A Novel Vaccine Designing Method Using In-silico Tools Against Scaffold Protein OPG 125 of Monkeypox Virus

By

Sadia Zafur Jannati

ID: 20146082

A thesis submitted to the School of Pharmacy in partial fulfillment
of the requirements for the degree of Bachelor of Pharmacy (Hons.)

School of Pharmacy

Brac University

April 2024

Declaration

It is hereby declared that,

1. The thesis submitted is my own original work while completing bachelor's 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 reference.

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. I have acknowledged all main sources of help.

Sadia Zafur Jannati

ID:20146082

Approval

The thesis titled "A Novel Vaccine Designing Method Using In-silico Tools against Scaffold Protein OPG 125 of Monkey Pox Virus" submitted by Sadia Zafur Jannati (20146082) of Spring 2020 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of Bachelor of Pharmacy (Hons.) on May 2, 2024.

Supervised By:

Supervisor:

Mohammad Kawsar Sharif Siam
Senior Lecturer
School of Pharmacy
Brac University

Approved By:

Program Director:

Professor Dr. Hasina Yasmin
Assistant Dean & Program Director
School of Pharmacy
Brac University

Departmental Head:

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

Ethics Statement

The study does not involve any kind of animal trial or human trial.

Abstract:

Recently, the number of patients with monkeypox disease is increasing at an alarming rate which has made the necessity of inventing a vaccine against this disease. As a result, this study is performed in order to design a vaccine model against the disease based on in silico computational tools. Scaffold protein OPG 125 has been finalized as the primary protein of this vaccine. CTL, HTL and B cell epitopes are filtered against their antigenic, allergic and toxic elements. The tool Vaxijen v2.0 has identified a 0.5312 antigenic score which has been increased from 0.5352. While evaluating biochemical properties, physical and chemical features were also examined. It shows a negative z score and gravy score which is required for this process. Additionally, 3D model assessment has also been carried out and it provides 80% coverage with a confidence of 100%. In the end molecular docking and immune simulation is performed by using Cluspro and C-ImmSimm.

Keywords: Monkeypox virus; In silico computational tools; Vaccine construction; Protein

Dedication

I want to dedicate this project to my family who supported me and stood by me all the time.

Acknowledgement

Firstly, I would like to thank to the Almighty Allah for everything. Moreover, I would like to thank my supervisor Mohammad Kawsar Sharif Siam, Senior Lecturer, School of Pharmacy, Brac University. Without his guidance it would be difficult to complete my project. His continuous assistance and feedback make my journey easier. Also I would like to thank my family, seniors and all my friends who were always by my side and support me throughout my journey.

Sadia Zafur Jannati

ID: 20146082

School of Pharmacy

Brac University

Table of Contents

List of Tables

List of Figures

List of Acronyms

CTL	Cytotoxic T Lymphocyte
GISAID	Global Initiative on Sharing All Influenza Data
HTL	Helper T Lymphocyte
MPXV	Monkey Pox Virus
MHC	Major Histocompatability Complex
ViPR	Virus Pathogen Resource

Chapter 1

Introduction

Monkeypox virus is a zoonotic orthopox DNA virus which initially emerged in humans in 1970 throughout the Democratic Republic of Congo (Thornhill et al., 2022). The *Orthopoxvirus* genus also includes other types of viruses which are cowpox, camelpox, variola and vaccinia. However, the World Health Organization confirmed monkeypox as the most widespread *Orthopoxvirus* in humans since smallpox eradication in 1980 (McCollum & Damon, 2013). After the first human case intermittent outbreaks were observed in West and Central Africa, mostly among rainforest-dwelling children (Mitjà et al., 2023). Though it was primarily found in the densely forested regions of West and Central Africa, there has been a recent global spread of the virus, resulting in outbreaks in several countries where it was not previously endemic. Additionally, the current outbreak of Mpox throughout the world began in May of 2022, and on July 23, 2022, it was designated a public health emergency of international concern. By August 2, 2023, 113 nations had 88,600 laboratory – confirmed cases and 152 deaths including 106 countries that had never reported mpox and among them immunocompromised patients and pregnant women are at higher risk (Alakunle et al., 2024). There are some treatments for monkeypox virus which includes drugs which are tecovirimat (orally or intravenously), brincidofovir (orally), cidofovir (intravenously or topically) and vaccinia immunoglobulin. However, these drugs have a lot of complexity (See, 2022). Moreover, there are two vaccines for monkeypox virus which are not efficacious and these are ACAM2000 and JYNNEOS. ACAM2000 has the risk of occurring myopericarditis and also other side effects additionally, to confirm JYNNNEOS's efficacy further testing is needed. Both of these vaccines are mainly licensed for the treatment of smallpox virus but they have the ability to reduce mpox infection. Despite the alarming rise of mpox patients, still there is no effective

treatment or therapy that is specifically approved by FDA for this virus (Ashley et al., 2024). Therefore, effective vaccinations are needed in order to lessen the mpox infection among individuals.

1.1 Genomic Characteristics and Attributes of Monkeypox Virus

Monkeypox virus belongs to the *Poxviridae* family and is composed of double stranded DNA genomes. This genome has a quite large length (approximately 197 kb) with a complex structure. Additionally, it contains 200 non-overlapping coding genes resulting in unique biological traits. In its genome there are two distinct clades which are clade I and clade II (Luna et al., 2023). Mpox Clade I isolates have a more homogeneous genome length (196–199 Kbp) than those in Clade II isolates (196–211 Kbp). Though gene synteny of the two clades are almost identical but there are some differences in terms of the contents (Alakunle et al., 2024). Clade I isolate expresses features like - significant clinical signs, increased transmissibility and increased rate of mortality. In contrast, clade II has two subgroups - subclade II a which shows minimal clinical signs, lower transmissibility and lower rate of mortality and subclade II b which contains current monkeypox cases spread in humans (Luna et al., 2023). Moreover, subclade II b has branched into some lineages (A, A.1, A.2, A.3, B.1) in accordance with the GISAID (Global Initiative on Sharing All Influenza Data) (Alakunle et al., 2024b). Among these lineages B.1 belongs to the strains of current MPXV cases. Additionally, this lineage exhibits a significant degree of divergence as a result of the numerous epidemiological occurrences that were responsible for the outbreaks that occurred in the past (Luna et al., 2023). To clearly understand the attributes of monkeypox virus one figure is attached below-

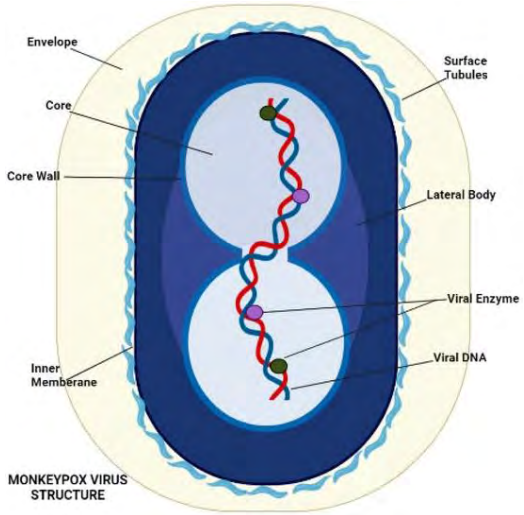

Figure 1*: Structure of Monkeypox virus* (Anwar et al., 2023)

1.2 Replication Cycle and Pathological Process of Monkeypox Virus

The replication process of MPXV mainly starts within the cytoplasm of any living cell. After that it attaches with the membrane of the living cell and creates a hole within the cell as a result central components leaked out. These components are packed together to form intracellular mature viruses (IMVs) and remain in the cytoplasm. When cellular disruption takes place these IMVs leave the cytoplasm as extracellular enveloped viruses. In terms of mature viruses they create an extra layer and leak out from the cell through exocytosis (Rabaan et al., 2023).

Among all the living cells, the human body is the most crucial host cell for the monkeypox virus and it has several ways to get inside the human body such as nasopharyngeal, oropharyngeal and intradermal route. Monkeypox viruses mainly work by destroying the pattern recognition receptors which are produced from the immune cells of the host body. These receptors bind with the ligand and activate some factors which produce inflammation, and these are- NF-κB, IRFs, AP-1 etc. intracellular adaptor proteins including MyD88, MAL, TRAM, TRIF, and SARM are essential for TLR signal transduction and intracellular immune responses. Any change in adapter proteins may

impair viral infection immunology. This disrupts the normal functioning of adaptor proteins, leading to the suppression of NF-κB, a transcription factor linked to inflammation. Consequently, the innate immune systems fail to recognize viruses due to this disease (Harapan et al., 2022).

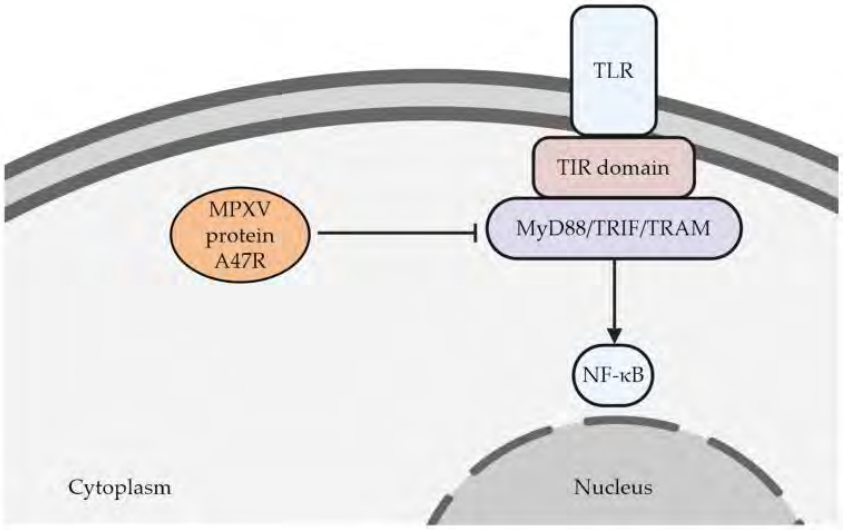

Figure 2*: Mechanism of Action of MPXV* (Harapan et al., 2022)

Chapter 2

Methodology

Vaccine design through in silico technique is done by some computational tools to predict and design a vaccine of any diseases. The whole process is done by a sequence of computational tools and the whole sequence is shown below-

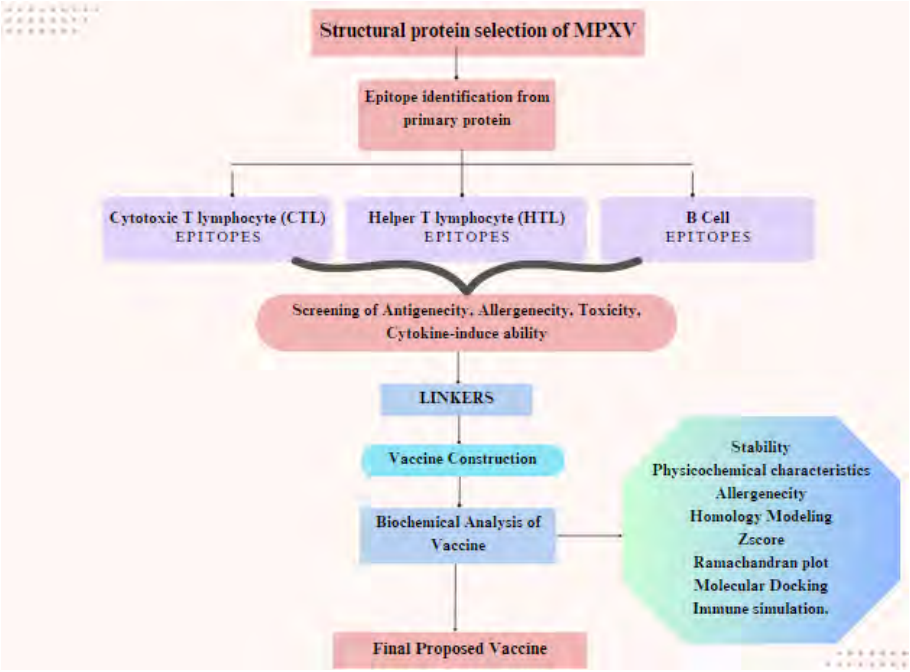

Figure 3*: Visual presentation of MPXV vaccine design by in silico technique.*

2.1 Protein Selection of MPXV

Primary protein of any disease can be selected by using various servers such as - NCBI (National Center for Biotechnology Information), ViPR (Virus Pathogen Resource), UniProtKB (The Universal Protein KnowledgeBase) etc. In this study for the selection of protein UniProtKB is used. UniprotKB is a large protein data bank as it contains information about protein's sequences and functions and is assembled with a set of databases (Apweiler et al., 2004). While selecting

protein only structural proteins were selected and all enzymatic proteins were dropped out. Finally, scaffold protein OPG 125 has been used to construct the vaccine for monkeypox virus. The amino acid length of this protein is 551 and it allows the initiation of transcription from late gene promoters by interacting with RNA polymerase. After selecting the appropriate protein the most important thing is to check whether the protein sequence contains any antigenic particles or not. Consequently, to check the antigenicity of the protein sequence Vaxijen v2.0 tool was used and a 0.5 threshold was set.

