Immuno-informatics aided -Design of Multi-epitope Based Vaccine Against *Nipah henipavirus* Glycoprotein G.

By

Sabrina Sultana Moon
17346028

A thesis submitted to the School of Pharmacy in partial fulfillment of the requirements for the degree of
Bachelors' of Pharmacy

School of Pharmacy, Brac University
Summer 2021

Declaration

It is hereby declared that

1. The thesis submitted is our own original work while completing degree at Brac University.
2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.
3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.
4. We have acknowledged all main sources of help.

Student's Full Name & Signature:

Sabrina Sultana Moon
17346028

Approval

The thesis/project titled “Immuno-informatics aided -Design of Multi-epitope Based Vaccine Against Nipah henipavirus Glycoprotein G” submitted by Sabrina Sultana Moon (17346028) of Summer 2017 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of Bachelor’s of Pharmacy.

Examining Committee:

Supervisor:
(Member)

_____
Mohammad Kawsar Sharif Siam
Senior Lecturer, School of Pharmacy
Brac University

Program Coordinator:
(Member)

_____
Namara Mariam Chowdhury
Lecturer, School of Pharmacy
Brac University

Assistant Dean:
(Member)

_____
Dr. Hasina Yasmin
Assistant Dean and Program Director , School of Pharmacy
Brac University

Departmental Head:
(Dean)

_____
Professor Dr. Eva Rahman Kabir
Dean, School of Pharmacy
Brac University

Ethics Statement

In working on this thesis, there were no unethical practices followed. There are no tests conducted on either humans or animals in this study. This was carried out in a manner that adhered to all applicable ethical norms.

Abstract

In the family Paramyxoviridae, the genome of Nipah henipavirus belongs. The Henipa virus is a descendant of the Hendra and Nipah viruses. The most recent Nipah henipavirus outbreak in Kerala occurred in 2018 in the Kozhikode district. Since there are no effective antiviral drugs to treat NiV illness, patients have no other therapy choices. In this study, we developed an immunoinformatics-based multi-epitope vaccine to guard against Nipah henipavirus glycoprotein G infection. Vipr's database service provided the glycoprotein G sequence in the form of a FASTA file. After that, we used several servers to choose ctl, htl, B cell epitopes, and other predictions. We found two of the CTL epitopes here (AMDEGYFAY and GIKQGDTLY) and two of the htl epitopes here (NPLVVNWRDNTVISR and VNPLVVNWRDNTVIS) show the highest binding affinity with the glycoprotein G sequence. The antigenicity of the entire protein sequence rises from 0.52 to 0.55 when all of the antigenic epitopes are included. The Z score is positive (-8.47), the GRAVY score is positive (-0.316), which suggests that the protein is hydrophilic, the confidence level is 100%, and the coverage rate is 52%, considered a good result for the development of a vaccine.

Keywords: Immuno-informatics; immune simulation; Toll - like Receptor 3; multi-epitopes.

Dedication

Dedicated to my parents

Acknowledgement

This project might not have been possible for me to complete without the help of many people. A special thanks to my advisor, Mohammad Kawsar Sharif Siam of the Department of Pharmacy at Brac University, who reviewed my revisions and helped me make sense of them. I received assistance and guidance primarily due to Rian Rafsan, a student in MS of biotechnology at Brac University.

Sabrina Sultana Moon
ID: 17346028
Department of Pharmacy
Brac University.

Table of Contents

List of Tables

List of Figures

List of Acronyms

TLR3 Toll-like Receptor 3

NiV Nipah virus

HeV Hendra virus

IFN Interferon

IL4 Interleukin 4

DC Dendritic cell

CTL Cytotoxic T lymphocyte

HTL Helper T lymphocyte

Chapter 1

Introduction & literature review:

When it comes to the diseases that pose a significant risk to the general population's health, those that are caused by viruses are of utmost significance. The Nipah virus, which is an RNA virus and a member of the Paramyxoviridae family, was discovered in Malaysia in 1998/99 (Skowron et al., 2022) . It is divided as a member of the genus Henipavirus, which also contains the Hendra virus (HeV) and the Cedar virus, the latter of which was only recently identified (Aditi & Shariff, 2019), (Shariff, n.d.). In addition to its high fatality rate in humans, the World Health Organization (WHO) considers it a worldwide health hazard because it is zoonotic, can be spread among humans (Skowron et al., 2022). An infection with NiV is linked to encephalitis, which is an inflammation of the brain, and it can produce symptoms ranging from mild to severe sickness and even result in death (*Nipah Virus Infection - Causes, Symptoms, Treatments and Preventions | TRUTEST Laboratories*, n.d.). In some areas of Asia, most notably Bangladesh and India, there is a near-yearly occurrence of outbreaks (*Nipah Virus (NiV) | CDC*, n.d.). A phylogenetic study based on the complete sequences of the N and G genes confirmed the existence of two major NiV clades currently in circulation. Raj et al., in their article, state that the NiV-MY clade was formed by samples gathered in Malaysia and Cambodia, while the NiV-BD clade was formed by samples obtained in Bangladesh and India. NiV strains found in Thailand had a wide range of genetic material (Singh et al., 2019).

1.1 Nipah henipa virus’ structure and genome:

Nipah virus infection can be diagnosed in the laboratory using the reverse transcriptase polymerase chain reaction (RT- PCR), performed on urine, cerebrospinal fluid, throat swabs and blood samples (Bruhn et al., 2014). After recovering, the patient needs to do several tests for IgG and IgM antibodies to establish the presence of Nipah virus infection (Blocquel et al.,

2013). Immunohisto- chemistry on tissues taken following an autopsy can confirm the condition (Mohammed et al., 2020). The NiV genome is a single-stranded, negative-sense RNA of approximately 18.2 kilobases in size. This RNA encodes six structural proteins: nucleoprotein (N), phosphoprotein (P), matrix protein (M), fusion protein (F), and attachment glycoprotein (G). In addition, this RNA also encodes the large protein, also known as the RNA polymerase protein (L). Aside from that, the P gene is responsible for the encoding of three nonstructural proteins called V and W proteins by RNA editing or an alternate open reading frame (C protein) (Sun et al., 2018).

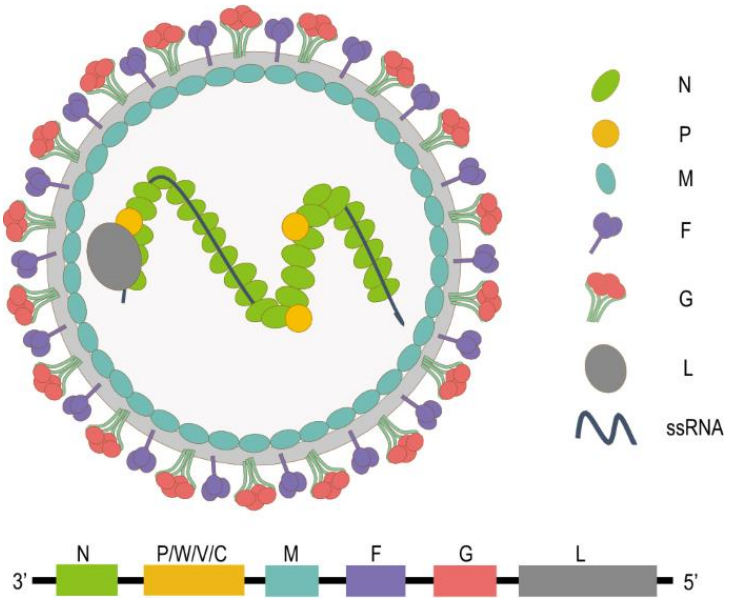


Figure 1 An illustration in the form of a schematic showing the genome's organisation and the virus's structure is located in the upper panel and lower panel. The utilisation of a wide range of hues indicates the manifestation of various genes or proteins.

Again because the genomic lengths of HeV and NiV are sufficiently distinct from those of other existing genera of paramyxoviruses and their protein homologies, the two viruses have been placed in a new genus. They stand out from other paramyxoviruses due to the vast range of hosts they can infect and the tremendous virulence they exhibit (Aguilar & Lee, 2011).

1.2 Replication cycle and pathogenesis of Nipah henipavirus:

In general, after the virus has bound to the host cell receptor, paramyxoviruses need the assistance of their specific bonding and fusion transmembrane glycoproteins in order to replicate inside the host cell (Maisner et al., 2018). Research on emerging paramyxoviruses, often known as henipaviruses, has received the most attention because of the comparatively increased mortality rates associated with these pathogens (Ringel et al., 2019). Viral life cycles have been studied in great detail, including host receptors, membrane fusion and viral entrance, replication and interferon (IFN) response (Bharaj et al., 2016). Each of these processes provides a possible target for creating antiviral medications that may be used in the treatment of patients (Aguilar & Lee, 2011).

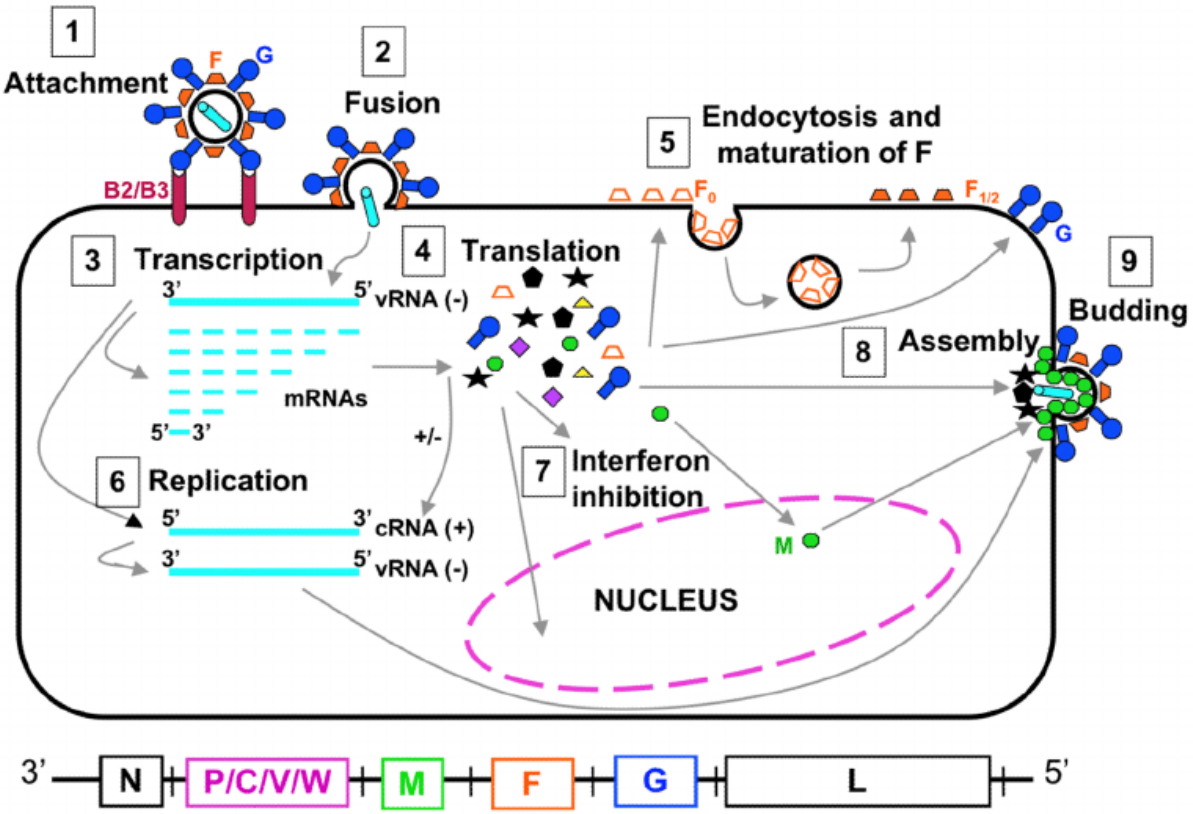


Figure 2 The virus enters the cell after attaching itself to the ephrinB2/B3 receptor (a) and then combining with the receptor (b). The negative RNA genome, also known as [vRNA(-)], acts as a guideline for the transcription of viral mRNAs, which occurs by an at attenuation gradient running from N to L in 3' to 5' length (c). At the very bottom of the picture is a representation of the henipavirus genomic RNA, oriented from 3' to 5'. On this RNA are the symbols N and L. mRNAs are converted into proteins (d), while at the same time, the vRNA(-) serves as a template for cRNA(+), which in response serves as a template for vRNA(-) genomes when they are replicated (e). New viral RNA genomes will be integrated into newly assembled viruses (f) during viral assembly. After translation (step d), several viral proteins participate in interferon (IFN) signalling pathways (step g), and the parent fusion protein (F0) will be endocytosed and developed (steps F1/2), respectively (h). The M (matrix) protein is principally responsible for orchestrating both the assembly (f) and budding I processes, and the N, P, C, M, F (fusion), and G (attachment) proteins are all included in the virus particles*(Aguilar & Lee, 2011)*.

Developing a vaccine for the Nipah virus is high on the WHO's priority list for research and development. This designation indicates that the virus is regarded as a category C agent for possible bioterrorism. A successful surveillance and detection system relies on having a better understanding of the virus's host range and data on resistance and pathogenesis in domestic livestock (Kasloff et al., 2019).

