In-Silico Based Design of Lysosome Membrane Protein 2 Targeted Multi-Epitope Vaccine Against Breast Cancer

By
Khadiza Hasan Roddur
ID: 19346067

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

School of Pharmacy
BRAC University
January 2026

Declaration

It is hereby declared that

1. The thesis submitted is my own original work while completing the degree at BRAC University.
2. This 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 that 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.

Khadiza Hasan Roddur
ID: 19346067

Approval

The thesis titled "In-Silico Based Design of Lysosome membrane protein 2 targeted multi-epitope Vaccine against Breast Cancer" submitted by Khadiza Hasan Roddur (19346067) of Spring 2020 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of Bachelor of Pharmacy (Hons.) on January 15, 2026.

Supervised By:

Supervisor:

Mohammad Kawsar Sharif Siam
Senior Lecturer
School of Pharmacy
BRAC University

Approved By:

Chairperson:

Dr. Shamind Neelotpol,PhD
Professor and Chairperson
School of Pharmacy
BRAC University

Acting Dean:

Professor F.M.Yusuf Haider, PhD
Acting Dean, School of Pharmacy
Department of Mathematics and Natural Sciences
BRAC University

Ethics Statement

All ethical standards were maintained in the completion of this thesis. While completing the research, no human or animal trials were involved.

Abstract

Although the clinical therapeutic approaches have improved, breast cancer remains one of the most widespread malignancies on a global scale and remains a major burden on the population's health. In silico approaches to the design of therapeutic cancer vaccines provide an economic and efficient solution to the generation of specific immune responses against cancer-associated antigens. Lysosome-Associated Membrane protein 2 (LAMP2) has been considered in the current study, where it is an antigen with the potential of creating a multi-epitope subunit vaccine targeting breast cancer by using various set of immunoinformatic tools. Amino-acid sequence of LAMP2 was obtained at UniProt database and interrogated to determine B-cell, cytotoxic T-lymphocyte (CTL), and helper T-lymphocyte (HTL) epitopes. Antigenicity, allergenicity, toxicity, and population coverage were rigorously screened on candidate epitopes, which guaranteed not only the vaccine safety but also the wide range of immunogenic efficacy. The chosen epitopes were then modeled and tested in physicochemical characteristics, secondary, tertiary structure, and molecular docking of Toll-like receptors (such as TLR2 and TLR3). Both epitopes were conjugated to the appropriate adjuvants and linker sequences in order to enhance immunogenicity. The computational simulations showed a stable interaction and a strong antigenic response, which means that LAMP2-derived epitopes can induce a strong immune response to breast cancer cells. In general, this in silico investigation demonstrates that LAMP2 can be an excellent immunogenic target and recommends additional refinement and next steps of validation, both in vivo and in vitro**.**

Keywords

Breast Cancer, Lysosome membrane protein, Molecular docking, In-silico drug design.

Dedication

I would like to dedicate this work to my parents, whose love and guidance have given me the wings to fly. Thank you for helping me build my own identity with confidence and strength.

Acknowledgement

I would like to thank Almighty Allah, who helped me achieve everything I have today.

Moreover, I would like to show my utmost gratitude to my supervisor, Senior Lecturer Mohammad Kawsar Sharif Siam, for guiding me in every step of this work. Without his support and motivation, I could not have reached this far with my thesis journey.