2.2 Selection of CTL (Cytotoxic T Lymphocyte) Epitopes:

CTL epitopes are selected using the NetCTL 1.2 tool from the protein sequence. This server contains information about MHC class-1 binding affinity, C-terminal proteasomal cleavage and tap transport efficiency. Mainly, it predicts the binding of Major Histocompatibility complex (MHC) class-I to the 12 supertypes of MHC (G et al., 2020). In this server primary protein sequence was submitted in fasta format and combined score were chosen in the sort by score part and other information was chosen by default. After getting appropriate alleles from the NetCTL 1.2 tool these peptides were submitted to another tool which was NETMHCpan 4.1. This server helps to predict the binding of MHC-I peptides with the specific CTL epitopes obtained from the previous server. In this server the peptide length has to be set at 9 mer peptides after that all the alleles are selected and have to put a tick mark on include BA (Binding Affinity) prediction. From this server only peptides with strong binding affinity are chosen. Furthermore, these peptides are screened out based on their antigenicity, allergenicity and toxicity. All peptides must have to be antigen, non-allergen and nontoxic. For checking antigenicity “Vaxijen v2.0”, for allergenicity “AllerTOP v.2.0” and for toxicity “ToxinPred” tool has been used.

2.3 Selection of HTL (Helper T Lymphocyte) Epitopes:

In terms of selecting the HTL epitopes NetMHCIIpan 4.0 tool has been used. It works by predicting the binding of peptides with MHC class-II alleles and is done through ANNs (Artificial Neural Networks) (Reynisson et al., 2020). In the server protein sequence is submitted in fasta format and 15 mer peptide is selected for peptides. After that only 20 alleles can be selected at a time and have to select the section; include BA prediction. Then again the rest of the alleles are selected and the other processes are the same. Peptides that express higher binding affinity only are selected for further selection. Additionally, ability of inducing cytokines are checked for these alleles since helper T cells have the property of inducing cytokines. On the contrary, interferon gamma (IFN-gamma), Interleukin-4 (IL4), Interleukin-10 (IL10) cytokines inducing ability are checked. Consequently, alleles showing positive results for IFN-gamma and inducers for IL4 and IL10 cytokines are selected for an additional screening process. These selections are done through various computational tools; “IFNepitope " for IFN-gamma, “IL4Pred " for IL4 inducers and “IL10Pred” for IL10 inducers. In addition, antigenicity, allergenicity and toxicity are checked of these inducers through the same servers used in case of CTL epitopes. Finally, alleles having antigenic, non-allergenic and nontoxic properties are screened for vaccine constructions.

2.4 Selection of B-cell Epitopes:

In the same manner as the previous epitopes were selected for the selection of B cell epitopes a computational tool has been used. To choose B cell epitopes “IEDB Analysis Resource” server has been used. In this server there is a section of “B cell epitope prediction” where a plain format of protein sequence is submitted. Epitopes having length from 7 to 25 are taken and antigenicity, allergenicity and toxicity are checked for these epitopes using the same servers as CTL epitopes. Epitopes which have a very low length cannot be taken because they don’t possess any effect as

well epitopes containing long length can produce toxicity or make the vaccine less effective as a result only those that have length 7 to 25 are chosen for vaccine construction. Epitopes that show antigen, non-allergen and non-toxic are further used to construct the vaccine.

2.5 Construction of MPXV Vaccine:

Vaccine is constructed by linking all the epitopes and the protein sequence. After selecting appropriate epitopes they are linked together and in between all the epitopes linkers are used to connect them. Linker EAAAK is used to connect the adjuvant with CTL epitopes. Additionally, AAY linker has been used to connect the CTL epitopes with each other. Moreover, to link CTL epitopes with HTL epitopes as well as to link HTL epitopes GPGPG linker is applied. Furthermore, for B Cell epitopes KK linker has been added (Martinelli, 2022).

2.6 Identification of Antigenicity, Allergenicity & Toxicity of the Constructed Vaccine:

At the end of vaccine construction, the first and foremost thing is to check the vaccine's antigenicity, allergenicity and toxicity. These are done by using various computational tools. "Vaxijen v.2.0" is used in order to verify the vaccine's antigenicity by setting a threshold of 0.5 (Doytchinova & Flower, 2007). Additionally, "AllergenOnline" is used to check allergenicity of the constructed vaccine. This server helps in recognition of allergenic cross reactions among the protein sequence and thus ensure safety of any protein sequence (Sircar et al., 2014).

Finally, assessment of toxic entities in the protein sequence is very crucial as it may deteriorate health. Identification of toxic entities in vaccines is done through the "T3DB (The Toxin and Toxin Target Database" server. This database contains information about toxin entities including their mechanism and reaction with any drugs (Wishart et al., 2014).

2.7 Evaluation of the Constructed Vaccine Based on Biochemical Features:

Biochemical features of the vaccine mainly include all the physical and chemical properties of the vaccine which is identified by the "ProtParam" tool. The physical and chemical properties include molecular weight of the protein sequence, number of amino acids, theoretical pI, amino acid composition, atomic composition, total number of atoms, estimated half-life, instability index, gravy and aliphatic index. Theoretical pI is basically the pH at which vaccine will be electrically neutral. In addition, instability index refers to the stability of the vaccine and it must have to be stable. Moreover, gravy abbreviates the grand average of hydropathicity. If the gravy score is positive it means the vaccine is hydrophobic in nature which is not good for our body. On the other hand, if the gravy score is negative it indicates that it has a better absorption ability and it will be less toxic.

2.8 Forecast of 3D Model of the Vaccine:

Phyre2 is a computational tool which is one of the best for the prediction of the 3D model for any protein sequence. Besides, it also anticipates functions and mutation of the protein sequence and also analyzes them. However, it can also forecast the binding site of the ligand of the protein sequence and evaluates the consequences of amino acid variants (Kelley et al., 2015). It shows results based on confidence and coverage in a percentage. Confidence refers to the reliance of predicting the right model and coverage means the ability of the server to predict the actual model. The protein sequence of the vaccine is given on the Phyre2 server and it gives the result as a pdB file through email which is provided along with the protein sequence.

2.9 Assessment of the Validity of 3D Model:

The SWISS-MODEL is a tool that allows for the automated 3D modeling of protein structures. This server can provide an authentic 3D model which has the highest similarity with the protein

sequences. It works through three interaction modes which are- first approach mode, alignment mode and project mode. Moreover, the modeling method mainly includes selection of the template, sequencing of target template, model building & finally, assessment (Schwede, 2003).

In addition, results are analyzed based on the ramachandran plot which gives a plain scenario on the protein structure and has φ-ψ angles which is shown in the graph (Ho & Brasseur, 2005). Moreover, 4 types of ramachandran regions are distinguished in the swiss-model expasy tool. Among these ramachandran favor and ramachandran outlier regions are mainly considered for vaccine construction.

2.10 Quality Evaluation of 3D Model:

ProSA-web is a computational tool which states protein structure analysis; identifies quality of 3D model. The ProSA-web application aims to motivate structure depositors to validate their structures (Wiederstein & Sippl, 2007). A pdB file which is given in the mail of Phyre2 server is submitted in the pro-SA web tool thus a z score value is obtained. The z score must have to give a negative score as the negative score of z score implies the amount of energy released from the model. The more negative the value is, the more energy is released, and the model is more stable.

2.11 Evaluation of Protein-Ligand Binding Interaction:

After constructing a vaccine, it is very important to check whether it binds with the body receptor or not. So, to assess the protein-ligand binding molecular docking is done by using various tools. In this study molecular docking is done through cluspro server. In the cluspro server two pdb files are required, one is for ligand and another one is for receptor. Ligand pdb file is found from Phyre2 mail and receptor pdb file is found in the RCSB PDB server which is a protein data bank. The output comprises 10 models of the ensuing structure that is produced after the interaction (Alekseenko et al., 2020). The best model is chosen based on the cluster energy among the models.

2.12 Determination of Vaccine's Immune Response:

In order to evaluate the immunological response provided by the in silico-based vaccine "C-IMMSIMM" web tool has been used. In addition, to characterize the immunogenicity and immunological response as well as to ensure the use of PSSM (position-specific scoring matrices) employed by machine learning, this server plays an important role (Motamedi et al., 2023). This also determines the dosing sequence of vaccines in the human body. It provides several graphs to show the dosing of the vaccine.

2.13 Remarks Concerning the Methodology:

The main aim of this research is to design a multiepitope based vaccine by using in silico computational tools. Along with these tools a complete evaluation of some relevant literature has been also performed based on the necessity.

Chapter 3

Results:

3.1 Antigenicity Identification of the Selected Protein:

Following the screening process of protein strains, scaffold protein OPG 125 has been selected finally for vaccine construction. Amino acid sequence of this protein is given below-

```
MNNTIINSLIGGDDFIKRSNVFAVDSQIPTLYMPQYISLSGVMTNDGPDNQAIASFEIRD
QYITALNHLVLSLELPEVKGMGRFGYVPYVGYKCINHVSVSSCNGVIWEIEGEELYNNCI
NNTIALKHSGYSSELNDISIGLTPNDTIKEPSTVYVYIKTPFDVEDTFSSLKLSDSKITV
TVTFNPVSDIVIRDSSFDFETFNKEFVYVPELSFIGYMVKNVQIKPSFIEKPRRVIGQIN
QPTATVTEVHAATSLSVYTKPYYGNTDNKFISYPGYSQDEKDYIDAYVSRLLDDLVIVSD
GPPTGYPESAEIVEVPEDGVVSIQDADVYVKIDNVPDNMSVYLHTNLLMFGTRKNSFIYN
ISKKFSAITGTYSDATKRTVFAHISHTINIIDTSIPVSLWTSQRNVYNGDNRSAESKAKD
LFINDPFIKGIDFKNKTDIISRLEVRFGNDVLYSENGPISRIYNELLTKSNNGTRTLTFN
FTPKIFFRPTTITANVSRGKDKLSVRVVYSTMDINHPIYYVQKQLVVVCNDLYKVSYDQG
VSITKIMGDNN
```

This protein has an antigen score of 0.5352 (Probable) with a threshold of 0.5 which has been identified through Vaxijen v2.0.

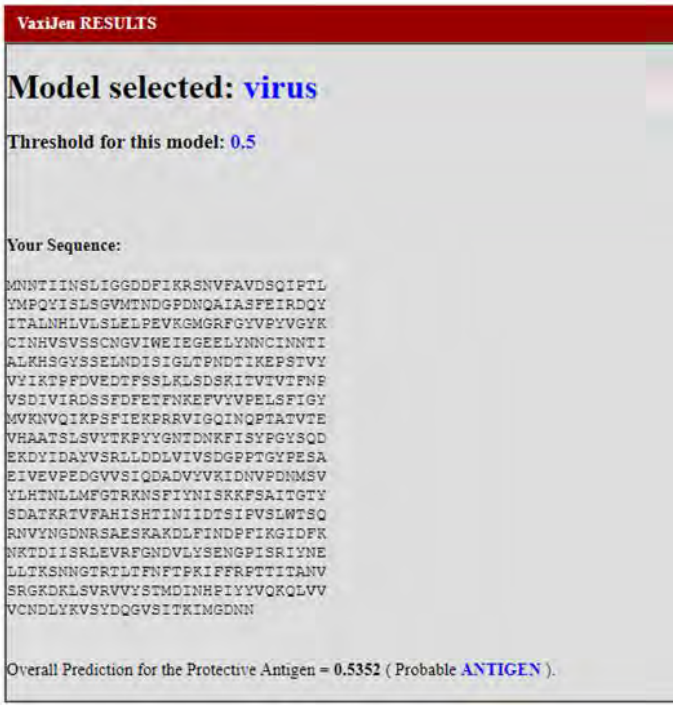

Figure 4*: Antigenicity score of Scaffold protein OPG 125* (Doytchinova & Flower, 2007)

3.2 CTL Epitopes Selection:

19 CTL epitopes have been identified from NetCTL1.2 server among them 13 epitopes have revealed strong bindings with the MHC class I peptides. On the contrary, these epitopes were selected based on antigenicity, allergenicity and toxicity.