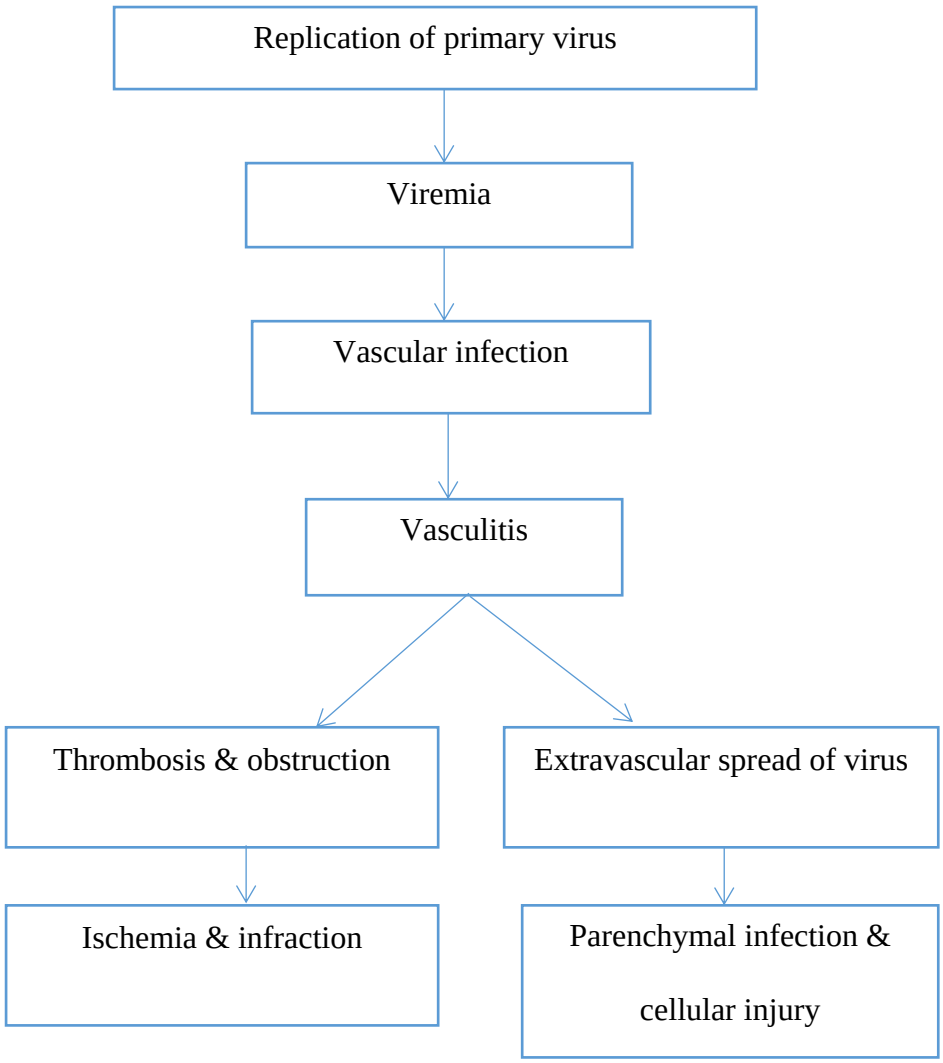


Figure 3 Nipah virus's role in the development of disease.

The structure of vasculitis and viral protein distribution indicates that endothelium infection occurs before transmural dissemination to primary smooth muscle in the tunica media. The observation shows that endothelial illness always comes first before transmural dissemination. A breakdown of the blood-brain barrier may allow the virus to move and extravascular CNS infection, as suggested by discovering many infected neurons in conjunction with necrotic plaques and vasculitic arteries. (Figure 3) (Raveendran et al., 2002).

Chapter 2

Materials and Method

2.1 Selection of Protein Sequence

Our team utilized the ViPR database tool for protein sequence selection. Virus Pathogen Database and Analysis Resource is the full name for what is known as ViPR. This repository serves as an integrated data and analytic storage for many viral families. It is supported by the Bioinformatics Resource Centers (BRC) initiative of the National Institute of Allergy and Infectious Diseases (NIAID) (NIAID) (Pickett, Sadat, et al., 2012). Through ViPR, we can locate information on the sequence of proteins, data regarding host factors, genes and proteins, genomic annotations, immunological epitopes (in addition to 3D structures), and other sorts of data (Pickett, Greer, et al., 2012).

2.2 Antigenicity identification

When creating a vaccine, the first thing scientists looked for was any antigenic components in the protein sequence. Antigenicity is the structure of antibody responses that a recipient creates in response to a virus. The recipient produces these antibody responses (Ye et al., 2020). Our team used the Vaxijen v2.0 server to make antigenicity predictions on the protein sequence, and we chose the threshold of 0.5. We kept the protein sequence in a format known as Fasta, and we have chosen the virus to act as the target organism. Without considering alignment, the Vaxijen service is the first to predict protective antigens of viral, bacterial, and neoplastic origin (Da Silva & Hughes, 1998).

2.3 Identification of Cytotoxic T lymphocyte Epitopes

We have used multiple epitopes for designing our vaccines like CTL, HTL and B cell epitopes. It is believed that cytotoxic T lymphocytes (CTLs) that identify particular viral peptides

(epitopes) offer the most efficient management over the reproduction and spread of viruses (Doytchinova & Flower, 2007).The portions of a protein known as epitopes or antigenic determinants have the potential to activate a cellular immune response mediated by T or B cells ("IMMUNOLOGY OF LAGOMORPHS," 1998). There are two T cell epitopes: CTL (Cytotoxic T lymphocyte) and HTL (helper T lymphocyte). Linear epitopes are the primary targets of the cytotoxic T cell-mediated response triggered by intracellular antigen processing ("IMMUNOLOGY OF LAGOMORPHS," 1998).

We considered several parameters for selecting CTL epitopes: antigenicity, toxicity, hydrophobicity, and solid binding capabilities, among other things, while determining which CTL epitopes to use. First, we decided to use the NetCtl 1.2 server to locate the peptide sequence that demonstrates effective binding with MHC (major histocompatibility complex) 1. Secondly, we used the NetMHC Pan 4.1 server to determine the strong bonds. Then for further filtration, we used the Toxinpred server to check the toxicity, and for antigenicity Vaxijen 2.0 server was used.

2.4 Helper T lymphocyte epitope Prediction:

The NetMHC II pan 4.0 server was the first application that we used to identify htl epitopes. We took twenty A1 alleles at a time, and all epitopes were retained while we set the binding affinity to 0.5x2 to find strong and weak binders. According to the anticipated score, the following is a list. Additional filtering of these sequences was done using the IFN pred, IL4 and IL10 pred servers. A total of just ten peptide sequences emerged from the filtration process, and they were found to be present in all sectors. Just two of these ten peptides were found to have antigenic characteristics, which we explored further. Consequently, we restricted the HTL epitope selection to these two peptide sequences.

2.5 B cell Prediction

To make predictions about B cells, we used the IEDB server. Epitopes on B cells and T cells can be determined with the help of this method (Vita et al., 2018). IEDB stands for the Immune Epitope Database, and it is a type of online interface that provides access to various epitope-related technologies that have been thoroughly described and validated (Soria-Guerra et al., 2015), (Olive et al., 2017). The protein sequence must first be entered in plain format for this service to function. The server then generates the B cell epitopes table, length score, and graph.

2.6 Antigenicity, Toxicity and Allergenecity Prediction of the Final Vaccine

We used Vaxijen 2.0 server to know the antigenicity of the sequence as before. Here we selected 0.5 threshold and applied the protein sequence.

We utilized the AllergenOnline program to perform the Allergenecity analysis. It shows that how many sequences are showing allergenecity with another substances and it also indicates the sequences. So that it will be helpful for us to develop the protein sequence and make better one.

After that, we chosen the T3DB server to determine whether or not the protein sequence was harmful.

2.7 Assessment of the Vaccine Using Biochemical techniques

Many factors can be used to evaluate a vaccine's stability, such as its molecular weight (under 100 kDa), whether or not the protein in it is hydrophilic or hydrophobic, and so on. We relied on the Protparam tool for all of our experiments.

2.8 Structural Analysis of the Vaccine

The Phyre 2 server was utilized in order to carry out the protein sequence structure analysis. The crystal three dimensional structure of the protein can be obtained from this server, which is also utilized for coverage and confidence analysis of the structure.

2.9 Assessment of 3D Modeled Structures

We were utilized two servers to assess the quality of the created models. They are, Prosa web, Swiss PDB Plotter.

Prosa web is used to determine whether or not a protein fulfils a set of predetermined energy requirements and analyzes the binding energies of each residue with the rest of the protein. A plot of it and the Z scores of empirically characterized structures in PDB show the latter's Z scores (Alexander et al., 1998).

To determine the Ramachandran plots of the structure, we utilized the Swiss PDB Plotter. The Ramachandran plot is a tool that can be utilized for the purpose of determining how well protein structures can be modeled. This server allowed for calculating the favoured number of the design, the amount of coverage that the system had, the number of bad bonds and wrong angles, the molprobity score, rotamer outliers, and a quality estimate of the structure.

2.10 Molecular Docking

We docked with the help of the patchdock server. Docking at the molecular level allows for the analysis of the interactions that take place between a ligand molecule and a receptor molecule, with the goal of determining the stability and interaction affinity of the complex that results (Silakari & Singh, 2021), (Stanzione et al., 2021). Here, we were required to input the PDB file of the protein and the name of the receptor to which the protein binds. As a receptor we used TLR3 protein, TLRs can be found not only in immune cells, such as macrophages and

dendritic cells (DCs), but also in non-immune cells, such as fibroblast cells and epithelial cells. TLRs play an important role in the innate immune system (Ahmadi et al., 2020). The usual location of TLR3 is in endosomes, and it is there that its luminal ectodomain (ECD) comes into contact with its ligands, the most important of which being dsRNA (Kawasaki & Kawai, 2014). Here we used 1ziw as TLR3 receptor which is a single chain and contains sequence from human (Zheng et al., 2021). The method of molecular docking known as PatchDock is one in which geometry plays a central role. This procedure aims to identify docking alterations that lead to a good shape complementarity between the molecules (Bell et al., 2006).

2.11 Dose Analysis of the Vaccine

We used the C-ImmSim server to analyze how the vaccination improved after it was administered (F Castiglione & Bernaschi, 2004). The C-ImmSim model represents, at the cellular level, the humoral reaction as well as the cytokine production of the immune system of a mammalian organism to the existence of antigens (such as viruses, bacteria, and so on), (mesoscopic scale) (Schneidman-Duhovny et al., 2005). It is capable of digesting peptides and displaying them on MHC class I and class II molecules, respectively (exogenous and endogenous pathways) (Filippo Castiglione, n.d.).The end result of this process is that some cells significantly extend their half-lives and live for a significantly longer period of time than other cells. In this study, we added the protein sequence in fasta format and administered three injection doses with a gap of 28 days between each one.

1 time–step = 8 hours = Day 1

84 time-steps = 672 hours = Day 28

168 time-steps= 1344 hours= Day 56

The periods we used were 1, 84, and 168. The graph of the provided antigens is displayed on this server, along with an indication of whether or not they are operating correctly.

Chapter 3

Results

3.1 Protein Sequence and its Antigenicity

The first thing that we did was search the family of the virus, which is called the Paramyxoviridae. Within this family, we chose the Orthoparamyxovirinae as a subfamily. After that, we decided on the genus henipavirus, which has another seven species. Finally, we decided on the Nipah henipavirus from among these seven species. This species of the Nipah-Henipa virus has a total of 168 genomes. We were able to determine the protein sequence by analyzing all of these genomes, and the sequence of the glycoprotein G protein was selected as the protein sequence. The protein sequence antigenicity score is 0.5215, indicating that it is likely antigenic.

3.2 CTL Epitopes Selection

From NetCtl 1.2 server we have found twelve CTL epitopes those showing efficient binding with MHC I alleles. They are:

PSKVIKSYY,STDNQAMIK,STASINENV,QTEGVSNLV,AMDEGYFAY,TVYHCSAVY,

CSAVYNNEF,SAVYNNEFY,AVYNNEFYY,GIKQGDTLY,LSDEENSKI,DSLGQPVFY.

After that, we used the NetMHC Pan 4.1 server to determine which of these 12 peptide sequences showed a strong connection with MHC I alleles. These steps were done to move on to the next stage of the filtration process. Our team discovered only five peptide sequences demonstrating a significant binding capacity with diverse alleles using this method.

Peptide Sequence	Alleles Covered	%Rank EL	%Rank BA
AMDEGYFAY	HLA-A*01:01	0.035	0.038
AVYNNEFYY	HLA-A*01:01	0.368	0.567
DSLGQPVFY	HLA-A*01:01	0.207	0.484
GIKQGDTLY	HLA-B*15:01	0.051	0.546
TVYHCSAVY	HLA-A*26:01	0.166	0.106

Table 1: MHC I alleles specific to CTL epitope shows strong binding affinity.

Then, we used a technology called Toxinpred to determine whether or not those five peptide sequences were poisonous. None of them had any toxic properties.

Query Peptides

Peptide ID	Peptide Sequence	SVM Score	Prediction	Hydrophobicity	Hydropathicity	Hydrophilicity	Charge	Mol wt
seq1	AMDEGYFAY	-0.83	Non-Toxin	0.03	-0.19	-0.38	-2.00	1066.26
seq2	AVYNNEFYY	-1.21	Non-Toxin	-0.05	-0.62	-0.89	-1.00	1182.37
seq3	DSLGQPVFY	-1.31	Non-Toxin	0.01	-0.03	-0.51	-1.00	1025.25
seq4	GIKQGDTLY	-0.83	Non-Toxin	-0.12	-0.60	-0.01	0.00	994.25
seq5	TVYHCSAVY	-0.78	Non-Toxin	0.06	0.60	-1.08	0.50	1042.29

1/1 50

Figure 4 Toxicity result of CTL epitopes no peptide sequence were found toxic.

Lastly, we used the Vaxijen 2.0 server with a 0.5 threshold to assess their antigenicity of them, and we discovered that just two of them were antigenic. Therefore, we decided to use this two-peptide sequence as the final CTL epitope. They have the values of AMDEGYFAY and GIKQGDTLY.

3.3 HTL epitopes

From NetMHC II Pan 4.0 server, we found twenty-five peptide sequences with high critical potential. The interferon-gamma pred tool filtered these twenty-five peptide sequences one more time for further filtration. Peptide sequences including these 19 peptides yielded positive

results. Then for further filtration, we used IL4 pred and IL10 pred. After all the steps were done, we found only two final HTL epitopes for our vaccine.

