRRKC	19/51 x 100 = 37.25
			VLIKDTLVE	27/51 x 100 = 52.94
TATLALQ	Too small		Too small	
FLSFSD	Too small		Too small	
IEDLLFASCHNLSLKKSYISVASLEINSCTLT	KSYISVASL	21/47x100 =44.68	LLFASCHNL	6/ 51 x 100 = 11.76
	CHNLSLKKS	1/47x100 =2.13	LFASCHNLS	4/51 x 100 = 7.84
			FASCHNLSL	3/51 x 100 = 5.88
			LSLKKSYIS	8/51 x 100 = 15.69
			LKKSYISVA	5 /51 x 100 = 9.8
			YISVASLEI	21/51 x 100 = 41.18
			VASLEINS	2/51 x 100 = 3.92
			LEINSCTLT	2/51 x 100 = 3.92

ATGLGCTAGVPFRGPLVTL	PLVTLKQTA	1/47x100 =2.13	IFLSHPFFR	8/51 x 100 = 15.69
KQTAAMLSGAVDSGVMPF	GPLVTLKQT	5/47x100= 10.64	FLSHPFFRL	3/51 x 100 = 5.88
HSAERLFQIKQQECAYR	PFRGPLVTL	2/47x100=4. 26	FFRLLSSVV	11/51 x 100 = 21.57
			FRLLSSVVE	30/51 x 100 = 58.82
			LLSSVVETH	7/51 x 100 = 13.73
			VVETHARSV	1/51 x 100 = 1.96
			VLSKVSAVY	4/51 x 100 = 7.84
LPTCLGGKTVISFQSLLTWDCHPFWYQVHPD	0		FQSLLTWDC	25/51 x 100 = 49.02
			VISFQSLLT	10/51 x 100 = 19.61
			LLTWDCHPF	4/51x 100 = 7.84
			LGGKTVISF	11/51 x 100 = 21.57
DQKVLSVLAS	0		0	
LEALSDLDPLVPHRLLVSESDVSKI	VPHRLLVSE	3/47x100=6. 383	LLVSESDVS	8/51 x 100 = 15.69
			LVPHRLLVS	37/51 x 100 = 72.55
			LVSESDVSK	1/51 x 100 = 1.96
			VPHRLLVSE	1/51 x 100 = 1.96
ARQAHLKSLGLE	0		0	