Table 1: *CTL Epitopes along with combined score.*

CTL Epitopes	Combined score	Sensitivity	Specificity
NTDNKFISY	3.6066	0.54	0.993
TMDINHPIY	3.0493	0.54	0.993
VSDGPPTGY	2.5041	0.54	0.993
HAATSLSVY	1.8504	0.54	0.993
GTRKNSFIY	1.4676	0.54	0.993
LVVVCNDLY	1.3488	0.54	0.993
TIALKHSGY	1.3052	0.54	0.993
SLSVYTKPY	1.2364	0.70	0.985
NVPDNMSVY	1.1929	0.70	0.985
ETFNKEFVY	1.1786	0.70	0.985
YSDATKRTV	1.1509	0.70	0.985
YIDAYVSRL	1.1124	0.70	0.985

VSIQDADVY	1.0917	0.70	0.985
LSVYTKPYY	1.0710	0.70	0.985
ASFEIRDQY	0.9950	0.74	0.980
CNDLYKVSY	0.9510	0.74	0.980
YLHTNLLMF	0.8728	0.80	0.970
ITALNHLVL	0.7899	0.80	0.970
VDSQIPTLY	0.7591	0.80	0.970

***Table 2**: Eventually screened CTL Epitopes.*

	PEPTIDES	ANTIGENICITY	ALLERGENICITY	TOXICITY
1	ASFEIRDQY	Antigen	Allergen	Non Toxin
2	ETFNKEFVY	Non Antigen	Allergen	Non Toxin
3	HAATSLSVY	Non Antigen	Non Allergen	Non Toxin
4	NTDNKFISY	Antigen	Non Allergen	Non Toxin
5	NVPDNMSVY	Non Antigen	Non Allergen	Non Toxin
6	SLSVYTKPY	Antigen	Non Allergen	Non Toxin
7	TIALKHSGY	Antigen	Non Allergen	Non Toxin
8	TMDINHPIY	Non Antigen	Allergen	Non Toxin
9	VDSQIPTLY	Non Antigen	Allergen	Non Toxin
10	VSDGPPTGY	Non Antigen	Allergen	Non Toxin
11	VSIQDADVY	Antigen	Allergen	Non Toxin
12	YIDAYVSRL	Non Antigen	Non Allergen	Non Toxin
13	YLHTNLLMF	Non Antigen	Non Allergen	Non Toxin

3.3 HTL Epitopes Identification:

In the NetMHCIIpan 4.0 server 112 epitopes have been found with strong binding accompanied by MHC class II alleles. Then, the epitopes were chosen based on the cytokine inducing ability. As a result, 14 epitopes are selected based on these criteria. Finally, they were screened out through antigenicity, allergenicity and toxicity testing where only 3 HTL epitopes have been passed and eventually selected for vaccine construction.

***Table 3**: Extracted HTL Epitopes*.

	Peptides	IFN- gamma	IL4 Cytokines	IL10 Cytokines	Anitigenicity	Allergenicity	Toxicity
1	ADVYVKIDNVPDNMS	NEGATIVE	IL4 inducer	IL10 non-inducer			
2	ATSLSVYTKPYYGNT	NEGATIVE	Non IL4 inducer	IL10 non-inducer			
3	ATVTEVHAATSLSVY	POSITIVE	IL4 inducer	IL10 non-inducer			
4	DADVYVKIDNVPDNM	NEGATIVE	IL4 inducer	IL10 non-inducer			
5	DDFIKRSNVFAVDSQ	NEGATIVE	IL4 inducer	IL10 non-inducer			
6	DDLVIVSDGPPTGYP	NEGATIVE	Non IL4 inducer	IL10 non-inducer			
7	DEKDYIDAYVSRLLD	POSITIVE	Non IL4 inducer	IL10 inducer			
8	DGVVSIQDADVYVKI	POSITIVE	Non IL4 inducer	IL10 non-inducer			
9	DINHPIYYVQKQLVV	NEGATIVE	Non IL4 inducer	IL10 inducer			
10	DLVIVSDGPPTGYPE	NEGATIVE	Non IL4 inducer	IL10 non-inducer			
11	DLYKVSYDQGVSITK	NEGATIVE	Non IL4 inducer	IL10 non-inducer			
12	DNKFISYPGYSQDEK	POSITIVE	IL4 inducer	IL10 inducer	Non Antigen	Non Allergen	Non Toxin
13	DQYITALNHLVLSLE	NEGATIVE	Non IL4 inducer	IL10 non-inducer			
14	DVYVKIDNVPDNMSV	NEGATIVE	IL4 inducer	IL10 non-inducer			
15	DYIDAYVSRLLDDLV	POSITIVE	Non IL4 inducer	IL10 inducer			
16	EDGVVSIQDADVYVK	POSITIVE	Non IL4 inducer	IL10 non-inducer			
17	EIRDQYITALNHLVL	NEGATIVE	Non IL4 inducer	IL10 non-inducer			
18	EKDYIDAYVSRLLDD	POSITIVE	Non IL4 inducer	IL10 inducer			
19	ENGPISRIYNELLTK	POSITIVE	IL4 inducer	IL10 inducer	Non Antigen	Allergen	Non Toxin
20	EPSTVYVYIKTPFDV	POSITIVE	IL4 inducer	IL10 inducer	Non Antigen	Non Allergen	Non Toxin
21	ESAEIVEVPEDGVVS	NEGATIVE	Non IL4 inducer	IL10 non-inducer			
22	FGNDVLYSENGPISR	NEGATIVE	IL4 inducer	IL10 inducer			
23	GDDFIKRSNVFAVDS	NEGATIVE	IL4 inducer	IL10 non-inducer			
24	GGDDFIKRSNVFAVD	NEGATIVE	IL4 inducer	IL10 non-inducer			
25	GNDVLYSENGPISRI	NEGATIVE	IL4 inducer	IL10 non-inducer			
26	GTRTLTFNFTPKIFF	NEGATIVE	Non IL4 inducer	IL10 non-inducer			
27	HTNLLMFGTRKNSFI	NEGATIVE	Non IL4 inducer	IL10 inducer			
28	IKPSFIEKPRRVIGQ	POSITIVE	IL4 inducer	IL10 non-inducer			
29	IKRSNVFAVDSQIPT	NEGATIVE	Non IL4 inducer	IL10 non-inducer			
30	IKTPFDVEDTFSSLK	NEGATIVE	Non IL4 inducer	IL10 non-inducer			
31	INHPIYYVQKQLVVV	POSITIVE	Non IL4 inducer	IL10 non-inducer			
32	INNTIALKHSGYSSE	NEGATIVE	Non IL4 inducer	IL10 inducer			
33	IPVSLWTSQRNVYNG	NEGATIVE	Non IL4 inducer	IL10 non-inducer			
34	IQDADVYVKIDNVPD	NEGATIVE	Non IL4 inducer	IL10 non-inducer			
35	IRDQYITALNHLVLS	NEGATIVE	Non IL4 inducer	IL10 non-inducer			

36	ISKKFSAITGTYSDA	NEGATIVE	IL4 inducer	IL10 inducer			
37	ISRIYNELLTKSNNG	NEGATIVE	IL4 inducer	IL10 inducer			
38	ISYPGYSQDEKDYID	POSITIVE	IL4 inducer	IL10 inducer	Non Antigen	Non Allergen	Non Toxin
39	KDYIDAYVSRLLDDL	POSITIVE	Non IL4 inducer	IL10 inducer			
40	KEPSTVYVYIKTPFD	POSITIVE	IL4 inducer	IL10 non-inducer			
41	KKFSAITGTYSDATK	NEGATIVE	IL4 inducer	IL10 non-inducer			
42	KNSFIYNISKKFSAI	POSITIVE	IL4 inducer	IL10 inducer	Non Antigen	Non Allergen	Non Toxin
43	KPRRVIGQINQPTAT	NEGATIVE	IL4 inducer	IL10 non-inducer			
44	KPSFIEKPRRVIGQI	POSITIVE	IL4 inducer	IL10 non-inducer			
45	KRSNVFAVDSQIPTL	NEGATIVE	IL4 inducer	IL10 non-inducer			
46	**KRTVFAHISHTINII**	**POSITIVE**	**IL4 inducer**	**IL10 inducer**	**Antigen**	**Non Allergen**	**Non Toxin**
47	KTPFDVEDTFSSLKL	NEGATIVE	Non IL4 inducer	IL10 non-inducer			
48	LHTNLLMFGTRKNSF	NEGATIVE	Non IL4 inducer	IL10 inducer			
49	LPEVKGMGRFGYVPY	NEGATIVE	Non IL4 inducer	IL10 inducer			
50	LSFIGYMVKNVQIKP	NEGATIVE	IL4 inducer	IL10 non-inducer			
51	LSVRVVYSTMDINHP	NEGATIVE	IL4 inducer	IL10 inducer			
52	LYKVSYDQGVSITKI	NEGATIVE	Non IL4 inducer	IL10 non-inducer			
53	LYMPQYISLSGVMTN	NEGATIVE	IL4 inducer	IL10 non-inducer			
54	MDINHPIYYVQKQLV	NEGATIVE	Non IL4 inducer	IL10 inducer			
55	MPQYISLSGVMTNDG	NEGATIVE	IL4 inducer	IL10 inducer			
56	MVKNVQIKPSFIEKP	NEGATIVE	IL4 inducer	IL10 non-inducer			
57	NGPISRIYNELLTKS	NEGATIVE	IL4 inducer	IL10 non-inducer			
58	NGTRTLTFNFTPKIF	NEGATIVE	Non IL4 inducer	IL10 non-inducer			
59	**NHPIYYVQKQLVVVC**	**POSITIVE**	**IL4 inducer**	**IL10 inducer**	**Antigen**	**Non Allergen**	**Non Toxin**
60	NISKKFSAITGTYSD	NEGATIVE	IL4 inducer	IL10 inducer			
61	NKFISYPGYSQDEKD	POSITIVE	IL4 inducer	IL10 inducer	Non Antigen	Non Allergen	Non Toxin
62	NNTIALKHSGYSSEL	NEGATIVE	Non IL4 inducer	IL10 inducer			
63	NSFIYNISKKFSAIT	POSITIVE	IL4 inducer	IL10 inducer	Non Antigen	Allergen	Non Toxin
64	NTDNKFISYPGYSQD	POSITIVE	IL4 inducer	IL10 inducer	Non Antigen	Non Allergen	Non Toxin
65	NVYNGDNRSAESKAK	NEGATIVE	Non IL4 inducer	IL10 non-inducer			
66	PESAEIVEVPEDGVV	NEGATIVE	Non IL4 inducer	IL10 non-inducer			
67	PGYSQDEKDYIDAYV	POSITIVE	Non IL4 inducer	IL10 inducer			
68	PQYISLSGVMTNDGP	NEGATIVE	IL4 inducer	IL10 non-inducer			
69	PRRVIGQINQPTATV	NEGATIVE	IL4 inducer	IL10 non-inducer			
70	PSFIEKPRRVIGQIN	POSITIVE	IL4 inducer	IL10 inducer	Non Antigen	Allergen	Non Toxin
71	PSTVYVYIKTPFDVE	POSITIVE	IL4 inducer	IL10 inducer	Non Antigen	Non Allergen	Non Toxin
72	PTTITANVSRGKDKL	NEGATIVE	Non IL4 inducer	IL10 non-inducer			
73	PVSDIVIRDSSFDFE	NEGATIVE	IL4 inducer	IL10 non-inducer			
74	PVSLWTSQRNVYNGD	NEGATIVE	Non IL4 inducer	IL10 non-inducer			
75	QDADVYVKIDNVPDN	POSITIVE	IL4 inducer	IL10 non-inducer			
76	QIKPSFIEKPRRVIG	POSITIVE	IL4 inducer	IL10 non-inducer			
77	QRNVYNGDNRSAESK	NEGATIVE	Non IL4 inducer	IL10 non-inducer			
78	RDQYITALNHLVLSL	NEGATIVE	Non IL4 inducer	IL10 non-inducer			
79	RFGNDVLYSENGPIS	NEGATIVE	IL4 inducer	IL10 non-inducer			
80	RFGYVPYVGYKCINH	NEGATIVE	IL4 inducer	IL10 inducer			
81	RKNSFIYNISKKFSA	POSITIVE	IL4 inducer	IL10 inducer	Non Antigen	Non Allergen	Non Toxin
82	RNVYNGDNRSAESKA	NEGATIVE	Non IL4 inducer	IL10 non-inducer			
83	RPTTITANVSRGKDK	POSITIVE	Non IL4 inducer	IL10 non-inducer			
84	RSNVFAVDSQIPTLY	NEGATIVE	Non IL4 inducer	IL10 non-inducer			
85	SAEIVEVPEDGVVSI	NEGATIVE	Non IL4 inducer	IL10 non-inducer			