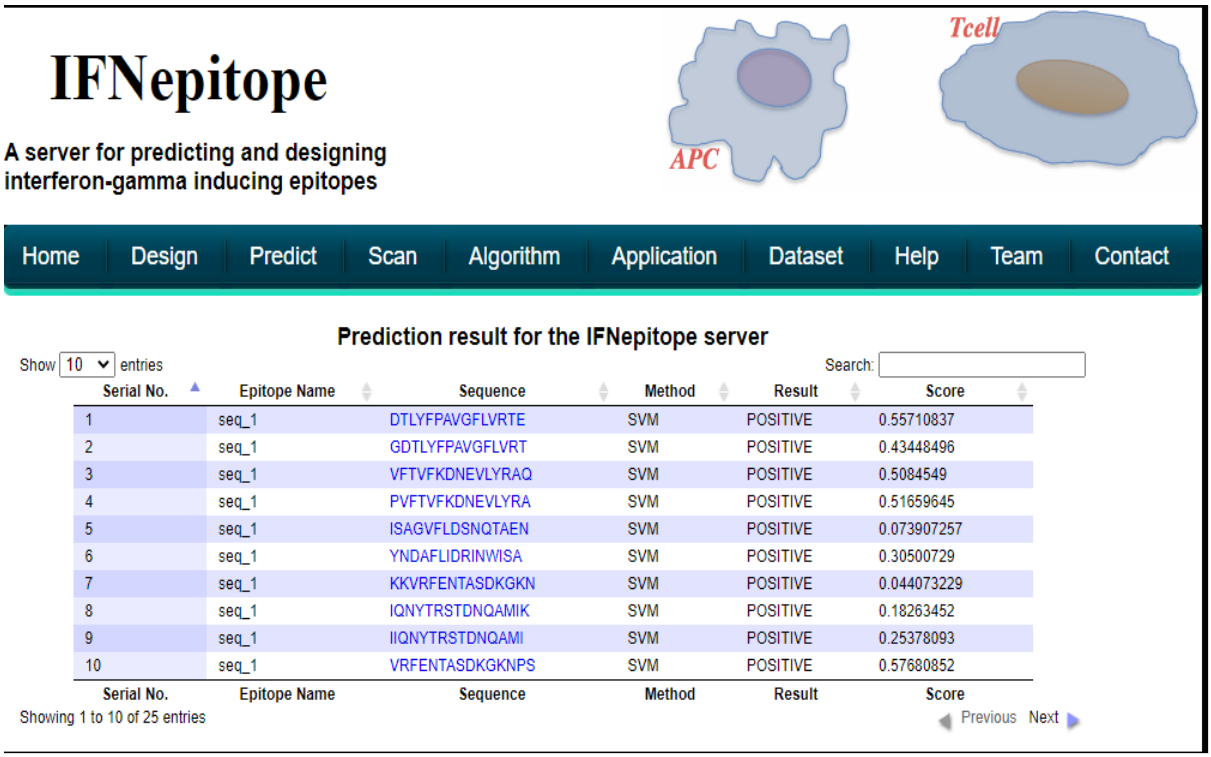


Figure 5 Interferon gamma positive results.

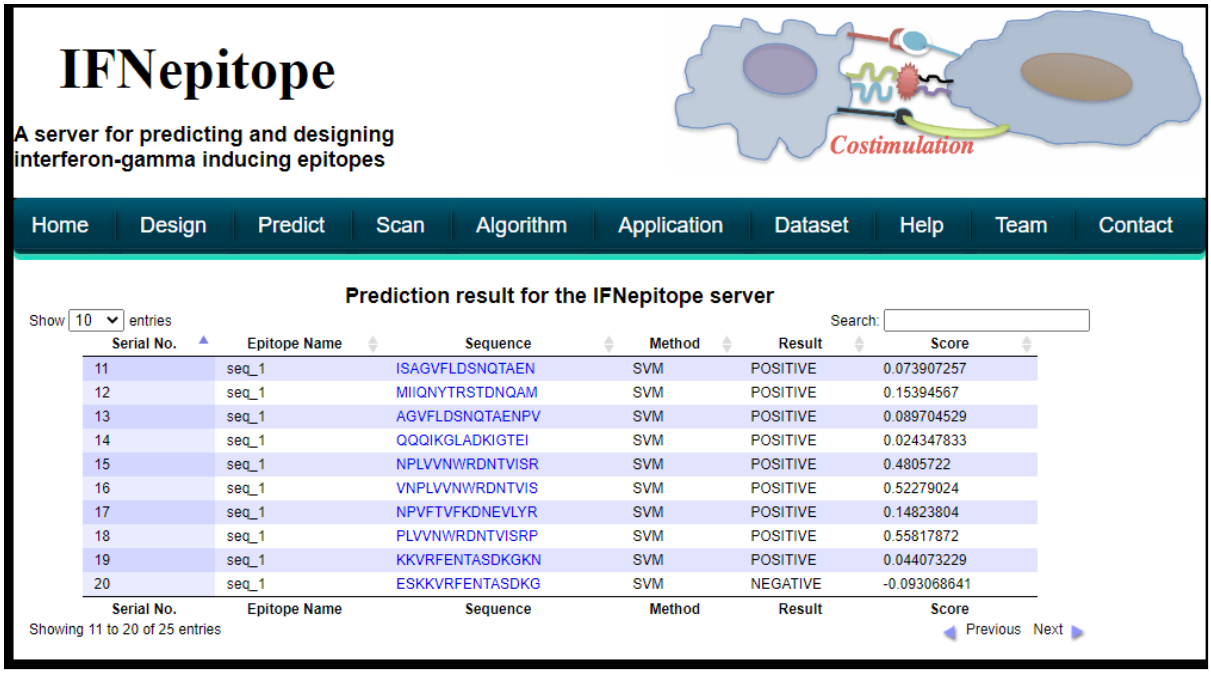


Figure 6 Interferon gamma positive results for rest of the sequences.

Query Peptides

Peptide ID	Peptide Sequence	SVM Score	Prediction	Hydrophobicity	Hydropathicity	Hydrophilicity	Charge	Mol wt
seq_1	DTLYFPAVGFLVRTE	0.83	IL4 inducer	0.02	0.48	-0.41	-1.00	1728.19
seq_1	GDTLYFPAVGFLVRT	0.83	IL4 inducer	0.07	0.69	-0.61	0.00	1656.13
seq_1	VFTVFKDNEVLYRAQ	1.29	IL4 inducer	-0.14	-0.04	-0.14	0.00	1829.30
seq_1	PVFTVFKDNEVLYRA	1.28	IL4 inducer	-0.10	0.09	-0.15	0.00	1798.28
seq_1	ISAGVFLDSNQTAEN	1.37	IL4 inducer	-0.06	-0.09	-0.12	-2.00	1565.87
seq_1	YNDAFLIDRINWISA	1.25	IL4 inducer	-0.03	0.15	-0.45	-1.00	1811.24
seq_1	KKVRFENTASDKGKN	1.47	IL4 inducer	-0.51	-1.81	1.12	3.00	1722.14
seq_1	IQNYTRSTDNQAMIK	1.31	IL4 inducer	-0.32	-1.11	0.11	1.00	1783.21
seq_1	IIQNYTRSTDNQAMI	1.26	IL4 inducer	-0.20	-0.55	-0.21	0.00	1768.20
seq_1	VRFENTASDKGKNPS	1.46	IL4 inducer	-0.39	-1.45	0.74	1.00	1649.98
seq_1	ISAGVFLDSNQTAEN	1.37	IL4 inducer	-0.06	-0.09	-0.12	-2.00	1565.87
seq_1	MIIQNYTRSTDNQAM	1.26	IL4 inducer	-0.23	-0.73	-0.18	0.00	1786.23
seq_1	AGVFLDSNQTAENPV	0.43	IL4 inducer	-0.06	-0.16	-0.12	-2.00	1561.88
seq_1	QQQIKGLADKIGTEI	1.48	IL4 inducer	-0.17	-0.51	0.30	0.00	1642.13
seq_1	NPLVVNWRDNTVISR	1.25	IL4 inducer	-0.23	-0.41	-0.13	1.00	1783.24
seq_1	VNPLVVNWRDNTVIS	1.25	IL4 inducer	-0.07	0.17	-0.43	0.00	1726.19
seq_1	NPVFTVFKDNEVLYR	1.28	IL4 inducer	-0.16	-0.27	-0.11	0.00	1841.31
seq_1	PLVVNWRDNTVISRP	1.25	IL4 inducer	-0.19	-0.28	-0.15	1.00	1766.25
seq_1	KKVRFENTASDKGKN	1.47	IL4 inducer	-0.51	-1.81	1.12	3.00	1722.14
seq_1	ESKKVRFENTASDKG	1.41	IL4 inducer	-0.45	-1.61	1.13	1.00	1696.05
seq_1	TESKKVRFENTASDK	1.45	IL4 inducer	-0.48	-1.63	1.10	1.00	1740.10

Figure 7 Interleukin 4 Positive Results.

ID	Seq	Score	Prediction
	-	-	-
seq_1	DTLYFPAVGFLVRTE	0.531202831505	IL10 inducer
seq_1	VFTVFKDNEVLYRAQ	0.39682971083	IL10 inducer
seq_1	PVFTVFKDNEVLYRA	0.325108814306	IL10 inducer
seq_1	ISAGVFLDSNQTAEN	0.941702216599	IL10 inducer
seq_1	ISAGVFLDSNQTAEN	0.941702216599	IL10 inducer
seq_1	AGVFLDSNQTAENPV	0.639527528401	IL10 inducer
seq_1	NPLVVNWRDNTVISR	0.691832781874	IL10 inducer
seq_1	VNPLVVNWRDNTVIS	0.403309452481	IL10 inducer
seq_1	PLVVNWRDNTVISRP	0.911508967872	IL10 inducer

Figure 8 Interleukin 10 Positive Results.

Peptide Sequences	Interferon Gamma	Inteleukin - 4	Inteleukin - 10	Antigenicity
DTLYFPAVGFLVRTE	Positive	Inducer	Inducer	0.34
GDTLYFPAVGFLVRT	Positive	Inducer	Non inducer	0.29
VFTVFKDNEVLYRAQ	Positive	Inducer	Inducer	0.32
PVFTVFKDNEVLYRA	Positive	Inducer	Inducer	0.44
ISAGVFLDSNQTAEN	Positive	Inducer	Inducer	0.24
YNDAFLIDRINWISA	Positive	Inducer	Non inducer	0.78
KKVRFENTASDKGKN	Positive	Inducer	Non inducer	0.19
IQNYTRSTDNQAMIK	Positive	Inducer	Non inducer	0.34
IIQNYTRSTDNQAMI	Positive	Inducer	Inducer	0.26
VRFENTASDKGKNPS	Positive	Inducer	Non inducer	0.33
ISAGVFLDSNQTAEN	Positive	Inducer	Inducer	0.32
MIIQNYTRSTDNQAM	Positive	Inducer	Non inducer	0.03
AGVFLDSNQTAENPV	Positive	Inducer	Inducer	0.23
QQQIKGLADKIGTEI	Positive	Inducer	Non inducer	0.41
NPLVVNWRDNTVISR	Positive	Inducer	Inducer	1.06
VNPLVVNWRDNTVIS	Positive	Inducer	Inducer	1.24
NPVFTVFKDNEVLYR	Positive	Inducer	Inducer	0.49
KKVRFENTASDKGKN	Positive	Inducer	Non inducer	0.13
ESKKVRFENTASDKG	Negative	Non inducer	Inducer	0.22
TESKKVRFENTASDK	Negative	Non inducer	Inducer	0.15

Figure 9 Final HTL epitopes selection from MHC II prediction. Only two peptide sequences were found those are IFN positive, IL4 & IL10 inducer.

3.4 B cell Epitopes:

There were nineteen different peptides discovered, but out of all of them, we chose just those whose length numbers were greater than 20. Therefore, we were only working with three primary epitopes for B cells and those are also antigenic. They are:

KVRFENTASDKGKNPSKVIKSYYGTMDIKKINEGLL,

NENVNEKCKFTLPPLKIHECNISCPNPLPFREYKPQTEGVSNLVGLPNNICLQKTSNQI

LKPKLISYTLPVVGQSG and

VRTEFKYNDSNCPIAECQYSKPENCRLSMGIRPNS.

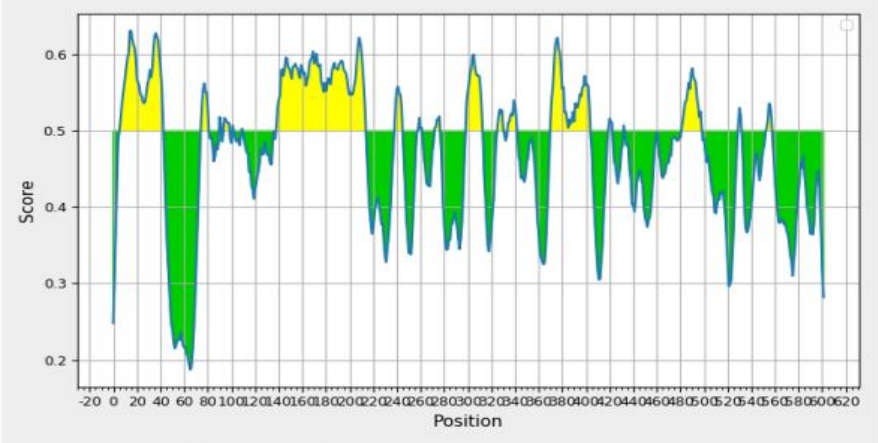


Figure 10 B cell result.

Predicted peptides:

No.	Start	End	Peptide	Length
1	7	43	KVRFENTASDKGKNPSKVIKSYYGTMDIKKINEGLLD	37
2	75	81	RSTDNQA	7

ols.iedb.org/bcell/result/

26/22, 10:05 PM B Cell Result

No.	Start	End	Peptide	Length
3	91	92	QQ	2
4	95	99	KGLAD	5
5	102	102	G	1
6	110	110	S	1
7	139	214	NENVNEKCKFTLPPLKIHECNISCPNPLPFREYKPQTEGVSNLVGLPNNICLQKTSNQILKPKLISYTLPVVGQSG	76
8	239	245	SCSRGVF	7
9	259	262	GDEV	4
10	273	277	PSNPN	5
11	300	313	VGDPILNSTYWSGS	14
12	326	331	NGESYN	6
13	335	342	FALRNIEK	8
14	371	404	VRTEFKYNDSNCPIAECQYSKPENCRLSMGIRPN	34
15	420	423	DEEN	4
16	433	433	D	1
17	482	499	NTVISRPGQSQCPRFNKC	18
18	530	532	QTA	3
19	554	558	EDTNA	5

Figure 11 B cell prediction result from them only 20 length above epitopes were taken.