86	SDIVIRDSSFDFETF	NEGATIVE	IL4 inducer	IL10 non-inducer			
87	SFIYNISKKFSAITG	NEGATIVE	IL4 inducer	IL10 non-inducer			
88	SIGLTPNDTIKEPST	NEGATIVE	Non IL4 inducer	IL10 non-inducer			
89	SKKFSAITGTYSDAT	NEGATIVE	IL4 inducer	IL10 inducer			
90	SNVFAVDSQIPTLYM	NEGATIVE	Non IL4 inducer	IL10 non-inducer			
91	SQRNVYNGDNRSAES	NEGATIVE	Non IL4 inducer	IL10 non-inducer			
92	SRIYNELLTKSNNGT	NEGATIVE	Non IL4 inducer	IL10 non-inducer			
93	STVYVYIKTPFDVED	POSITIVE	Non IL4 inducer	IL10 inducer			
94	SVRVVYSTMDINHPI	NEGATIVE	IL4 inducer	IL10 inducer			
95	SYPGYSQDEKDYIDA	POSITIVE	Non IL4 inducer	IL10 inducer			
96	TDNKFISYPGYSQDE	NEGATIVE	IL4 inducer	IL10 inducer			
97	TEVHAATSLSVYTKP	NEGATIVE	Non IL4 inducer	IL10 non-inducer			
98	TKPYYGNTDNKFISY	POSITIVE	IL4 inducer	IL10 non-inducer			
99	**TKRTVFAHISHTINI**	**POSITIVE**	**IL4 inducer**	**IL10 inducer**	**Antigen**	**Non Allergen**	**Non Toxin**
100	TRKNSFIYNISKKFS	NEGATIVE	IL4 inducer	IL10 inducer			
101	TRTLTFNFTPKIFFR	NEGATIVE	Non IL4 inducer	IL10 non-inducer			
102	TVTEVHAATSLSVYT	NEGATIVE	Non IL4 inducer	IL10 non-inducer			
103	TVTFNPVSDIVIRDS	NEGATIVE	Non IL4 inducer	IL10 non-inducer			
104	TVYVYIKTPFDVEDT	POSITIVE	Non IL4 inducer	IL10 inducer			
105	VKNVQIKPSFIEKPR	POSITIVE	IL4 inducer	IL10 non-inducer			
106	VSDIVIRDSSFDFET	NEGATIVE	IL4 inducer	IL10 non-inducer			
107	VSLWTSQRNVYNGDN	NEGATIVE	Non IL4 inducer	IL10 non-inducer			
108	VTEVHAATSLSVYTK	NEGATIVE	Non IL4 inducer	IL10 inducer			
109	VTVTFNPVSDIVIRD	NEGATIVE	Non IL4 inducer	IL10 non-inducer			
110	VYVKIDNVPDNMSVY	NEGATIVE	IL4 inducer	IL10 non-inducer			
111	VYVYIKTPFDVEDTF	NEGATIVE	Non IL4 inducer	IL10 inducer			
112	YIKTPFDVEDTFSSL	NEGATIVE	Non IL4 inducer	IL10 non-inducer			
113	YMPQYISLSGVMTND	NEGATIVE	IL4 inducer	IL10 inducer			
114	YPGYSQDEKDYIDAY	POSITIVE	Non IL4 inducer	IL10 inducer			
115	YTKPYYGNTDNKFIS	NEGATIVE	IL4 inducer	IL10 non-inducer			
116	YVKIDNVPDNMSVYL	NEGATIVE	IL4 inducer	IL10 inducer			

3.4 B cell Epitopes Finding:

Finding B cell Epitopes is easier than CTL and HTL epitopes since they are screened out based on antigenicity, allergenicity and toxicity only. From the B Cell Epitope Prediction tools a total of 23 epitopes were detected among them based on required length 12 epitopes were picked. Additionally, they were sorted by toxicity, allergenicity and antigenicity tests and consequently, 2 epitopes are approved to be used in vaccine preparation.

***Table 4**: Screened B cell epitopes*

NO.	Start	End	Peptide Sequence	Length	Antigenicity	Allergenicity	Toxicity
1	78	85	VKGMGRFG	8	Antigen	Allergen	Non Toxin
2	111	121	EGEELYNNCIN	11	Non Antigen	Allergen	Non Toxin
3	196	208	SFDFETFNKEFVY	13	Antigen	Allergen	Non Toxin
4	273	282	YPGYSQDEKD	10	Antigen	Allergen	Non Toxin
5	301	307	GPPTGYP	7	Non Antigen	Non Allergen	Non Toxin
6	355	363	NSFIYNISK	9	Non Antigen	Allergen	Non Toxin
7	402	419	SQRNVYNGDNRSAESKAK	18	Antigen	Non Allergen	Non Toxin
8	433	440	FKNKTDII	8	Antigen	Non Allergen	Non Toxin
9	465	472	ELLTKSNN	8	Antigen	Allergen	Non Toxin
10	492	500	ITANVSRGK	9	Non Antigen	Allergen	Non Toxin
11	512	521	MDINHPIYYV	10	Non Antigen	Allergen	Non Toxin
12	535	542	VSYDQGVS	8	Non Antigen	Allergen	Non Toxin

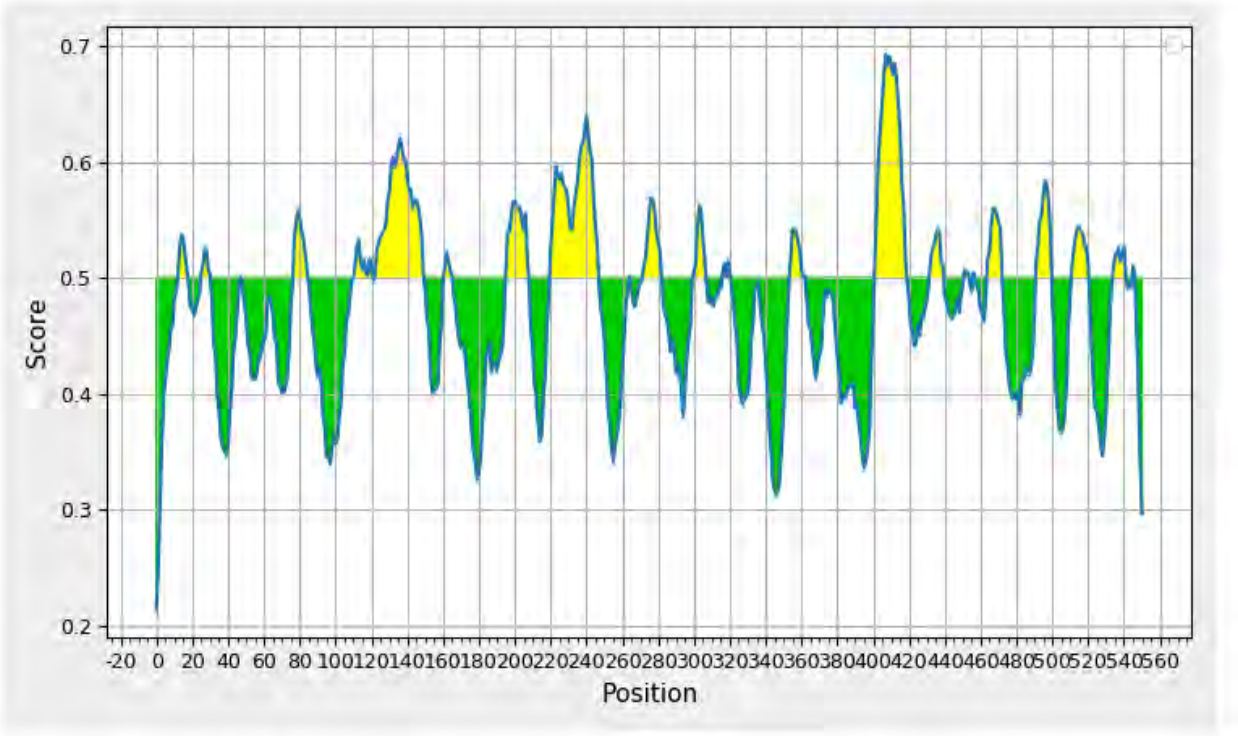

***Figure 5**: Graphical Overview of B cell Epitopes* (Hamza, 2021)

In figure 5, 0.5 is the threshold point and lines above 0.5 are pointed by yellow color and those below 0.5 are highlighted as green color. On the contrary, yellow-colored lines indicate epitopes which can be referred to as B cell epitopes. On the other hand, those epitopes which are not recognized as B cell epitopes are indicated through green-colored regions (Hamza, 2021).

3.5 Development of Vaccine Sequence:

Vaccine is constructed by assembling the constituents of the vaccine with the help of linkers. Adjuvant which is the protein sequence, 3 CTL epitopes, 3 HTL epitopes and 2 B cell epitopes along with appropriate linkers the main vaccine sequence is created. The whole sequence is given below-

MNNTIINSLIGGDDFIKRSNVFAVDSQIPTLYMPQYISLSGVMTNDGPDNQAIASFEIRD
QYITALNHLVLSLELPEVKGMGRFGYVPYVGYKCINHVSVSSCNGVIWEIEGEELYNNCI
NNTIALKHSGYSSELNDISIGLTPNDTIKEPSTVYVYIKTPFDVEDTFSSLKLSDSKITV
TVTFNPVSDIVIRDSSFDFETFNKEFVYVPELSFIGYMVKNVQIKPSFIEKPRRVIGQIN
QPTATVTEVHAATSLSVYTKPYYGNTDNKFISYPGYSQDEKDYIDAYVSRLLDDLVIVSD
GPPTGYPESAEIVEVPEDGVVSIQDADVYVKIDNVPDNMSVYLHTNLLMFGTRKNSFIYN
ISKKFSAITGTYSDATKRTVFAHISHTINIIDTSIPVSLWTSQRNVYNGDNRSAESKAKD
LFINDPFIKGIDFKNKTDIISRLEVRFGNDVLYSENGPISRIYNELLTKSNNGTRTLTFN
FTPKIFFRPTTITANVSRGKDKLSVRVVYSTMDINHPIYYVQKQLVVVCNDLYKVSYDQG
VSITKIMGDNNEAAAKNTDNKFISYAAYSLSVYTKPYAAYTIALKHSGYGPGPGKRTVFA
HISHTINIIGPGPGNHPIYYVQKQLVVVCGPGPGTKRTVFAHISHTINIKKSQRNVYNGD
NRSAESKAKKKFKNKTDII

3.6 Evaluation of Antigenic, Allergic and Toxic Entities in Vaccine Sequence:

The vaccine sequence has shown probable antigen with an antigenic score of 0.5312 along with a threshold of 0.5 after checking it through Vaxijen v2.0.

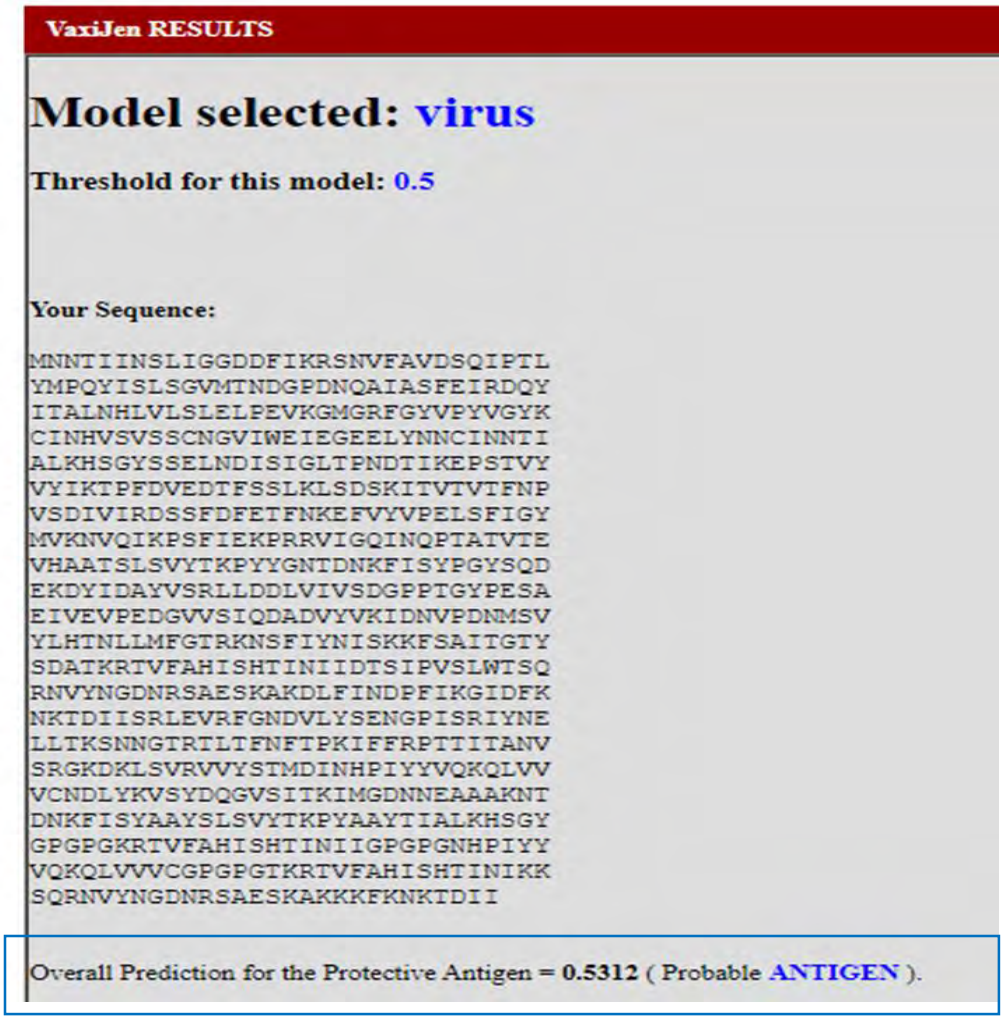

***Figure 6**: Antigenicity of constructed vaccine* (Doytchinova & Flower, 2007)

Moreover, in the "AllergenOnline" database the result shows the number of sequences with hits is 0 which indicates there is no allergic element in the vaccine sequence.

Database	AllergenOnline Database v22 (May 25, 2023)
Input Query	>FASTA MNNTIINSLIGGDDFIKRSNVFAVDSQIPTLYMPQYISLSGVMTNDGPDNQAIASFEIRD QYITALNHLVLSLELPEVKGMGRFGYVPYVGYKCINHVSVSSCNGVIWEIEGEELYNNCI NNTIALKHSGYSSELNDISIGLTPNDTIKEPSTVYVYIKTPFDVEDTFSSLKLSDSKITV TVTFNPVSDIVIRDSSFDFETFNKEFVYVPELSFIGYMVKNVQIKPSFIEKPRRVIGQIN QPTATVTEVHAATSLSVYTKPYYGNTDNKFISYPGYSQDEKDYIDAYVSRLLDDLVIVSD GPPTGYPESAEIVEVPEDGVVSIQDADVYVKIDNVPDNMSVYLHTNLLMFGTRKNSFIYN ISKKFSAITGTYSDATKRTVFAHISHTINIIDTSIPVSLWTSQRNVYNGDNRSAESKAKD LFINDPFIKGIDFKNKTDIISRLEVRFGNDVLYSENGPISRIYNELLTKSNNGTRTLTFN FTPKIFFRPTTITANVSRGKDKLSVRVVYSTMDINHPIYYVQKQLVVVCNDLYKVSYDQG VSITKIMGDNNEAAAKNTDNKFISYAAYSLSVYTKPYAAYTIALKHSGYGPGPGKRTVFA HISHTINIIGPGPGNHPIYYVQKQLVVVCGPGPGTKRTVFAHISHTINIKKSQRNVYNGD NRSAESKAKKKFKNKTDII
Length	679
Number of 80 mers	600
Number of Sequences with hits	0

***Figure 7**: Allergenicity result of the vaccine* (Sircar et al., 2014)

Furthermore, there is no toxic element found in the vaccine sequence. After checking it through the T3DB server the server shows the result “your search returned no result” which refers to no toxic element in the sequence.

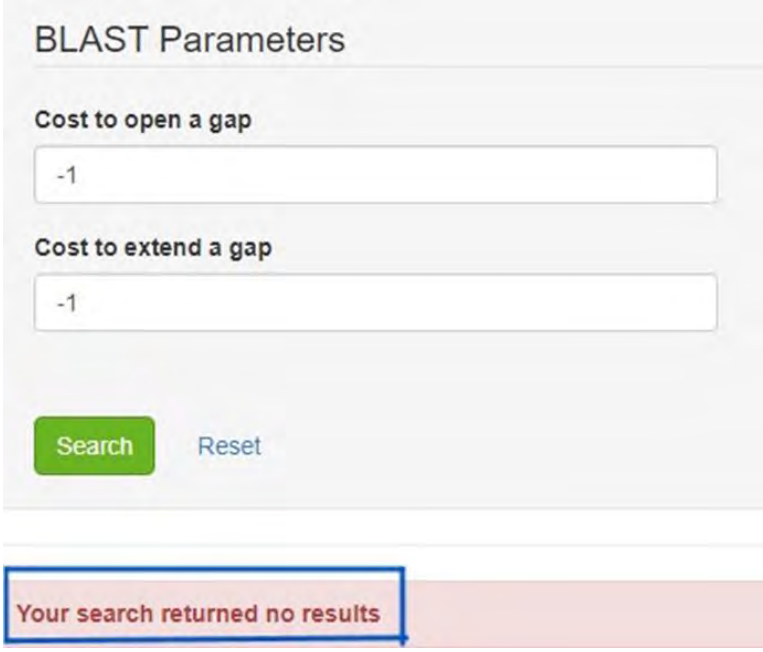

***Figure 8**: Toxicity result for vaccine model* (Wishart et al., 2014)

3.7 Biochemical Feature Analysis:

Biochemical characteristics of any vaccine include several properties which are checked through computational tools.

To identify the 3D model of the vaccine Phyre2 is used which shows 100% confidence with a coverage of 80%.

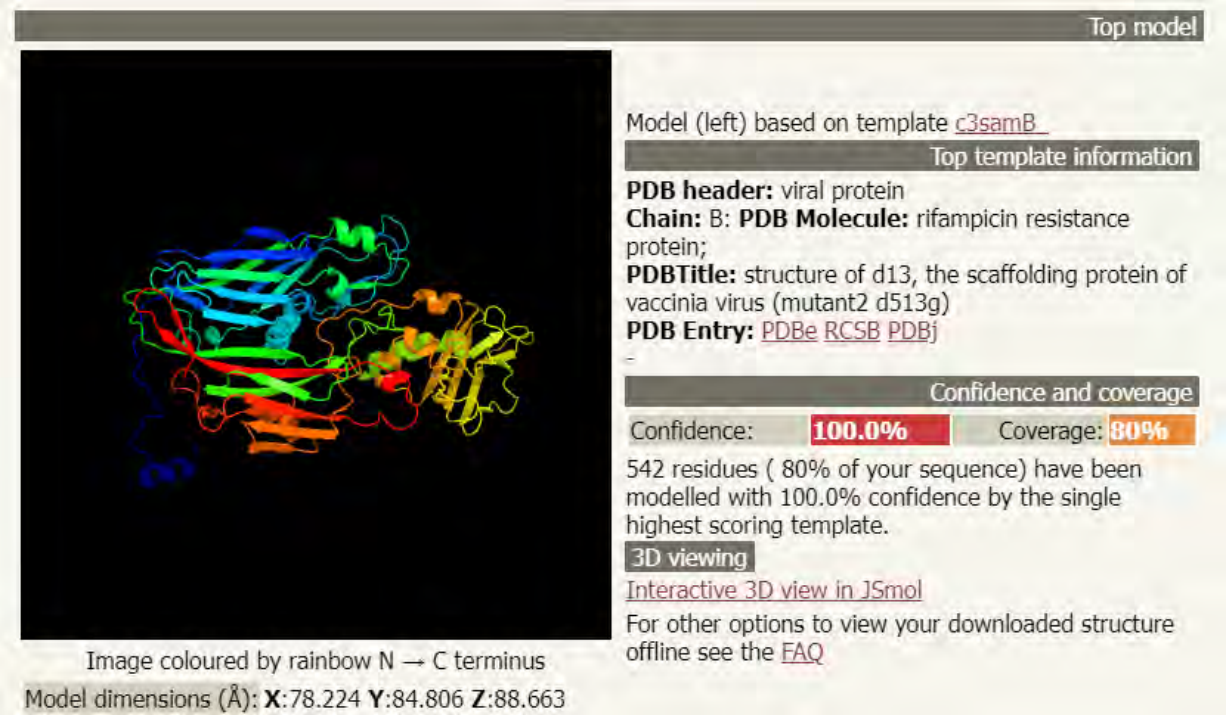

***Figure 9**: Vaccine's score in Phyre 2* (Kelley et al., 2015)

Physicochemical properties of any protein sequence is evaluated with the help of ProtParam tool. In protparam it has been identified that the vaccine sequence is 75924 dalton in weight with an amino acid number of 679. Moreover, the value of theoretical pI is 7.61 with a GRAVY of -0.255. Furthermore, the instability index and aliphatic score is respectively 29.61 and 86.91.

Number of amino acids: 679

Molecular weight: 75924.00

Theoretical pI: 7.61

Amino acid composition: CSV format

Ala (A)	29	4.3%
Arg (R)	22	3.2%
Asn (N)	52	7.7%
Asp (D)	42	6.2%
Cys (C)	5	0.7%
Gln (Q)	16	2.4%
Glu (E)	26	3.8%
Gly (G)	40	5.9%
His (H)	14	2.1%
Ile (I)	65	9.6%
Leu (L)	35	5.2%
Lys (K)	47	6.9%
Met (M)	9	1.3%
Phe (F)	30	4.4%
Pro (P)	34	5.0%
Ser (S)	61	9.0%
Thr (T)	50	7.4%
Trp (W)	2	0.3%
Tyr (Y)	41	6.0%
Val (V)	59	8.7%
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): 68
Total number of positively charged residues (Arg + Lys): 69

Atomic composition:

Carbon	C	3420
Hydrogen	H	5314
Nitrogen	N	890
Oxygen	O	1036
Sulfur	S	14

Formula: $C_{3420}H_{5314}N_{890}O_{1036}S_{14}$
Total number of atoms: 10674

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 29.61
This classifies the protein as stable.

Aliphatic index: 86.91

Grand average of hydropathicity (GRAVY): -0.255

***Figure 10**: Chemical & Physical feature's result.*

3.8 Determination of Z-score and Ramachandran Plot:

The ProSA-web tool shows the result of z-score -8.12 which is within the range according to some papers. In the graph of residues there are two regions which are X-ray crystallography distinguished by light blue color and Nuclear Magnetic Resonance region identified by dark blue color. The black dot in the graph must have to lie in the blue-colored region which indicates a valid result (Kwofie et al., 2018). In this study the black dot is plotted in the light blue region which is X-ray crystallography.

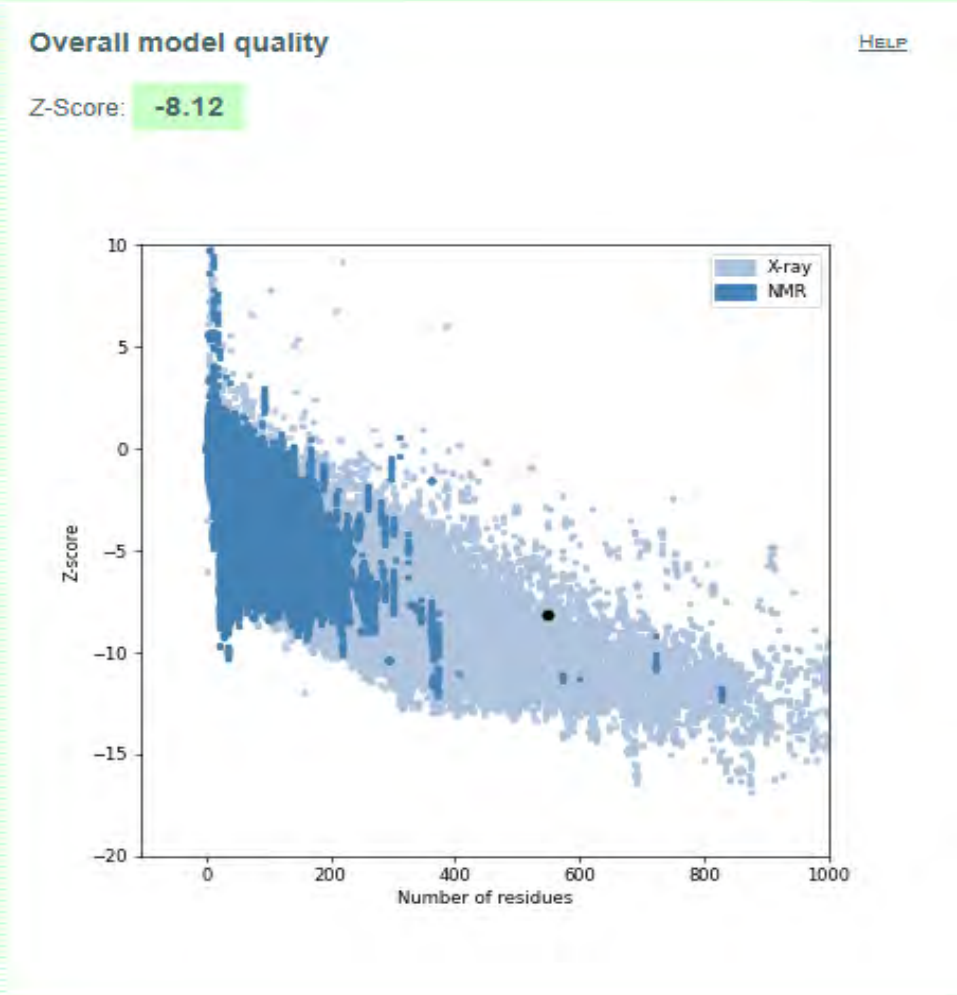

Figure 11: *Graphical view of Z-score plot* (Kwofie et al., 2018)

Another graph is given here which indicates residue scores stating that no steric hindrance is present between the residues specifying the model as a stable one (Roy et al., 2011).