3.5 Final Protein Sequence and its Antigenicity, Allergenecity and Toxicity

We combined the primary adjuvant with all of the identified CTL, HTL, and B cell epitopes using a linker. Adjuvant-linked CTL epitopes (EAAAK), CTL epitopes linked together (AAY), CTL epitopes linked to HTL epitopes (GPGPG), and B-cell epitopes linked to HTL epitopes (KK) were all used in this study. Here is our final protein sequence:

MPTESKKVRFENTASDKGKNPSKVIKSYYGTMDIKKINEGLLDSKILSAFNTVIALLG SIVIIVMNIMIIQNYTRSTDNQAMIKDALQSIQQQIKGLADKIGTEIGPKVSLIDTSSTITI PANIGLLGSKISQSTASINENVNEKCKFTLPPLKIHECNISCPNPLPFREYKPQTEGVSN LVGLPNNICLQKTSNQILKPKLISYTLPVVGQSGTCITDPLLAMDEGYFAYSHLEKIGS CSRGVFKQRIIGVGEVLDRGDEVPSLFMTNVWTPSNPNTVYHCSAVYNNEFYYVLC AVSVVGDPILNSTYWSGSLMMTRLAVKPKNNGESYNQHQFALRNIEKGKYDKVMP YGPSGIKQGDTLYFPAVGFLVRTEFKYNDSNCPIAECQYSKPENCRLSMGIRPNSHYI LRSGLLKYNLSDEENSKIVFIEISDQRLSIGSPSKIYDSLGQPVFYQASFSWDTMIKFGD VQTVNPLVVNWRDNTVISRPGQSQCPRFNKCPEVCWEGVYNDAFLIDRINWISAGVF LDSNQTAENPVFTVFKDNEVLYRAQLASEDTNAQKTITNCFLLKNKIWCISLVEIYDT GDNVIRPKLFAVKIPEQCTEAAAKAMDEGYFAYAAYGIKQGDTLYGPGPGNPLVVN WRDNTVISRGPGPGVNPLVVNWRDNTVISKKKVRFENTASDKGKNPSKVIKSYYGT MDIKKINEGLLKKNENVNEKCKFTLPPLKIHECNISCPNPLPFREYKPQTEGVSNLVG LPNNICLQKTSNQILKPKLISYTLPVVGQSGKKVRTEFKYNDSNCPIAECQYSKPENC RLSMGIRPNS.

The antigenicity of our adjuvant was measured at 0.5215. Following the addition of antigenic and non-antigenic CTL, HTL, and B cell epitopes, the antigenicity of the final protein decreased to 0.51, which indicated the presence of a plausible antigen even though the score was reduced. Therefore, we decided to include solely antigenic CTL, HTL, and B cell epitopes in our construct. After that, the antigenicity score went up to 0.5522, reflecting an increase.

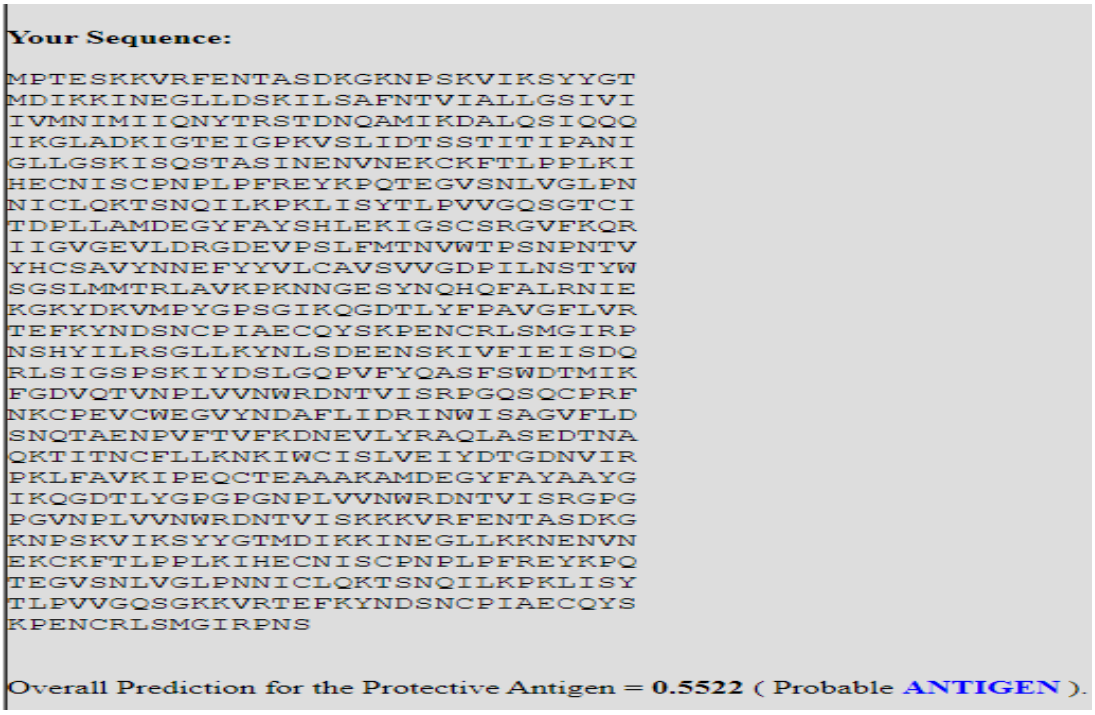


Figure 12 Antigenicity result of the final protein sequence was found 0.5522 which is probable antigenic.

In addition to that, there is no evidence that this protein sequence is allergenic, and it does not include any toxins either.

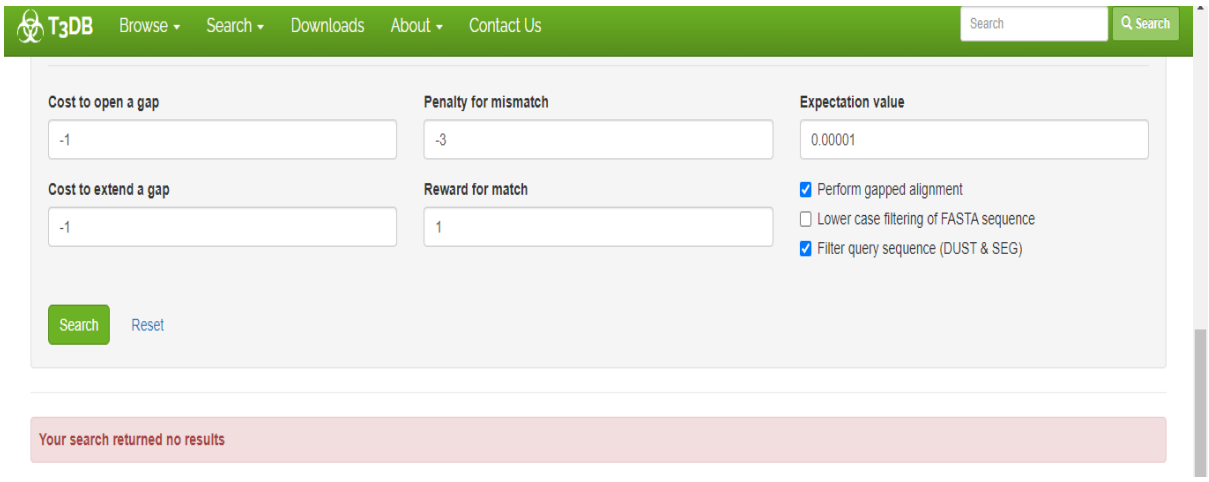
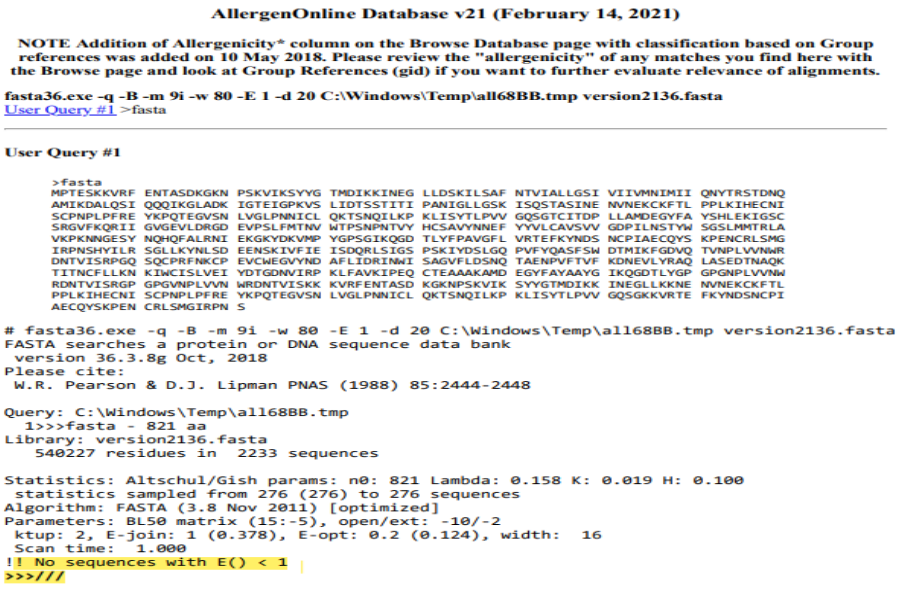


Figure 13 Toxicity result. There was no toxicity found.

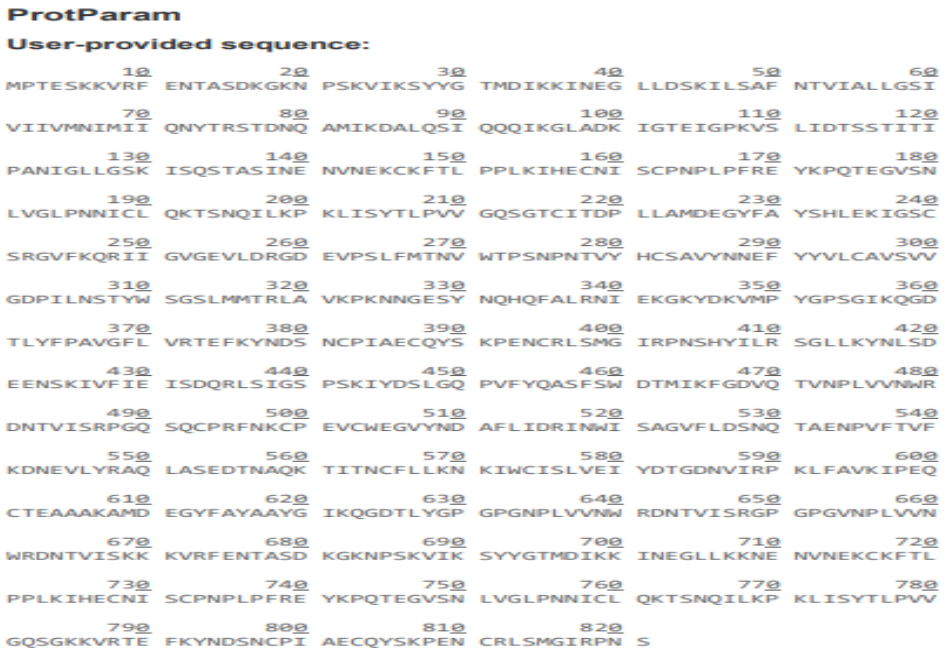


```
                    AllergenOnline Database v21 (February 14, 2021)

  NOTE Addition of Allergenicity* column on the Browse Database page with classification based on Group
  references was added on 10 May 2018. Please review the "allergenicity" of any matches you find here with
  the Browse page and look at Group References (gid) if you want to further evaluate relevance of alignments.

fasta36.exe -q -B -m 9i -w 80 -E 1 -d 20 C:\Windows\Temp\all68BB.tmp version2136.fasta
User Query #1 >fasta

User Query #1

      >fasta
      MPTESKKVRF ENTASDKGKN PSKVIKSYYG TMDIKKINEG LLDSKILSAF NTVIALLGSI VIIVMNIMII QNYTRSTDNQ
      AMIKDALQSI QQQIKGLADK IGTEIGPKVS LIDTSSTITI PANIGLLGSK ISQSTASINE NVNEKCKFTL PPLKIHECNI
      SCPNPLPFRE YKPQTEGVSN LVGLPNNICL QKTSNQILKP KLISYTLPVV GQSGTCITDP LLAMDEGYFA YSHLEKIGSC
      SRGVFKQRII GVGEVLDRGD EVPSLFMTNV WTPSNPNTVY HCSAVYNNEF YYVLCAVSVV GDPILNSTYW SGSLMMTRLA
      VKPKNNGESY NQHQFALRNI EKGKYDKVMP YGPSGIKQGD TLYFPAVGFL VRTEFKYNDS NCPIAECQYS KPENCRLSMG
      IRPNSHYILR SGLLKYNLSD EENSKIVFIE ISDQRLSIGS PSKIYDSLGQ PVFYQASFSW DTMIKFGDVQ TVNPLVVNWR
      DNTVISRPGQ SQCPRFNKCP EVCWEGVYND AFLIDRINWI SAGVFLDSNQ TAENPVFTVF KDNEVLYRAQ LASEDTNAQK
      TITNCFLLKN KIWCISLVEI YDTGDNVIRP KLFAVKIPEQ CTEAAAKAMD EGYFAYAAYG IKQGDTLYGP GPGNPLVVNW
      RDNTVISRGP GPGVNPLVVN WRDNTVISKK KVRFENTASD KGKNPSKVIK SYYGTMDIKK INEGLLKKNE NVNEKCKFTL
      PPLKIHECNI SCPNPLPFRE YKPQTEGVSN LVGLPNNICL QKTSNQILKP KLISYTLPVV GQSGKKVRTE FKYNDSNCPI
      AECQYSKPEN CRLSMGIRPN S

# fasta36.exe -q -B -m 9i -w 80 -E 1 -d 20 C:\Windows\Temp\all68BB.tmp version2136.fasta
FASTA searches a protein or DNA sequence data bank
 version 36.3.8g Oct, 2018
Please cite:
 W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

Query: C:\Windows\Temp\all68BB.tmp
  1>>>fasta - 821 aa
Library: version2136.fasta
   540227 residues in  2233 sequences

Statistics: Altschul/Gish params: n0: 821 Lambda: 0.158 K: 0.019 H: 0.100
 statistics sampled from 276 (276) to 276 sequences
Algorithm: FASTA (3.8 Nov 2011) [optimized]
Parameters: BL50 matrix (15:-5), open/ext: -10/-2
 ktup: 2, E-join: 1 (0.378), E-opt: 0.2 (0.124), width:  16
 Scan time:  1.000
!! No sequences with E() < 1
>>>///
```

Figure 14 Allergenecity result. No Allergenecity was found.