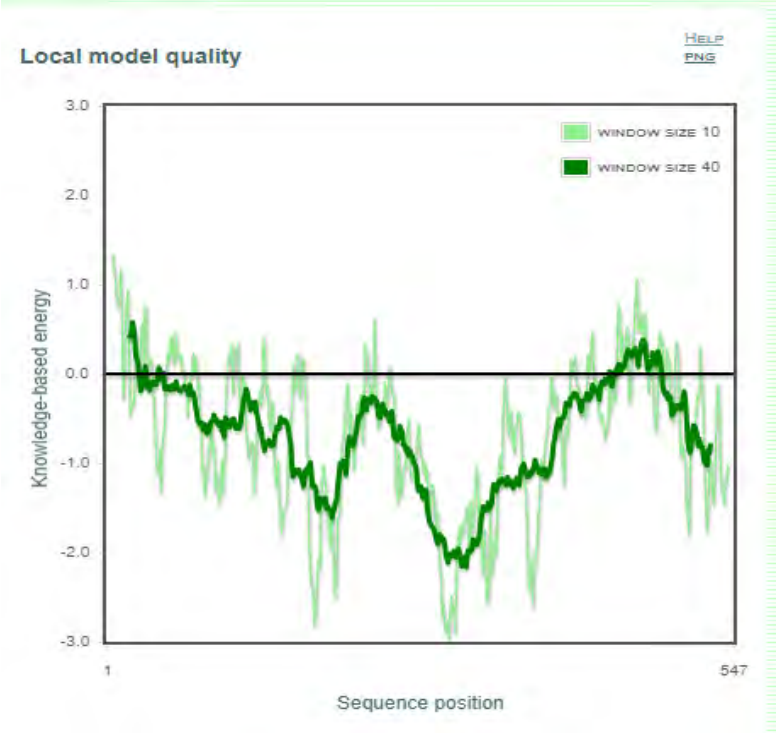

***Figure 12**: Residue Score plotting graph* (Roy et al., 2011)

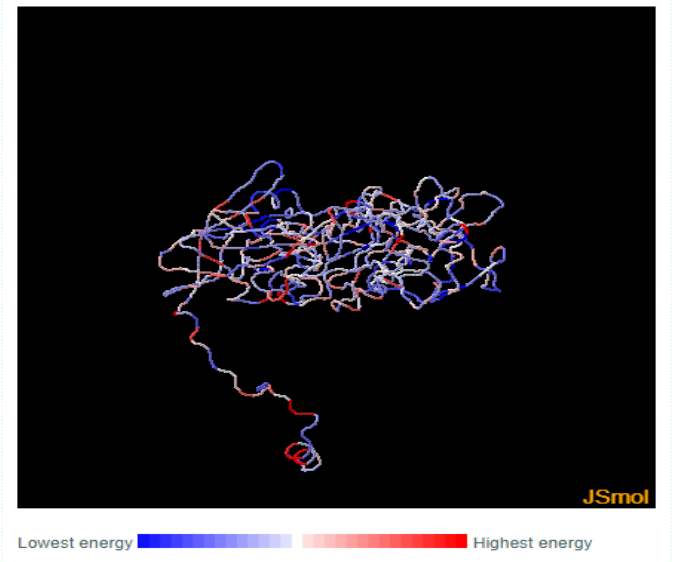

***Figure 13**: Energy Model of the residues* (Roy et al., 2011)

On the contrary, identification of Ramachandran Plot is done through the Swiss Model expasy tool where a regional figure is presented as the result included in figure 14. In this figure the structural view of the vaccine is pointed and also how molecules are scattered throughout the region is displayed. Moreover, molprobity result is given based on the 4 regions of the Ramachandran plot shown in figure 15. Ramachandran's outlier of this model is 95.96% whereas, ramachandran favored region is only 0.18% which is a very good score according to the papers. In addition, the percentage of rotamer outliers is 0.00 and C- Beta deviation is 0. Bad angle and bad bond score is respectively, 28/4435 and 18/6036.

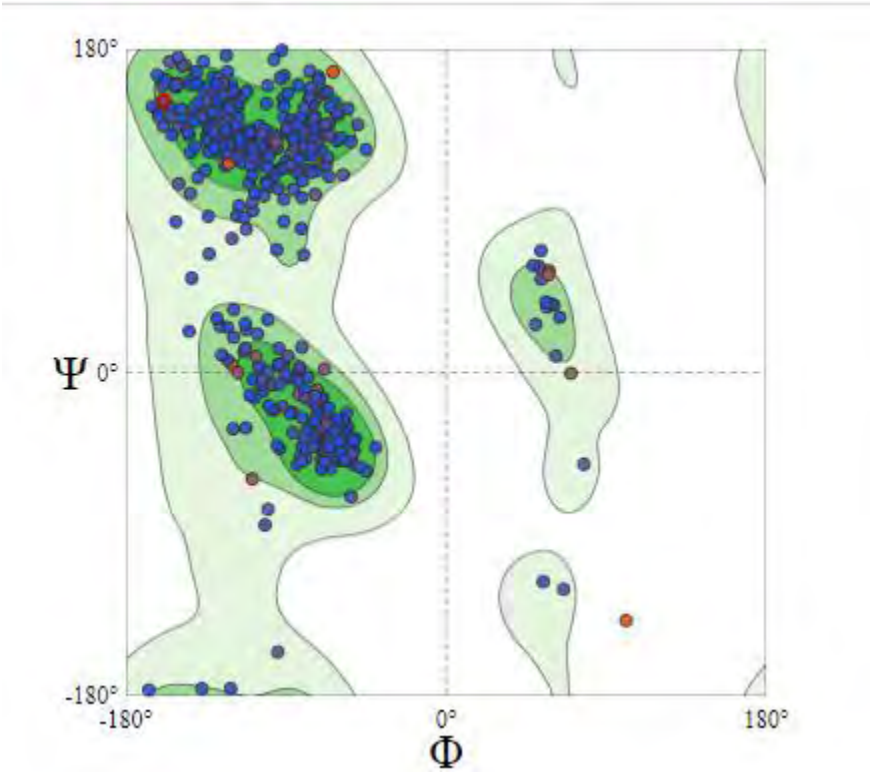

Figure 14*: Ramachandran region visual presentation* (Schwede, 2003)

MolProbity Score	2.16
Clash Score	24.68
(A440 ILE-A475 ARG), (A233 ARG-A246 VAL), (A234 ARG-A258 TYR), (A162 PHE-A171 LEU), (A113 GLU-A234 ARG), (A89 TYR-A126 LEU), (A359 TYR-A392 ASP), (A271 ILE-A412 ARG), (A131 TYR-A278 GLN), (A258 TYR-A526 VAL), (A84 PHE-A191 VAL), (A113 GLU-A260 LYS), (A67 ASN-A68 HIS), (A344 HIS-A347 LEU), (A255 LEU-A529 CYS), (A344 HIS-A346 ASN), (A132 SER-A276 TYR), (A512 MET-A513 ASP), (A131 TYR-A276 TYR), (A442 ARG-A456 ASN), (A297 ILE-A306 TYR), (A287 TYR-A291 LEU), (A230 GLU-A253 THR), (A47 GLY-A48 PRO), (A512 MET-A516 HIS), (A194 ASP-A194 ASP), (A113 GLU-A258 TYR), (A40 SER-A211 GLU)	
Ramachandran Favoured	95.96%
Ramachandran Outliers	0.18%
A513 ASP	
Rotamer Outliers	0.00%
C-Beta Deviations	0
Bad Bonds	28 / 4435
A386 HIS, A250 HIS, A68 HIS, A97 HIS, A344 HIS, A516 HIS, A383 HIS, A128 HIS, A242 PRO, A458 PRO, A48 PRO, A302 PRO, A29 PRO, A51 GLN-A52 ALA, A210 PRO, A76 PRO, A34 PRO, A274 PRO, A426 PRO, A18 ARG-A19 SER	
Bad Angles	18 / 6036
(A11 GLY-A12 GLY), (A128 HIS-A129 SER), (A129 SER-A130 GLY), (A51 GLN-A52 ALA), (A408 ASN-A409 GLY), (A317 GLU-A318 ASP), (A514 ILE-A515 ASN), A344 HIS, A97 HIS, A68 HIS, A250 HIS, A386 HIS, A516 HIS, A383 HIS, A128 HIS, (A89 TYR-A90 VAL)	
Cis Non-Proline	2 / 520
(A129 SER-A130 GLY), (A408 ASN-A409 GLY)	

***Figure 15**: MolProbity scores* (Schwede, 2003)

3.9 Molecular Docking:

In terms of molecular docking, “ClusPro 2.0: Protein-Protein Docking” server has been used in this study. Molecular docking refers to the binding of vaccine ligands with the receptor of the human body. For scaffold protein the best binding receptor is Toll like receptor 4 and the pdb format of the TLR4 receptor model is chosen from the rcsb protein data bank. In the ClusPro server after completion of docking 29 docked results have been shown. Moreover, cluster scores of each model are given with their respective energies. Among them, the model which has the lowest energy is chosen as the docked model of MPXV vaccine. Consequently, model 3 has the lowest

energy which is -1138.8 and cluster member no is 35. The docked figure of this complex is shown below-

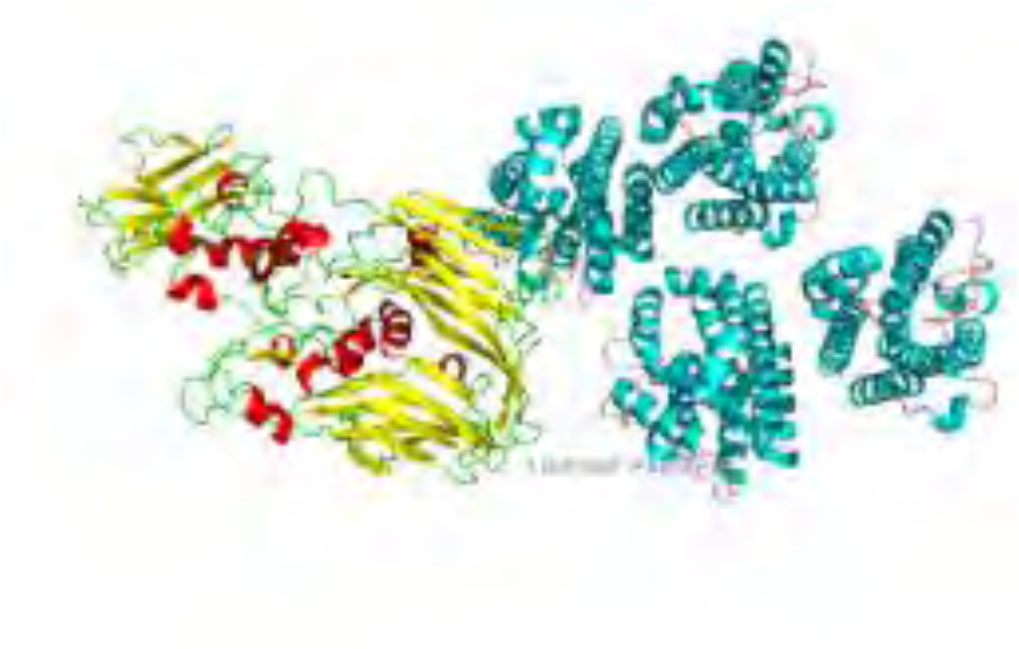

Figure 16*: Docked model of the vaxine* (Alekseenko et al., 2020)

Moreover, the table containing all the model's score is given below-

***Table 5**: Cluster Score from ClusPro.*

Cluster	Members	Representative	Weighted Score
0	39	Center	-782.2
		Lowest Energy	-911.3
1	38	Center	-826.8
		Lowest Energy	-888.8
2	37	Center	-980.6
		Lowest Energy	-1099.4
3	35	Center	-1138.8
		Lowest Energy	-1138.8
4	30	Center	-965.1
		Lowest Energy	-994.2
5	28	Center	-888.2
		Lowest Energy	-888.2
6	26	Center	-833.7
		Lowest Energy	-1039.9
7	26	Center	-930.6
		Lowest Energy	-952.9
8	22	Center	-800.2
		Lowest Energy	-907.7
9	20	Center	-794.5
		Lowest Energy	-918.6
10	19	Center	-770.8
		Lowest Energy	-934.8

3.10 Immune Simulation:

Immune simulation is done by C-ImmSim web browser to check the antibody producing ability of the vaccine and immune response. In this server immunity is expressed through various graphical figures. The main figure expresses the antibody producing ability after injecting 3 doses of the vaccine. The black line of the graph indicates the antigen, other coloured lines are antibodies which are pointed in the figure. On day 0, there is no antibody and the level of antigen is highest. Gradually, the level of antibody after injecting each dose has been increased, conversely the level

of antigen has been decreased in the second dose. As the level of graphical line for antibodies has increased it refers to increasing the level of antibody producing ability in our body.