3.6 Biochemical Analysis Result of the Vaccine:

The formula for this compound is as follows: C4084H6457N1089O1220S39. The number of amino acids in this compound is 821, and its molecular weight is 91.584 kDa, which indicates that it is acceptable because its size is less than 100 kDa. Theoretical pI of the the vaccine model is 8.83 which indicates that the vaccine model is slightly basic in nature. The total number of atoms in the protein is 12889, and its instability index (II) was calculated to be 33.21. Since an index value of 40 or higher indicates that the protein is unstable, this value indicates that the protein is stable. Index of aliphatic compounds: 84.96 this stability index high means it is stable in different temperatures. It can be deduced from the value of -0.316 for the grand average of hydropathicity (GRAVY) that the protein is hydrophilic. M is located at the beginning of the series under consideration (Met). It is estimated that the half-life is thirty hours (mammalian reticulocytes, in vitro). >10 hours (in vitro), >20 hours (in vivo yeast).

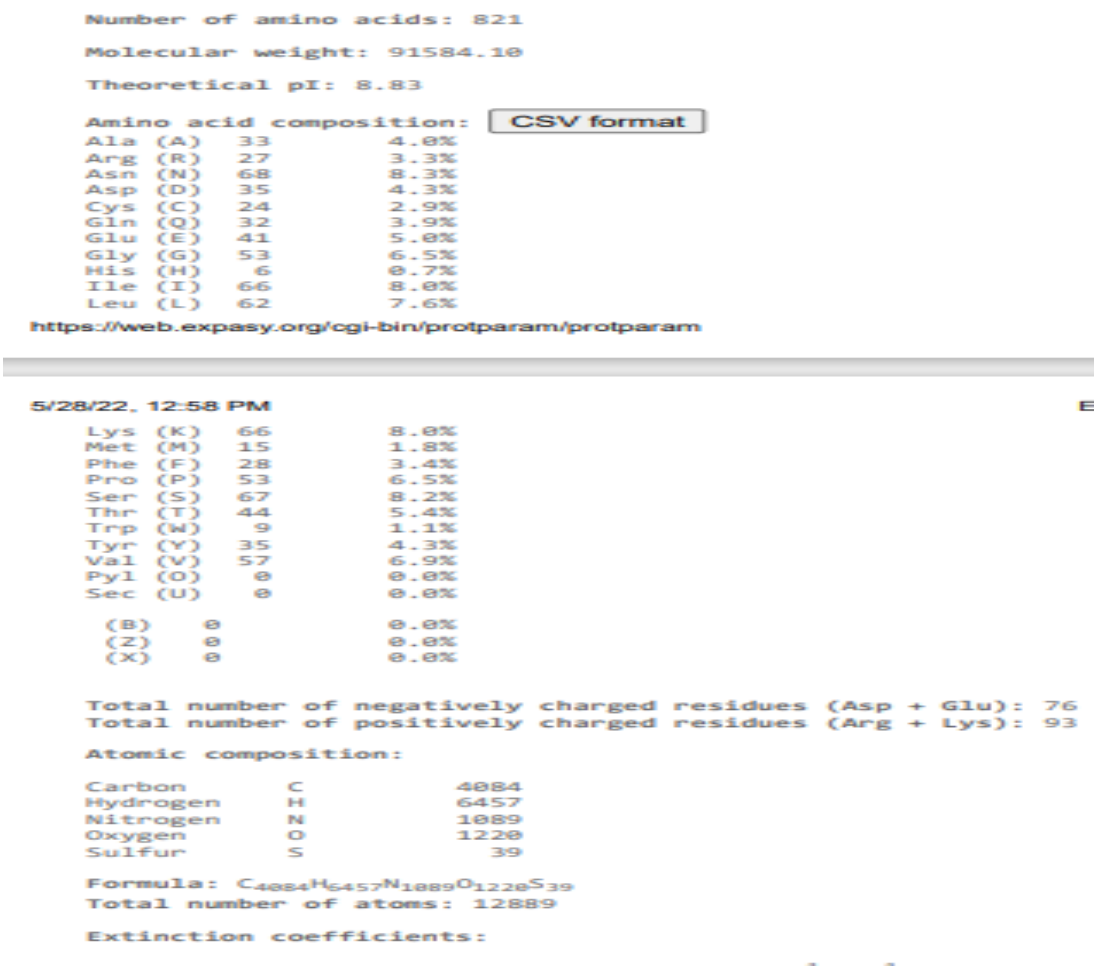


ProtParam

User-provided sequence:

```
        10         20         30         40         50         60
MPTESKKVRF ENTASDKGKN PSKVIKSYYG TMDIKKINEG LLDSKILSAF NTVIALLGSI

        70         80         90        100        110        120
VIIVMNIMII QNYTRSTDNQ AMIKDALQSI QQQIKGLADK IGTEIGPKVS LIDTSSTITI

       130        140        150        160        170        180
PANIGLLGSK ISQSTASINE NVNEKCKFTL PPLKIHECNI SCPNPLPFRE YKPQTEGVSN

       190        200        210        220        230        240
LVGLPNNICL QKTSNQILKP KLISYTLPVV GQSGTCITDP LLAMDEGYFA YSHLEKIGSC

       250        260        270        280        290        300
SRGVFKQRII GVGEVLDRGD EVPSLFMTNV WTPSNPNTVY HCSAVYNNEF YYVLCAVSVV

       310        320        330        340        350        360
GDPILNSTYW SGSLMMTRLA VKPKNNGESY NQHQFALRNI EKGKYDKVMP YGPSGIKQGD

       370        380        390        400        410        420
TLYFPAVGFL VRTEFKYNDS NCPIAECQYS KPENCRLSMG IRPNSHYILR SGLLKYNLSD

       430        440        450        460        470        480
EENSKIVFIE ISDQRLSIGS PSKIYDSLGQ PVFYQASFSW DTMIKFGDVQ TVNPLVVNWR

       490        500        510        520        530        540
DNTVISRPGQ SQCPRFNKCP EVCWEGVYND AFLIDRINWI SAGVFLDSNQ TAENPVFTVF

       550        560        570        580        590        600
KDNEVLYRAQ LASEDTNAQK TITNCFLLKN KIWCISLVEI YDTGDNVIRP KLFAVKIPEQ

       610        620        630        640        650        660
CTEAAAKAMD EGYFAYAAYG IKQGDTLYGP GPGNPLVVNW RDNTVISRGP GPGVNPLVVN

       670        680        690        700        710        720
WRDNTVISKK KVRFENTASD KGKNPSKVIK SYYGTMDIKK INEGLLKKNE NVNEKCKFTL

       730        740        750        760        770        780
PPLKIHECNI SCPNPLPFRE YKPQTEGVSN LVGLPNNICL QKTSNQILKP KLISYTLPVV

       790        800        810        820
GQSGKKVRTE FKYNDSNCPI AECQYSKPEN CRLSMGIRPN S
```

Figure 15 ProtParam tool result.



```
Number of amino acids: 821

Molecular weight: 91584.10

Theoretical pI: 8.83

Amino acid composition:   CSV format
Ala (A)   33         4.0%
Arg (R)   27         3.3%
Asn (N)   68         8.3%
Asp (D)   35         4.3%
Cys (C)   24         2.9%
Gln (Q)   32         3.9%
Glu (E)   41         5.0%
Gly (G)   53         6.5%
His (H)    6         0.7%
Ile (I)   66         8.0%
Leu (L)   62         7.6%
```

https://web.expasy.org/cgi-bin/protparam/protparam

5/28/22, 12:58 PM

```
Lys (K)   66         8.0%
Met (M)   15         1.8%
Phe (F)   28         3.4%
Pro (P)   53         6.5%
Ser (S)   67         8.2%
Thr (T)   44         5.4%
Trp (W)    9         1.1%
Tyr (Y)   35         4.3%
Val (V)   57         6.9%
Pyl (O)    0         0.0%
Sec (U)    0         0.0%

 (B)    0           0.0%
 (Z)    0           0.0%
 (X)    0           0.0%

Total number of negatively charged residues (Asp + Glu): 76
Total number of positively charged residues (Arg + Lys): 93

Atomic composition:

Carbon      C          4084
Hydrogen    H          6457
Nitrogen    N          1089
Oxygen      O          1220
Sulfur      S            39

Formula: C4084H6457N1089O1220S39
Total number of atoms: 12889

Extinction coefficients:
```

Figure 16 ProtParam tool result.

Extinction coefficients:

Extinction coefficients are in units of M^{-1} cm^{-1}, at 280 nm measured in water.

Ext. coefficient 103150
Abs 0.1% (=1 g/l) 1.126, assuming all pairs of Cys residues form cystines

Ext. coefficient 101650
Abs 0.1% (=1 g/l) 1.110, assuming all Cys residues are reduced

Estimated half-life:

The N-terminal of the sequence considered is M (Met).

The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
>20 hours (yeast, in vivo).
>10 hours (Escherichia coli, in vivo).

Instability index:

The instability index (II) is computed to be 33.21
This classifies the protein as stable.

Aliphatic index: 84.96

Grand average of hydropathicity (GRAVY): -0.316

Figure 17 ProtParam tool result.

3.7 Vaccines 3D Model Assessment:

From phyre 2 server we found the PDB format of the model and the structures of the protein. Here we found that our structure is showing 100% confidence with 52% coverage and 427 residues.

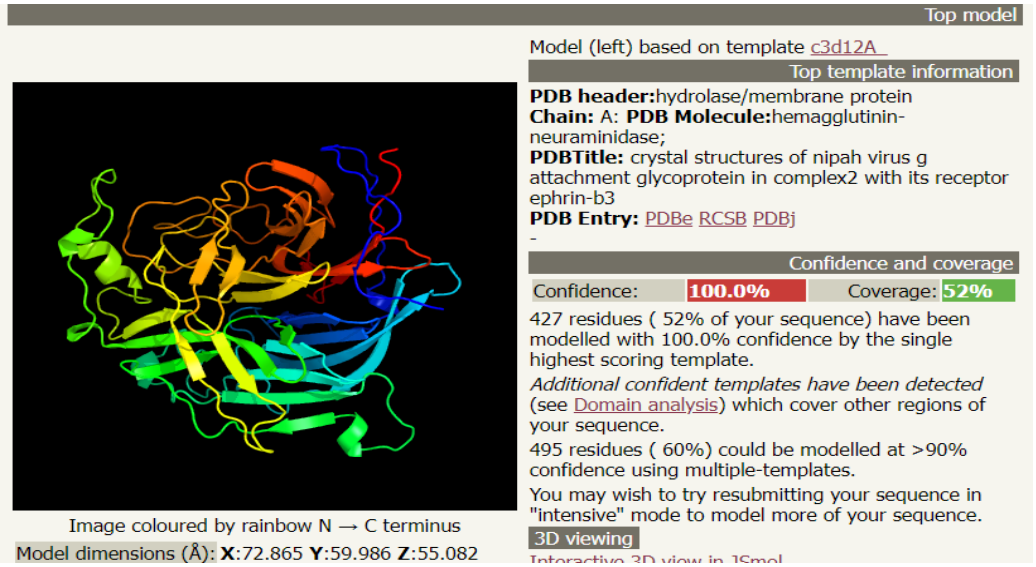


Figure 18 Homology Model of the Structure, 52% coverage was found here.

The Z-score of the model was the next thing that we determined. A positive outcome is that the black spot can be located between the grey line and the blue line, as shown here. The fact that this score, which is -8.47, is significantly lower than the average indicates more activity.

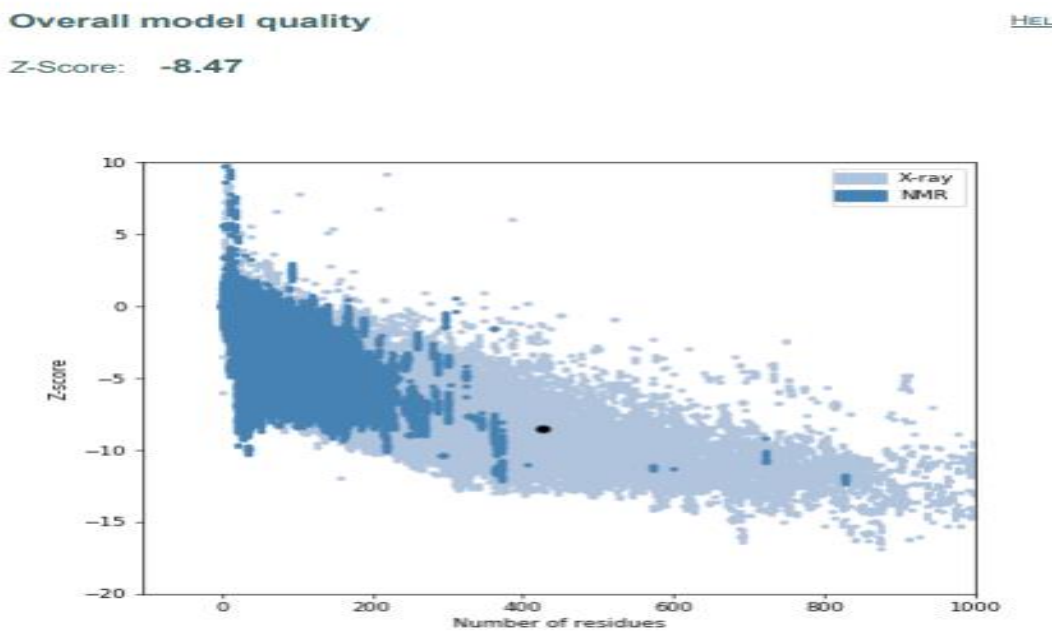


Figure 19 Z - Score of the Model.

Figure 20 Local Model Quality.

Next we have done quality assessment of the structure by checking the Ramachandran favoured score which is 89.65% and only 1.18% is outlier (328 GLU, 474 PRO, 258 ARG, 421 GLU,

338 ARG). Roatamer outlier is zero. Molprobidity score is 2.73. But there are some bad bonds exist (17 / 3470 382 CYS- 395 CYS, 565 CYS- 574 CYS, 387 CYS- 499 CYS, 282 CYS- 295 CYS, 406 HIS, 333 HIS, 233 HIS, 281 HIS, 323 PRO, 185 PRO, 189 CYS- 601 CYS, 493 CYS503 CYS, 303 PRO) and some bad angles too (11 / 4722) (216 CYS- 240 CYS), (493 CYS- 503 CYS), (382 CYS- 395 CYS), (565 CYS- 574 CYS), (387 CYS- 499 CYS), (189 CYS- 601 CYS), 333 HIS, 406 HIS, 281 HIS, 233 HIS).

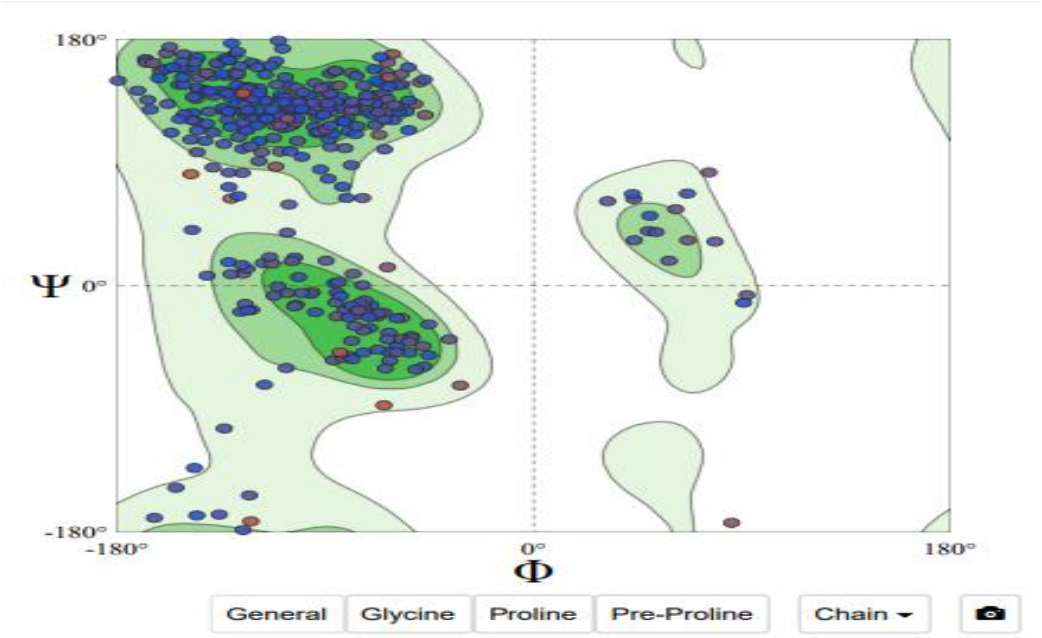


Figure 21 Ramachandran Plots Result 89.65% are covered and 1.18% outlier.

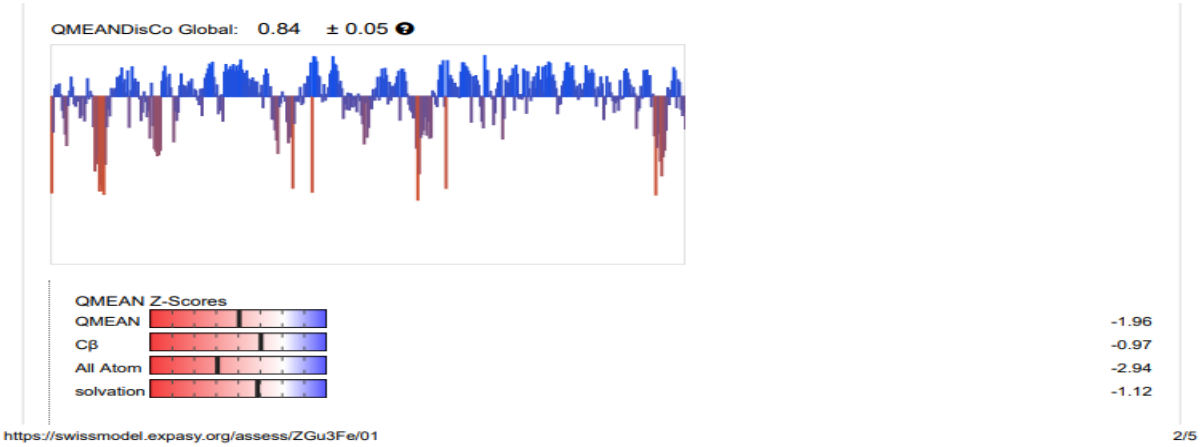


Figure 22 Quality Estimation.

3.8 Immune Simulation:

We have done our immune simulation system by using c-immsim server. From this server we got to know the immunological reaction as well as antigenicity. Here we gave three doses and found some graphs which represents the working system of the vaccine.

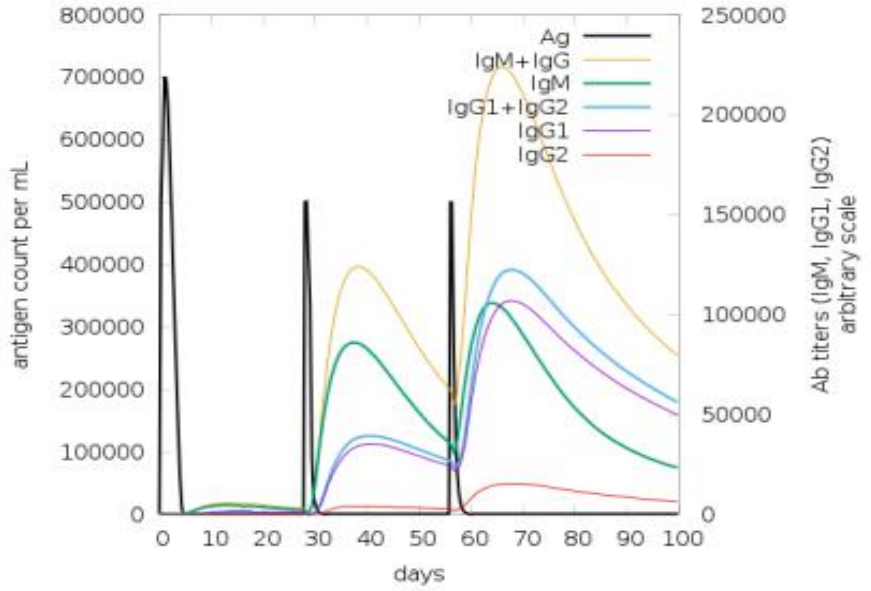


Figure 23 Antibodies, immunocomplexes, and the virus are involved.

After administering the first dose, antigenicity increased among seven lac individuals, and so did the total amount of IgM + IgG, which included an increase in IgG1 and IgG2 levels. This information was gleaned from the graph. After that, when administering the second and third doses, antigenicity generation began to decrease because, by that point, the majority of the amount had already been covered by the first dose.

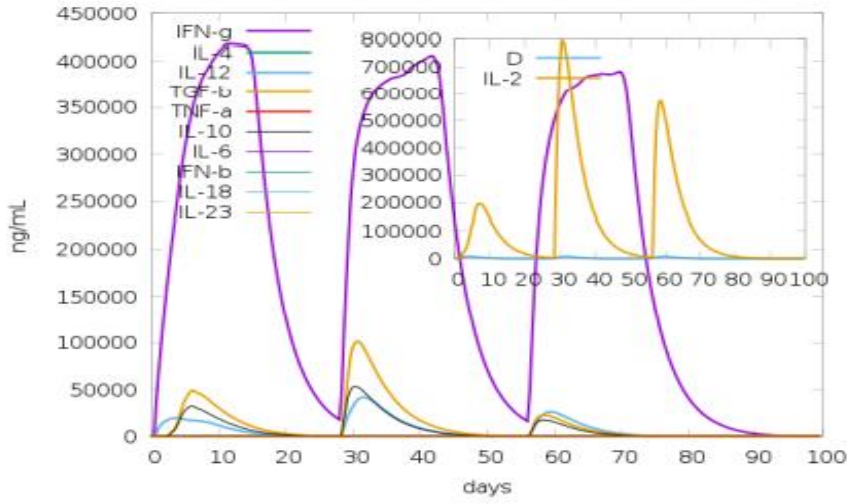


Figure 24 Concentration of Interferon gamma, cytokines and interleukins increased after giving three doses.

From (Figure 24) it can be assured that after giving the doses the effect of interferon gamma, interleukins are increased. In the third dose the amount is decreased because already most of the antibody was produced.

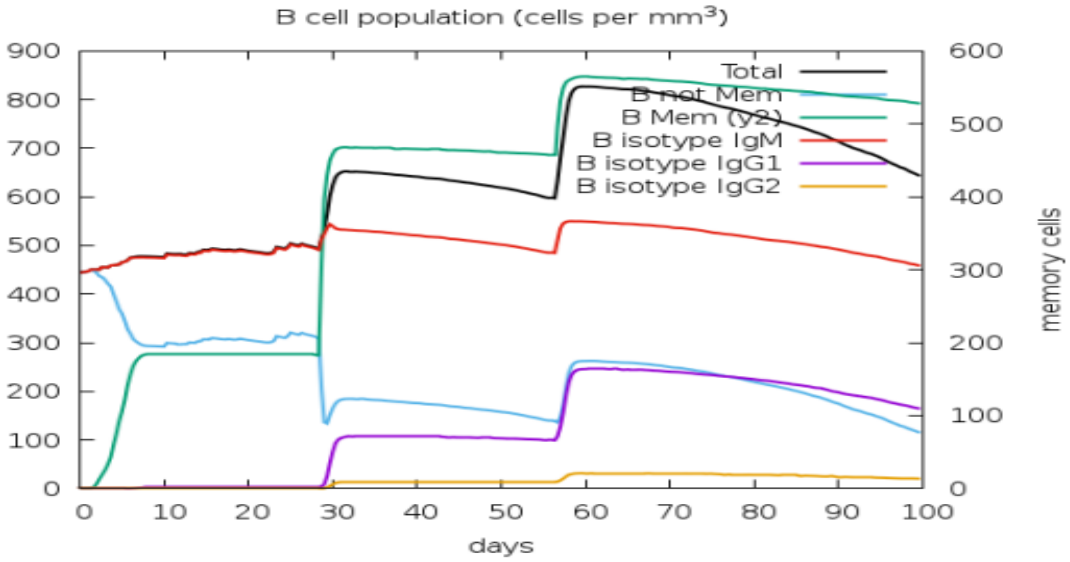


Figure 25 B cell growth among the population was increased.

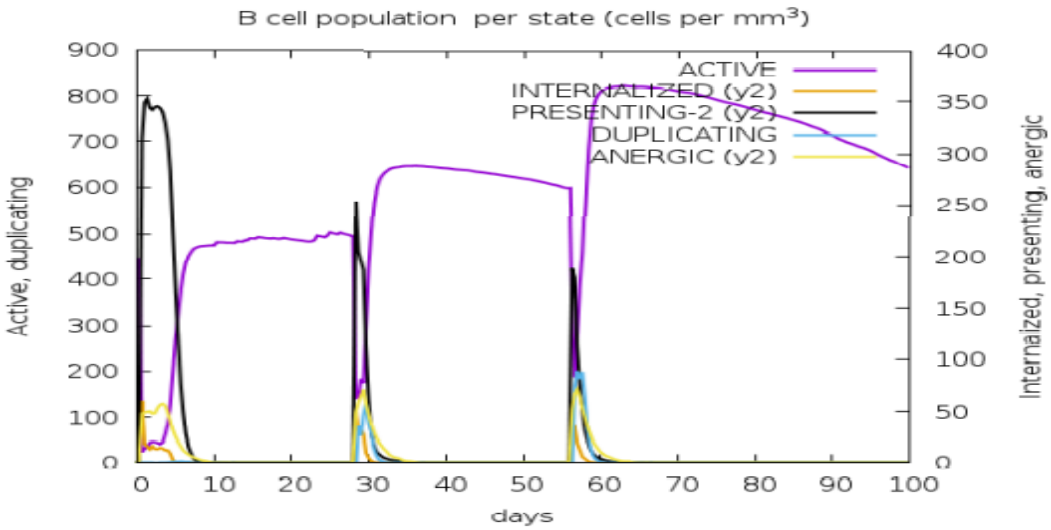


Figure 26 B cell population growth per state increased.

In this (Figure 25 & 26) it is clearly shown that the memory B cells growth also increased.

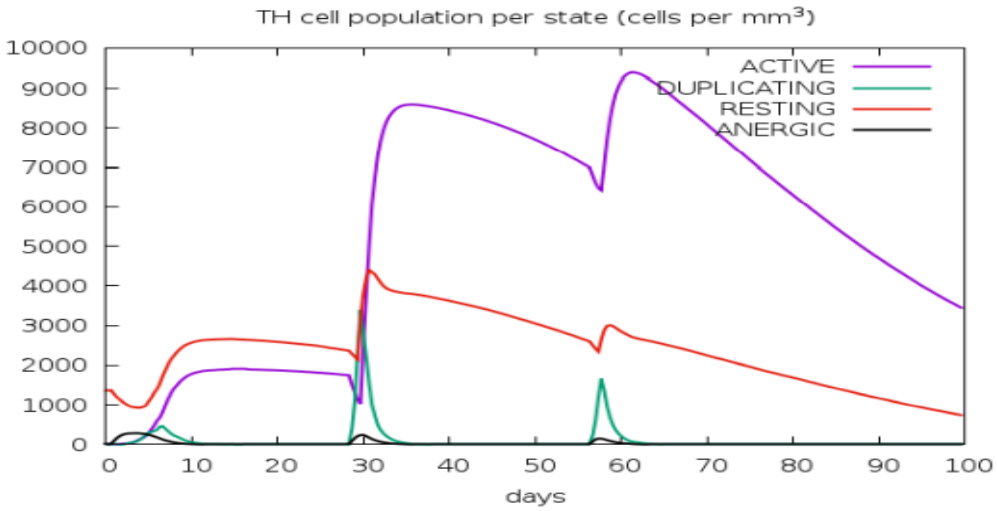


Figure 27 Cytotoxic T cell population growth per state increased.