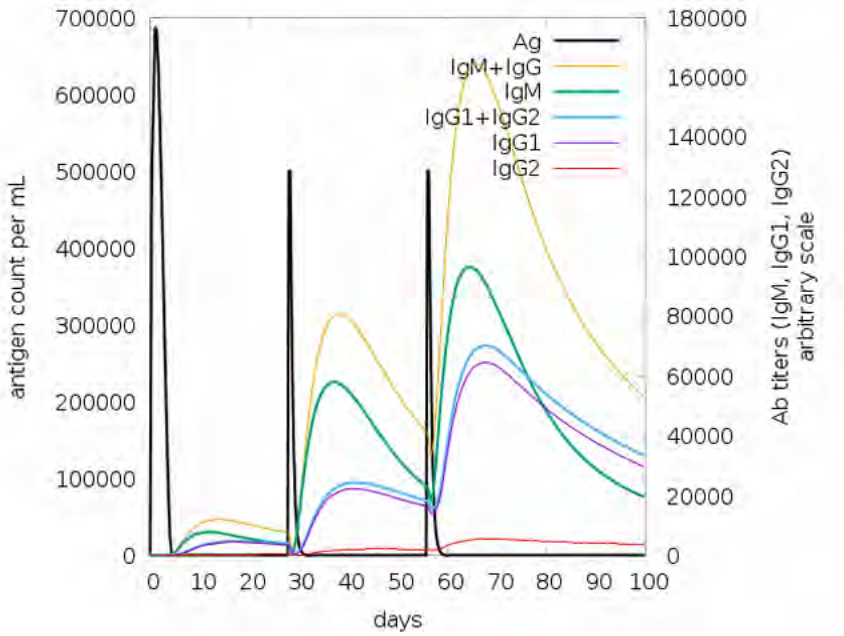

Figure 17: *Antigen count per day graph* (Motamedi et al., 2023)

Moreover, C-ImmSimm also provides a graphical view for B lymphocytes which include total count, memory cell and also antibodies.

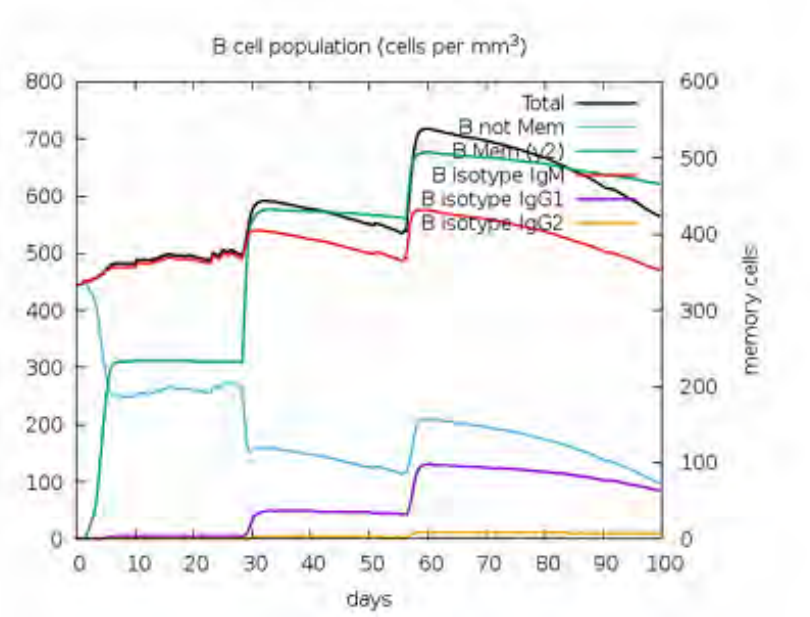

***Figure 18 (a)**: B Cell Population Graph* (Motamedi et al., 2023)

Another graphical presentation has been shown in figure 18 (b) expressing B cell population per entity state where active B cells are indicated by purple line.

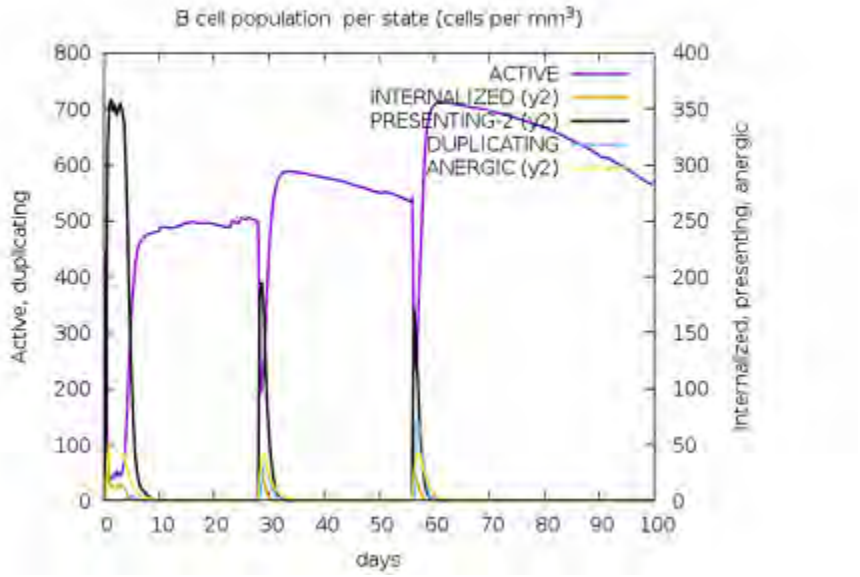

***Figure 18 (b)**: B cell population per state graph* (Motamedi et al., 2023)

According to figure Plasma B Lymphocyte count is observed. It is shown as per isotypes.

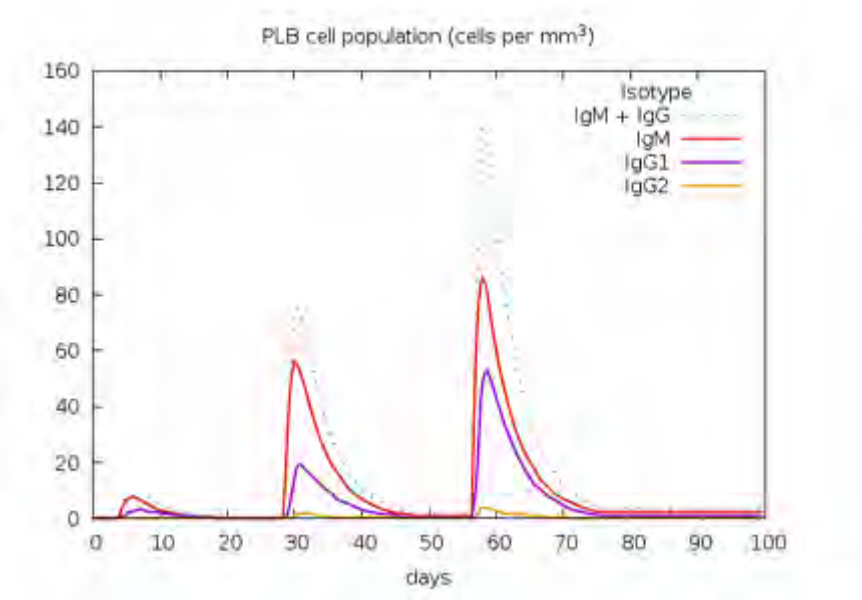

***Figure 19**: Diagram for Plasma B Lymphocyte cell population* (Motamedi et al., 2023)

Moreover, there are other graphs for helper T cells. Total 3 graphs are presented for explaining Helper T cells. One of them is for CD4 T Helper cell population, second one is for the Helper T cell population per state and the last one is for CD4 T regulatory cell population per state.

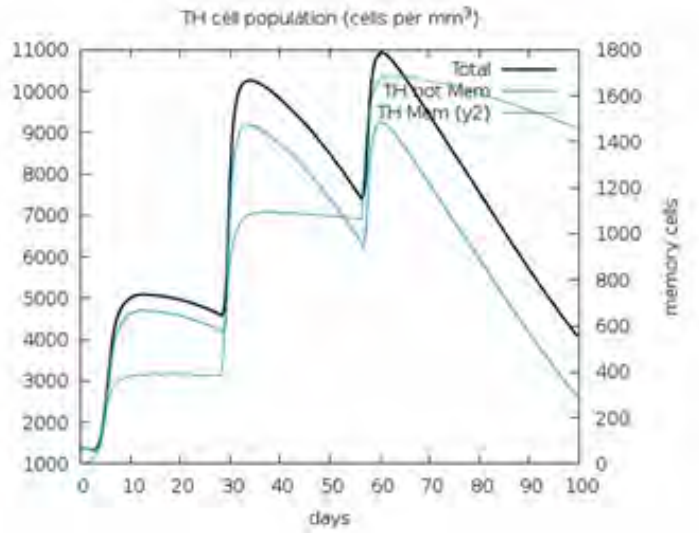

***Figure 20**: Graph of Helper T cell Population* (Motamedi et al., 2023)

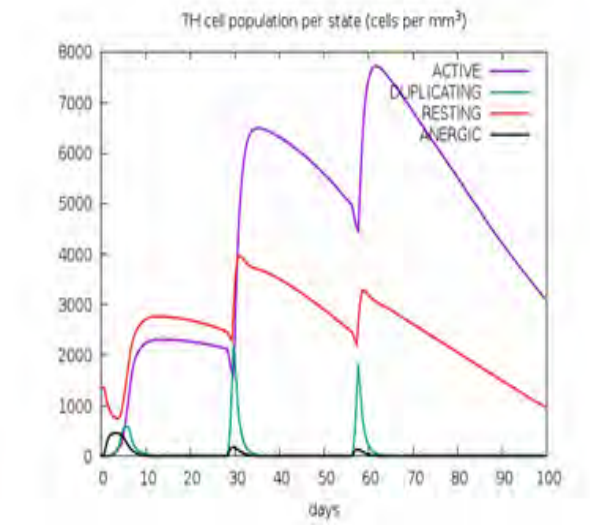

Figure 21 (a): *Schematic presentation of Helper T cell*

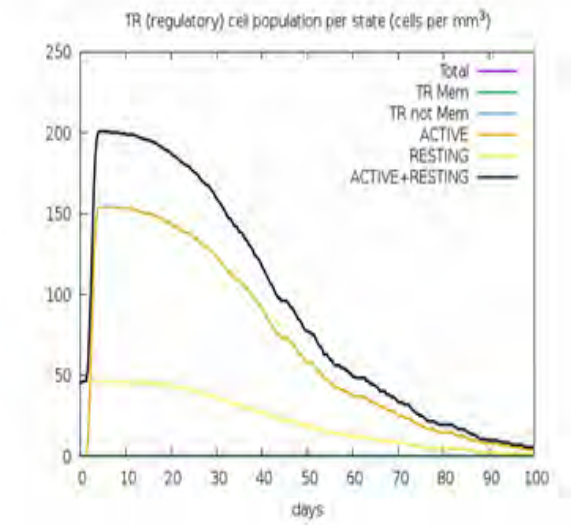

Figure 21 (b): *Regulatory T cell population per state.*

Grapes are also presented for explaining CTL (Cytotoxic T Lymphocyte) cells. Graph presenting the population is a scattered one and the other one is for population per state which is explained through lines.

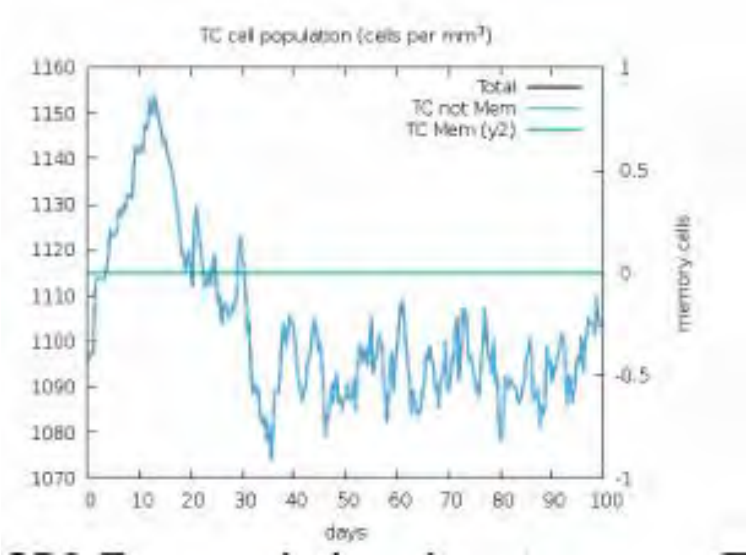

Figure 22 (a)*: CD8-CTL Population Graph* (Motamedi et al., 2023)

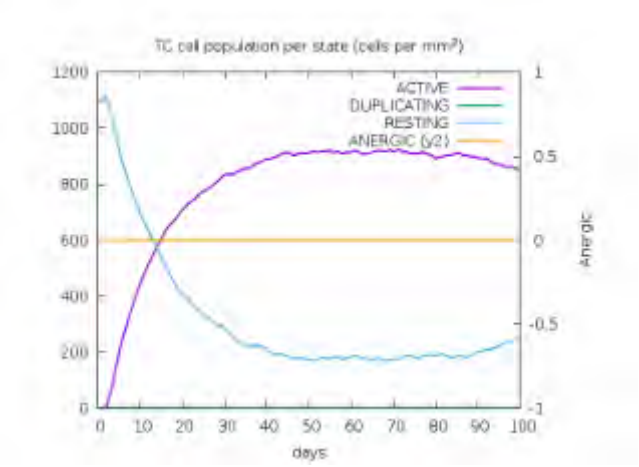

Figure 22 (b): *CD8-CTL count per state graph* (Motamedi et al., 2023)

Furthermore, in our immune system there are some other cells which are very crucial for expressing the immunity which are natural killer cells, macrophages, dendritic cells and epithelial cells. For each of these cells a graphical view is given for proper understanding of these cells in our body. Figure 23 on dendritic cells, only the total count of dendritic cells is included. In terms of dendritic cells it presents the graphical view for MHC class I & II molecules. In addition, curves in macrophages are for MHC class II molecules which covers the total count, active as well the resting molecules. Furthermore, in figure 26 which is highlighting epithelial cells are shows the active molecules only.