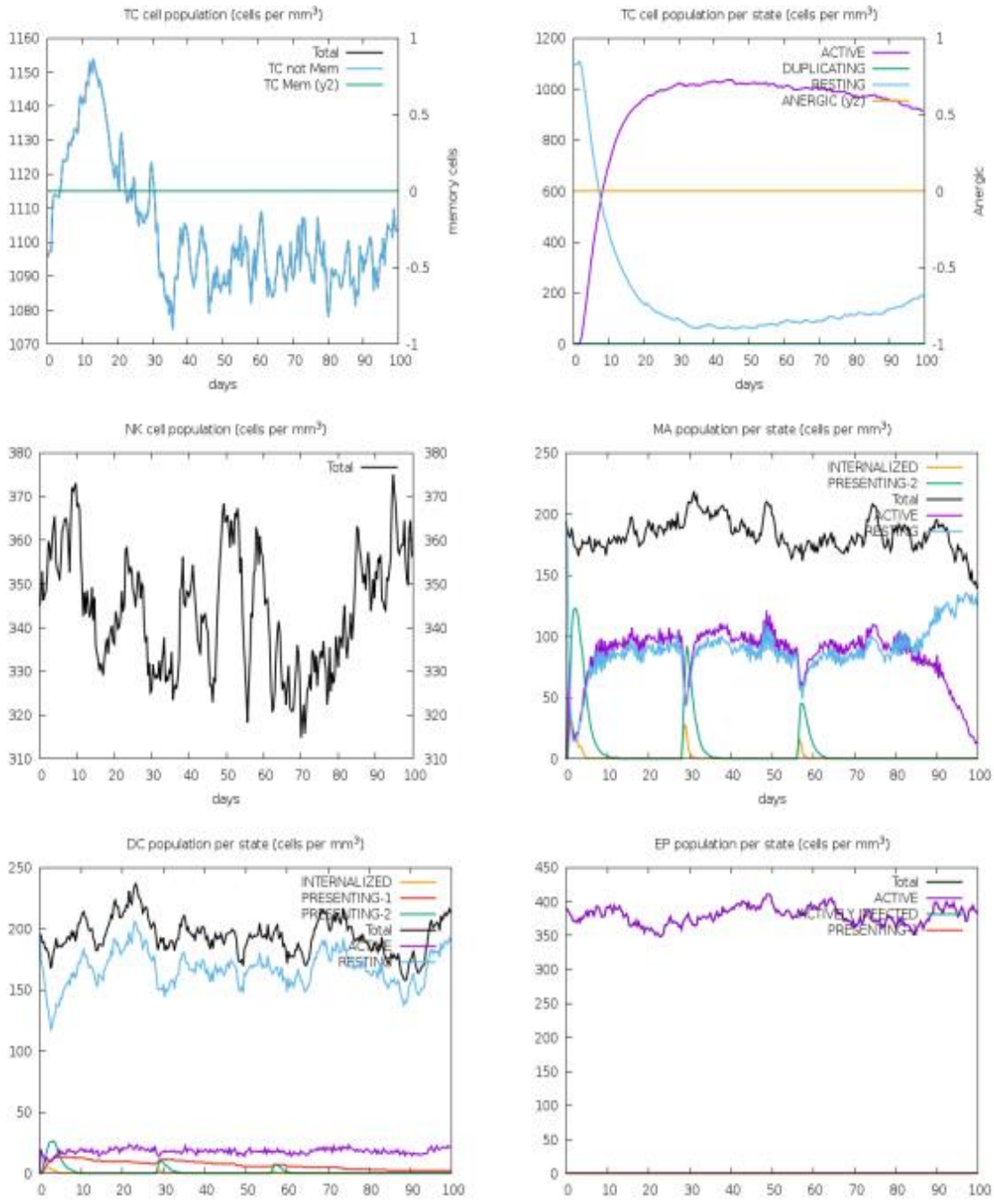


Figure 28 The symbols used in the legend are shown in the figure above.

Based on figure 28, it is possible to be sure that the immunity of the vaccine's epitopes is growing among the population, particularly after the administration of the first dose. The first graph shows the t cell's result per population; the second graph is about TC cell per population state, the third graph is about neural crest cell population, the fourth graph is about macrophages, and the fifth is about dendritic cells. The last one is about the epithelial cell.

3.9 Molecular Docking:

We have done our molecular docking by using patchdock server. Here we used TLR3 receptor binder (1ziw) and the pdb format of the protein.

Molecular Docking Algorithm Based on Shape Complementarity Principles

[About PatchDock] [Web Server] [Download] [Help] [FAQ] [References]

Receptor	Ligand	Complex Type	Clustering RMSD	User e-mail	Receptor Site	Ligand Site	Distance Constraints
1ziw	Student_p2.pdb	Default	4.0	sabrina.sultana.moon@g.bracu.ac.bd	-	-	-

Solution No	Score	Area	ACE	Transformation	PDB file of the complex
1	18946	2682.30	463.57	0.86 -0.02 -2.29 89.63 -1.34 164.26	result.1.pdb
2	18262	2614.60	463.10	0.63 1.24 1.38 -13.51 -141.65 28.94	result.2.pdb
3	18028	3035.20	464.37	1.89 0.87 2.09 174.63 -53.79 50.23	result.3.pdb
4	17770	2347.20	419.95	2.17 0.97 2.07 178.58 -29.12 23.81	result.4.pdb
5	17410	3357.10	484.40	3.04 -0.76 -0.51 -72.81 -80.97 31.73	result.5.pdb
6	16844	2164.40	400.91	3.11 -0.72 3.04 145.24 66.10 22.72	result.6.pdb
7	16782	2588.30	366.80	0.46 -0.16 2.65 45.45 -88.92 158.54	result.7.pdb
8	16764	2365.80	484.98	1.49 0.67 2.39 162.63 -56.01 96.02	result.8.pdb
9	16670	2357.90	441.47	-0.05 -0.32 1.64 -74.96 -45.66 132.61	result.9.pdb
10	16582	2422.70	401.08	3.00 -0.86 -2.58 53.42 103.44 44.52	result.10.pdb
11	16560	2485.60	430.60	-1.12 -0.46 -1.17 113.71 80.34 14.62	result.11.pdb
12	16508	3376.00	464.53	0.56 0.99 1.64 9.87 -148.71 60.40	result.12.pdb
13	16408	2165.00	341.61	3.05 -1.09 -2.70 69.33 95.00 64.66	result.13.pdb
14	16380	2322.30	370.48	1.72 0.02 1.00 74.18 -102.54 142.04	result.14.pdb
15	16340	3065.80	403.02	-1.27 -0.46 -1.53 139.95 55.18 -3.41	result.15.pdb
16	16152	3313.20	437.56	1.15 -0.08 -2.60 66.77 14.14 168.75	result.16.pdb
17	15878	2336.40	457.95	-1.17 1.38 -1.32 145.26 -55.42 -72.51	result.17.pdb
18	15774	2442.20	348.33	-3.10 -1.23 3.04 135.53 54.90 73.52	result.18.pdb
19	15754	2514.20	443.31	-2.93 1.18 1.05 153.90 -13.07 -41.31	result.19.pdb
20	15754	2010.40	471.09	1.80 0.48 2.15 164.18 -48.09 96.39	result.20.pdb

show next 20 »»

Figure 29 Patchdock result from these result 1, score 18946 were taken.

Based on this outcome (fig. 29), we decided to go with the option that received the highest score, which was solution number one. This binding encompassed 2682.30 square meters and had a total of 463.57 ACE; its transformation is 0.86-0.02 -2.29 89.63 -1.34 164.26.

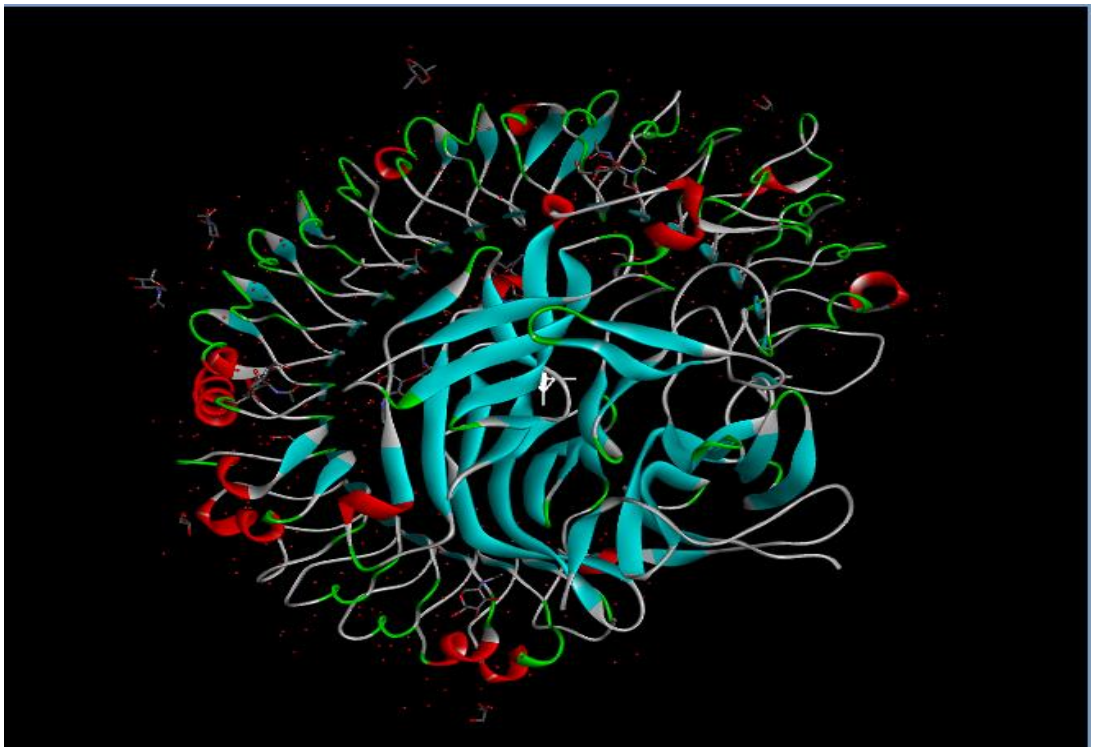


Figure 30 Three dimensional structure of protein - protein binding.

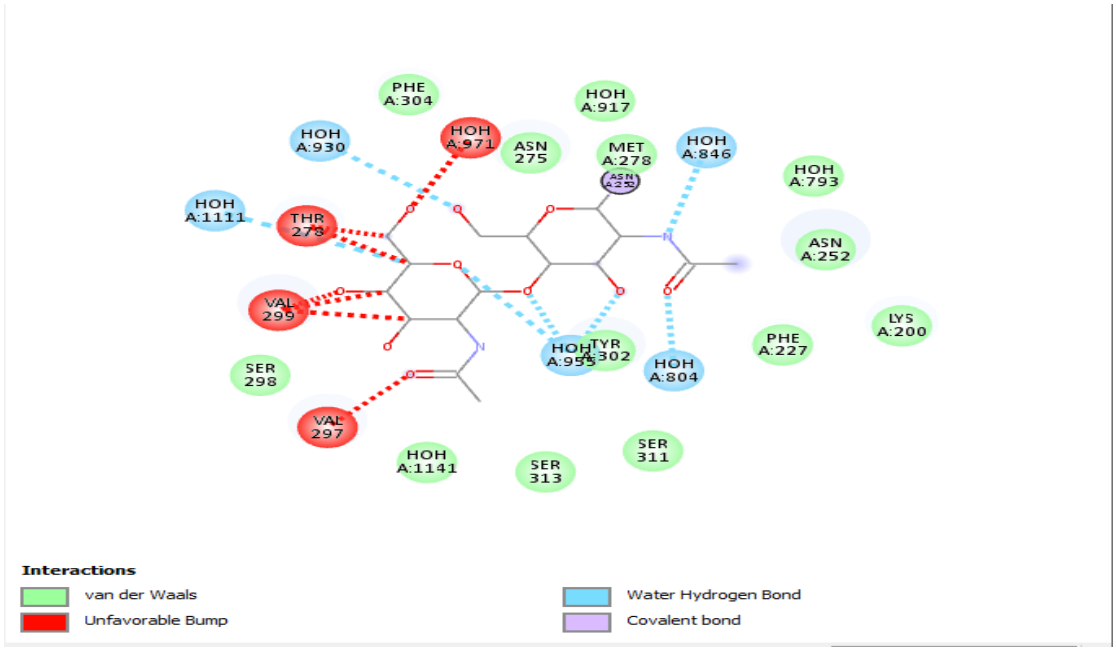


Figure 31 Two dimensional structure of protein - protein binding.

Chapter 4

Discussion

Recent studies indicate that we were successful in developing a multi-epitope-based vaccination against Nipah henipavirus glycoprotein G by employing an in silico strategy, which is founded on the application of computational technology. We find highly antigenic vaccines using this strategy, which also results in less adverse reactions. In this case, the chosen protein sequence demonstrates a strong binding affinity with two epitopes on cytotoxic T lymphocytes, two epitopes on helper T lymphocyte cells, and three epitopes on B cells. The antigenicity of the protein sequence enhanced as a result of the addition of several antigenic epitopes. Additionally, this final protein sequence does not exhibit any harmful consequences, and subsequent injections result in a rise in the population's overall antibody levels. This sequence has a confidence level of 100 percent while also having coverage of 52 percent, which indicates that it is a good model.