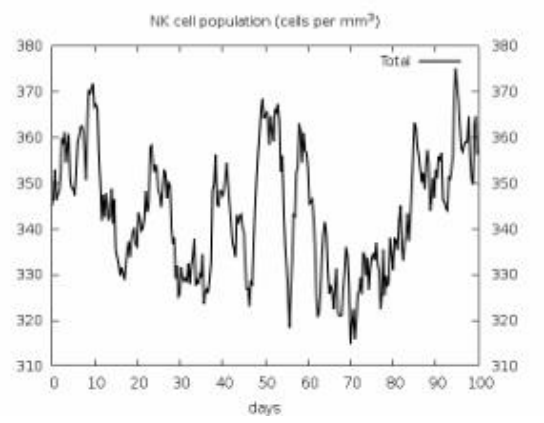

Figure 23*: Graph of Natural Killer Cells* (Motamedi et al., 2023)

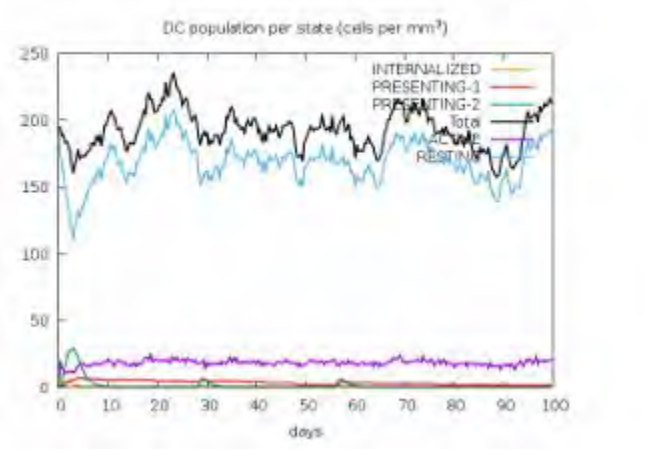

***Figure 24**: Graphical Overview of Dendritic Cells (Motamedi et al., 2023)*

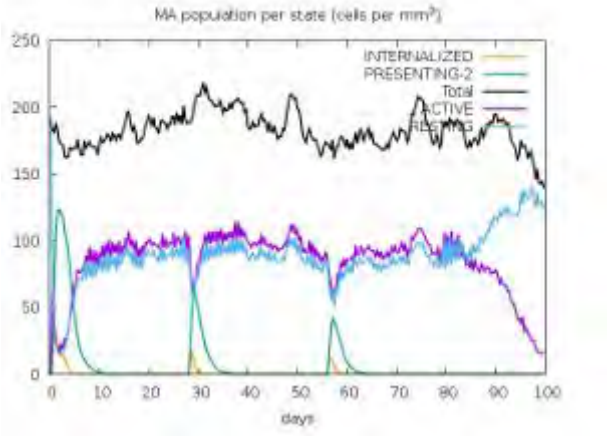

***Figure 25**: Macrophages population per state (Motamedi et al., 2023)*

***Figure 26**: Total Count of Epithelial cells* (Motamedi et al., 2023)

At last, cytokines and interleukins concentration states have been also shown through a figure which is highlighted in figure 27. Here, D refers to the danger sign and IL refers to interleukin.

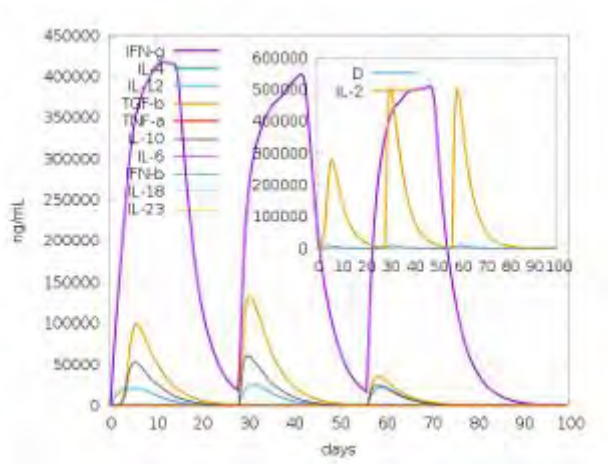

Figure 27: *Concentration between interleukin and cytokine* (Motamedi et al., 2023)

Chapter 4

Discussion:

In silico multi epitope vaccine design means arrangement of the vaccine model with the help of multiple epitopes by using a significant number of computational tools.

As the selection of primary protein, scaffold protein OPG 125 has been chosen which has an amino acid length of 551. It is proven that scaffold protein is a very important protein for the pox viruses as it provides the basic platform of these viruses. Consequently, by targeting this protein it will be very easy to destroy the basic structure of the virus and their life cycle as well. On top of that, this protein was particularly chosen based on its high antigenic score which is shown at figure 3. As a result, selection of this protein will be very beneficial in terms of vaccine construction.

After that, in our immune system 3 types of cells are very important and play a very crucial role which are- Cytotoxic T Lymphocyte, Helper T Lymphocyte and B Cells. CTL helps by killing all the infected cells, HTL shows their action by providing stimulation to the immune cells as well they store information about all types of infection so that it can be applied in future. Moreover, they have the capability of inducing cytokines which also has an impact on the immune system. Finally, B cells prove their activity by producing a vast amount of antibody against the antigen. Epitopes of all these cells are connected together and assembled the vaccine's protein sequence. However, these epitopes can't have any non-antigenic, allergic and toxic properties which will further make the vaccine sequence inefficient. A total number of 3 CTL, 3 HTL and 2 B cell epitopes are used in this in silico vaccine model.

On the contrary, it's mostly important to check all the physical and chemical properties of the constructed vaccine sequence. All these criteria are assessed and highlighted in figure 10. Additionally, structural assessment is another major feature for this vaccine and it shows an

excellent outcome which is pointed out in figure 14 & 15. Furthermore, the quality of the 3D vaccine model has been also assessed by determining the Z score value.

Last but not the least, the main two features that must be evaluated in a vaccine are immune simulation and molecular dynamic simulation. Molecular docking is conducted to check the binding ability of the vaccine with the body receptor. Result of this binding ability is presented in table 5. Finally, capability of antibody production of this vaccine against the antigen is examined through the C-ImmSimm tool and the results are explained by using graphs (figure 16-28) provided by this server.

To sum up, the outbreak of monkeypox virus has become a serious issue throughout the world which is why a vaccine is badly needed to treat monkeypox disease. In addition, based on the proposed vaccine model it can be predicted that further necessary steps can be taken to prepare an effective vaccine against monkeypox virus.

Chapter 5

Conclusion:

The main purpose of this study was to come up with an in silico vaccine model against monkeypox virus targeting specifically the protein of monkeypox virus. According to research, there are two licensed vaccines against smallpox which are additionally used to treat monkeypox disease. This model is established based on its nontoxic, non-allergic and antigenic features. Besides, the 3D model of this vaccine has proven excellent coverage with full confidence. Additionally, all the physical and chemical tests are done to check the vaccine's functioning ability. Moreover, structural assessment of this vaccine model has also been determined. On top of that, whether this vaccine has a powerful binding ability with the human body receptor or not is also tested by doing molecular docking. Finally, immune simulation has been conducted to check the antibody producing ability. By following this sequential process, the vaccine model of mpxv has been constructed; however, more research is needed to make the vaccine available in the market. In addition, in vivo and in vitro tests are a must for making the best use of this vaccine. Thus, it will be beneficial for people and rate of mpxv cases will be treated properly as a result life risk of individuals will be reduced.

References:

Alakunle, E., Kolawole, D., Diaz-Cánova, D., Alele, F., Adegboye, O. A., Moens, U., & Okeke, M. I. (2024). A comprehensive review of monkeypox virus and mpox characteristics. *Frontiers in Cellular and Infection Microbiology*, *14*.

Alekseenko, A., Ignatov, M., Jones, G., Sabitova, M., & Kozakov, D. (2020). Protein–Protein and Protein–Peptide Docking with ClusPro Server. In *Methods in molecular biology*

Anwar, F., Haider, F., Khan, S., Ahmad, I., Ahmed, N., Imran, M., Rashid, S., Ren, Z., Khattak, S., & Ji, X. (2023). Clinical Manifestation, transmission, pathogenesis, and diagnosis of Monkeypox Virus: A Comprehensive review. *Life*, *13*.

Ashley, C., Broni, E., Wood, C. D., Okuneye, T., Ojukwu, M. T., Dong, Q., Gallagher, C. J., & Miller, W. A. (2024). Identifying potential monkeypox virus inhibitors: an in silico study targeting the A42R protein. *Frontiers in Cellular and Infection Microbiology*, *14*.

Hamza, A. (2021). Immunoinformatics Prediction of Epitope Based Peptide Vaccine against Clostridium perfringens Fructose Bisphosphate Aldolase Protein. *www.academia.edu*. https://www.academia.edu/68421244/Immunoinformatics_Prediction_of_Epitope_Based_Peptide_Vaccine_against_Clostridium_perfringens_Fructose_Bisphosphate_Aldolase_Protein

Harapan, H., Ophinni, Y., Megawati, D., Frediansyah, A., Mamada, S. S., Salampe, M., Emran, T. B., Winardi, W., Fathima, R., Sirinam, S., Sittikul, P., Stoian, A. M. M., Nainu, F., & Sallam, M. (2022). Monkeypox: A comprehensive review. *Viruses*, *14*.

Ho, B., & Brasseur, R. (2005). The Ramachandran plots of glycine and pre-proline. *BMC Structural Biology*, *5*.

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., & Sternberg, M. J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols*, *10.*

Kwofie, S. K., Dankwa, B., Odame, E., Agamah, F. E., Doe, L. P. A., Teye, J., Agyapong, O., Miller, W. A., Mosi, L., & Wilson, M. D. (2018). In Silico Screening of Isocitrate Lyase for Novel Anti-Buruli Ulcer Natural Products Originating from Africa. *Molecules/Molecules Online/Molecules Annual*, *23*.

Luna, N., Muñoz, M., Bonilla-Aldana, D. K., Patiño, L. H., Kasminskaya, Y., Paniz-Mondolfi, A., & Ramírez, J. D. (2023). Monkeypox virus (MPXV) genomics: A mutational and phylogenomic analyses of B.1 lineages. *Travel Medicine and Infectious Disease*, *52*.

McCollum, A. M., & Damon, I. K. (2013). Human monkeypox. *Clinical Infectious Diseases/Clinical Infectious Diseases (Online. University of Chicago. Press)*, *58*.

Mitjà, O., Ogoina, D., Titanji, B. K., Galvan, C., Muyembe, J., Marks, M., & Orkin, C. (2023). Monkeypox. *Lancet*, *401*.

Rabaan, A. A., Alasiri, N. A., Aljeldah, M., Alshukairiis, A. N., AlMusa, Z., Alfouzan, W., Abuzaid, A. A., Alamri, A. A., Al-Afghani, H. M., Al-Baghli, N., Al-Qahtani, N. M., Al-Baghli, N. A., Almoutawa, M. Y., Alawi, M., Alabdullah, M. A., Bati, N. a. A., Alsaleh, A. A., Tombuloğlu, H., Arteaga-Livias, K., . . . Imran, M. (2023b). An updated review on Monkeypox viral disease: Emphasis on Genomic diversity. *Biomedicines*, *11*.

Reynisson, B., Alvarez, B., Paul, S., Peters, B., & Nielsen, M. (2020). NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. *Nucleic Acids Research*, *48*.

Roy, S., Maheshwari, N., Chauhan, R., Sen, N., & Sharma, A. (2011). Structure prediction and functional characterization of secondary metabolite proteins of Ocimum. *Bioinformation*, *6*.

Schwede, T. (2003). SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Research*, *31*.

See, K. C. (2022). Vaccination for monkeypox virus infection in humans: A review of Key considerations. *Vaccines*, *10*.

Thornhill, J., Barkati, S., Walmsley, S., Rockstroh, J., Antinori, A., Harrison, L. B., Palich, R., Nori, A., Reeves, I., Habibi, M. S., Apea, V., Boesecke, C., Vandekerckhove, L., Yakubovsky, M., Sendagorta, E., Blanco, J. L. S., Florence, É., Moschese, D., Maltêz, F., . . . Orkin, C. (2022). Monkeypox Virus Infection in Humans across 16 Countries — April–June 2022. *New England Journal of Medicine/˜the œNew England Journal of Medicine*, *387*.

Wiederstein, M., & Sippl, M. J. (2007). ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Research*, *35*.