The structure and coverage of the dataset can be increased to further develop this method. To create the dataset for this investigation, we chose a portion of the known epitopes and protective antigens. We will be capable of developing a better, more complete, and speedier vaccine creation tool with a larger dataset and more epitope combination possibilities. The existing model can still handle the majority of the current cases and give an effective vaccine design despite the restricted datasets that are currently accessible.

Moreover, from figure 18, we found the Z score of -8.47; and there is a black spot which is spotted inside the X-ray region and it indicates that it is an acceptable score according to the algorithm. The negative Z score indicates that here the amino acids residues, the only part of the molecule that has a favorable characteristic is the N-terminal region.

Then we found the Grand average hydropathicity score of -0.316 means the protein is hydrophilic and thermostable. So it can be considered as a good result because for constructing a vaccine water solubility is an important part.

Figure 20, therefore, demonstrates that the immunization was beneficial to the majority of the Ramachandran plot with only a tiny amount of outliers. A value greater than 85 percent of the total is considered an adequate Ramachandran plot score. This score represents the size of regions that are energetically favoured. Since our finding score is 89.65 percent, it has been determined that it is an excellent score.

Figure (22-28) represents that the vaccine may produce good immunity to the population as the antibody, IFN, Interleukin and B, T cells graphs are uphill. From figure 23 it shows that after giving the doses the antigenicity of the vaccine starts reducing and antibody (IgG and IgM) starts producing in the body. Figure 24 shows that according to the doses the interferon gamma starts decreasing because the antibody almost produced in the body. Figure 25 & 26 indicates that B cells increased highly according to the doses. Figure 27 shows that Cytotoxic T cells were increased following the dose interval.

Based on these results, we may generally assume that future research and development with this vaccine is possible.

Chapter 5

Conclusion

Since the first outbreak of the Nipah virus in humans that was recorded, about 15 years have passed. Since then, Bangladesh has had outbreaks of the Nipah virus approximately yearly. Nipah viruses continue to be a concern to human and animal health, necessitating a multidisciplinary strategy centered on therapeutic and preventative intervention techniques.

Numerous facets of public health need to be improved if there is any chance of successfully preventing the spread of the Nipah virus. Because the transmission of NiV is ongoing from one year to the next, we decided to concentrate on avoiding this disease. From 2020 all over the world is suffering for COVID-19 virus after that in Kerala there was an outbreak of Zika virus and lastly in 2021 there was a case found of Nipah virus. So it can be a next outbreak also.

The only method to quickly improve public health and contain the current pandemic is through vaccinations. As Nipah virus attacks continuously and the results of it can be very serious like effects in brain, coma and even death. So as a prevention making of a vaccine of NiV is very important. Here we chose immunoinformatics method to design a vaccine model. The benefits of this method is helping with the urgently needed quick detection of antigenic patterns from pathogen immune response-inducing proteins.

The vaccine that we have created is currently in the primary stage; there are many areas need to improve like the coverage of the vaccine model, Ramachandran outlier can be reduced to have more developed vaccine model. After that a significant amount of more research is required to develop this vaccine, including clinical trials and in vivo and in vitro testing.

References

1. Aditi, & Shariff, M. (2019). Nipah virus infection: A review. *Epidemiology and Infection*, *147*. https://doi.org/10.1017/S0950268819000086
2. Aguilar, H. C., & Lee, B. (2011). Emerging paramyxoviruses: Molecular mechanisms and antiviral strategies. *Expert Reviews in Molecular Medicine*, *13*. https://doi.org/10.1017/S1462399410001754
3. Ahmadi, F., Dorosti, H., Ghasemi, Y., & Nezafat, N. (2020). In Silico Design of Epitope-Based Allergy Vaccine Against Bellatella germanica Cockroach Allergens. *International Journal of Peptide Research and Therapeutics*, *26*(4), 1739–1749. https://doi.org/10.1007/S10989-019-09980-7
4. Alexander, J., Fikes, J., Hoffman, S., Franke, E., Sacci, J., Appella, E., Chisari, F. V., Guidotti, L. G., Chesnut, R. W., Livingston, B., & Sette, A. (1998). The optimization of helper T lymphocyte (HTL) function in vaccine development. *Immunologic Research*, *18*(2), 79–92. https://doi.org/10.1007/BF02788751
5. Bell, J. K., Askins, J., Hall, P. R., Davies, D. R., & Segal, D. M. (2006). The dsRNA binding site of human Toll-like receptor 3. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(23), 8792–8797. https://doi.org/10.1073/pnas.0603245103
6. Bharaj, P., Wang, Y. E., Dawes, B. E., Yun, T. E., Park, A., Yen, B., Basler, C. F., Freiberg, A. N., Lee, B., & Rajsbaum, R. (2016). *The Matrix Protein of Nipah Virus Targets the E3-Ubiquitin Ligase TRIM6 to Inhibit the IKKε Kinase-Mediated Type-I IFN Antiviral Response*. https://doi.org/10.1371/journal.ppat.1005880
7. Blocquel, D., Beltrandi, M., Erales, J., Barbier, P., & Longhi, S. (2013). Biochemical and structural studies of the oligomerization domain of the Nipah virus phosphoprotein: Evidence for an elongated coiled-coil homotrimer. *Virology*, *446*(1–2), 162–172.

https://doi.org/10.1016/j.virol.2013.07.031

8. Bruhn, J. F., Barnett, K. C., Bibby, J., Thomas, J. M. H., Keegan, R. M., Rigden, D. J., Bornholdt, Z. A., & Saphire, E. O. (2014). Crystal Structure of the Nipah Virus Phosphoprotein Tetramerization Domain. *Journal of Virology*, *88*(1), 758–762. https://doi.org/10.1128/jvi.02294-13
9. Castiglione, F, & Bernaschi, M. (2004). C-immsim: playing with the immune response. *Proceedings of the Sixteenth …*, *August 2004*, 1–7. http://www.math.ucsd.edu/~helton/MTNSHISTORY/CONTENTS/2004LEUVEN/CDROM/papers/316.pdf
10. Castiglione, Filippo. (n.d.). *Filippo Castiglione C-ImmSim: a model of immunological models C-ImmSim: a model of immunological models*.
11. Da Silva, J., & Hughes, A. L. (1998). Conservation of cytotoxic T lymphocyte (CTL) epitopes as a host strategy to constrain parasite adaptation: evidence from the nef gene of human immunodeficiency virus 1 (HIV-1). *Molecular Biology and Evolution*, *15*(10), 1259–1268. https://doi.org/10.1093/OXFORDJOURNALS.MOLBEV.A025854
12. Doytchinova, I. A., & Flower, D. R. (2007). VaxiJen: A server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics*, *8*(1), 1–7. https://doi.org/10.1186/1471-2105-8-4/TABLES/2
13. IMMUNOLOGY OF LAGOMORPHS. (1998). *Handbook of Vertebrate Immunology*, 223–260. https://doi.org/10.1016/B978-012546401-7/50008-1
14. Kasloff, S. B., Leung, A., Pickering, B. S., Smith, G., Moffat, E., Collignon, B., Embury-Hyatt, C., Kobasa, D., & Weingartl, H. M. (2019). Pathogenicity of Nipah henipavirus Bangladesh in a swine host. *Scientific Reports*, *9*(1). https://doi.org/10.1038/S41598-019-40476-Y

15. Kawasaki, T., & Kawai, T. (2014). Toll-like receptor signaling pathways. *Frontiers in Immunology*, *5*(SEP), 461. https://doi.org/10.3389/FIMMU.2014.00461/BIBTEX
16. Maisner, A., Ringel, M., Behner, L., Heiner, A., Sauerhering, L., & Maisner, A. (2018). The Journal of Infectious Diseases Replication of a Nipah Virus Encoding a Nuclear-Retained Matrix Protein. *The Journal of Infectious Diseases ®*, *221*(S4), 389–394. https://doi.org/10.1093/infdis/jiz440
17. Mohammed, A. A., Shantier, S. W., Mustafa, M. I., Osman, H. K., Elmansi, H. E., Osman, I. A. A., Mohammed, R. A., Abdelrhman, F. A., Elnnewery, M. E., Yousif, E. M., Mustafa, M. M., Elfadol, N. M., Abdalla, A. I., Mahmoud, E., Yagaub, A. A., Ahmed, Y. A., & Hassan, M. A. (2020). Epitope-Based Peptide Vaccine against Glycoprotein G of Nipah Henipavirus Using Immunoinformatics Approaches. *Journal of Immunology Research*, *2020*. https://doi.org/10.1155/2020/2567957
18. *Nipah Virus (NiV) | CDC*. (n.d.). Retrieved June 8, 2022, from https://www.cdc.gov/vhf/nipah/index.html
19. *Nipah Virus Infection - Causes, Symptoms, Treatments and Preventions | TRUTEST Laboratories*. (n.d.). Retrieved June 8, 2022, from https://www.trutestlab.com/blog/nipah-virus-infection-causes-symptoms-treatments-and-preventions
20. Olive, D., Fleri, W., Paul, S., Kumar Dhanda, S., Mahajan, S., Xu, X., Peters, B., & Sette, A. (2017). The Immune Epitope Database and Analysis Resource in Epitope Discovery and Synthetic Vaccine Design. *Front. Immunol*, *8*, 278. https://doi.org/10.3389/fimmu.2017.00278
21. Pickett, B. E., Greer, D. S., Zhang, Y., Stewart, L., Zhou, L., Sun, G., Gu, Z., Kumar, S., Zaremba, S., Larsen, C. N., Jen, W., Klem, E. B., & Scheuermann, R. H. (2012). Virus pathogen database and analysis resource (ViPR): a comprehensive bioinformatics

database and analysis resource for the coronavirus research community. *Viruses*, *4*(11), 3209–3226. https://doi.org/10.3390/V4113209

22. Pickett, B. E., Sadat, E. L., Zhang, Y., Noronha, J. M., Squires, R. B., Hunt, V., Liu, M., Kumar, S., Zaremba, S., Gu, Z., Zhou, L., Larson, C. N., Dietrich, J., Klem, E. B., & Scheuermann, R. H. (2012). ViPR: an open bioinformatics database and analysis resource for virology research. *Nucleic Acids Research*, *40*(Database issue), D593. https://doi.org/10.1093/NAR/GKR859
23. Raveendran, A. V., Sadanandan, S., Thulaseedharan, N. K., Sajeeth Kumar, K. G., Pallivalappil, B., & Anoop Kumar, A. S. (2002). Nipah Virus Infection: Pathology and Pathogenesis of an Emerging Paramyxoviral Zoonosis. *The American Journal of Pathology*, *161*(6), 2153–2167. https://doi.org/10.1016/S0002-9440(10)64493-8
24. Ringel, M., Heiner, A., Behner, L., Halwe, S., Sauerhering, L., Becker, N., Dietzel, E., Sawatsky, B., Kolesnikova, L., & Maisner, A. (2019). Nipah virus induces two inclusion body populations: Identification of novel inclusions at the plasma membrane. *PLoS Pathogens*, *15*(4). https://doi.org/10.1371/JOURNAL.PPAT.1007733
25. Schneidman-Duhovny, D., Inbar, Y., Nussinov, R., & Wolfson, H. J. (2005). PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Research*, *33*(suppl_2), W363–W367. https://doi.org/10.1093/NAR/GKI481
26. Shariff, M. (n.d.). *Epidemiology and Infection Nipah virus infection: A review*. https://doi.org/10.1017/S0950268819000086
27. Silakari, O., & Singh, P. K. (2021). Molecular docking analysis: Basic technique to predict drug-receptor interactions. *Concepts and Experimental Protocols of Modelling and Informatics in Drug Design*, 131–155. https://doi.org/10.1016/B978-0-12-820546-4.00006-4
28. Singh, R. K., Dhama, K., Chakraborty, S., Tiwari, R., Natesan, S., Khandia, R., Munjal,

A., Vora, K. S., Latheef, S. K., Karthik, K., Singh Malik, Y., Singh, R., Chaicumpa, W., & Mourya, D. T. (2019). Nipah virus: epidemiology, pathology, immunobiology and advances in diagnosis, vaccine designing and control strategies – a comprehensive review. *The Veterinary Quarterly*, *39*(1), 26. https://doi.org/10.1080/01652176.2019.1580827

29. Skowron, K., Bauza-Kaszewska, J., Grudlewska-Buda, K., Wiktorczyk-Kapischke, N., Zacharski, M., Bernaciak, Z., & Gospodarek-Komkowska, E. (2022). Nipah Virus–Another Threat From the World of Zoonotic Viruses. *Frontiers in Microbiology*, *12*, 4267. https://doi.org/10.3389/FMICB.2021.811157/BIBTEX
30. 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. https://doi.org/10.1016/J.JBI.2014.11.003
31. Stanzione, F., Giangreco, I., & Cole, J. C. (2021). Use of molecular docking computational tools in drug discovery. *Progress in Medicinal Chemistry*, *60*, 273–343. https://doi.org/10.1016/BS.PMCH.2021.01.004
32. Sun, B., Jia, L., Liang, B., Chen, Q., & Liu, D. (2018). Phylogeography, Transmission, and Viral Proteins of Nipah Virus. *Virologica Sinica*, *33*(5), 385. https://doi.org/10.1007/S12250-018-0050-1
33. Vita, R., Mahajan, S., Overton, J. A., Kumar Dhanda, S., Martini, S., Cantrell, J. R., Wheeler, D. K., Sette, A., & Peters, B. (2018). The Immune Epitope Database (IEDB): 2018 update. *Nucleic Acids Research*, *47*, 339–343. https://doi.org/10.1093/nar/gky1006
34. Ye, W., Schmitt, N. C., Ferris, R. L., & Allen, C. T. (2020). Improving responses to immunotherapy in head and neck squamous cell carcinoma. *Improving the Therapeutic*

Ratio in Head and Neck Cancer, 107–133. https://doi.org/10.1016/B978-0-12-817868-3.00006-8

35. Zheng, X., Li, S., & Yang, H. (2021). Roles of Toll-Like Receptor 3 in Human Tumors. *Frontiers in Immunology*, *12*, 1368. https://doi.org/10.3389/FIMMU.2021.667454/BIBTEX