Khadiza Hasan Roddur

ID: 19346067

School of Pharmacy

BRAC university

Table of contents

List of Figures

List of Tables

Chapter 1

Introduction

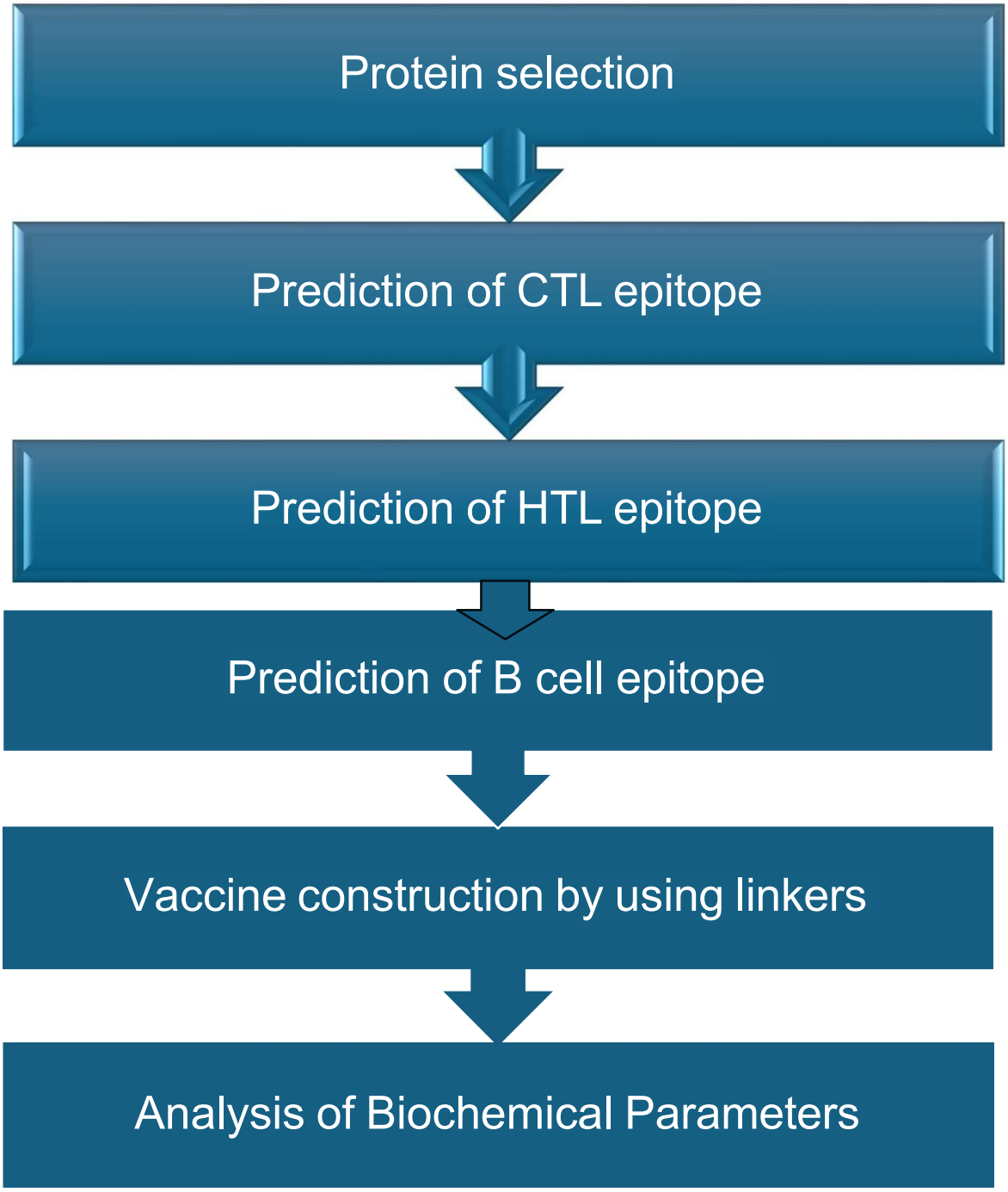
Breast cancer is one of the most common illnesses in the world, which still remains to be a major issue of public health among women. According to data published by the World Health Organization (WHO), more than 2.3 million women were diagnosed with breast cancer in 2020, and approximately 685,000 deaths were observed in the world (Prawiningrum, Paramita, and Panigoro et al., 2022). Although the disease has made tremendous progress in initial diagnosis and combined treatment options such as surgery, chemotherapy, radiotherapy, targeted therapy, and hormone therapy, but recurrence and resistance to therapy remain significant clinical issues (Sung et al., 2021). These shortcomings have prompted growing enthusiasm for the use of cancer immunotherapy, especially vaccine-based immunotherapy, as a method of generating long term and tumor specific immune responses at minimal systemic toxicity. Moreover, the recurrence of the tumor and metastasis may continue to undermine the success rate of the treatment (Zhou et al., 2024).

The function of a breast cancer vaccine is to stimulate the cytotoxic T lymphocytes (CTLs) and helper T cells to kill cancer cells by stimulating the immune system through exposure to an antigen fragment or peptide of neoplasm protein (Gupta et al., 2025). In addition, HER2, MUC1, LAMP1, LAMP2, and AKT1 have been studied in preclinical studies and early-phase clinical trials. For example, peptide vaccines based on the HER2 have demonstrated encouraging immune reactions. Still, because of neoplasm immune evasion and the suppressive tumor microenvironment, the majority of trial vaccination candidates have little clinical efficacy despite good immunogenicity. Of course, computational immunology, or immunoinformatics, has emerged as a revolutionary approach in vaccine design to overcome these challenges. Sometimes, before laboratory validation, researchers can identify possible epitopes, kind of, model three-dimensional protein structures, evaluate molecular interactions, and simulate immune responses by using in-silico approaches (Prawiningrum et al., 2022). The development of a breast cancer vaccine is justified by its potential to prevent tumor recurrence and offer targeted, long-term immune protection. A therapeutic vaccine may have less toxicity, more specificity, and the capacity to create immunological memory as compared to traditional treatments. Furthermore, vaccine design becomes more accurate and effective when combined with computational prediction methods (Prawiningrum et al., 2022)

Thus, it's important to understand that, before clinical translation, more or less in-silico predictions need to be validated through in vitro and in vivo investigations. In conclusion, breast cancer remains one of the significant health issues that requires a new treatment method. The introduction of immunotherapy, particularly vaccine development, has shown a bright way to long-term immunity and the possibility of eliminating recurrence. Researchers may rationally develop and assess multi-epitope vaccines that target important tumor antigens while addressing immunogenicity, allergenicity, and population coverage with the aid of immunoinformatics. Aiming to contribute to the current global endeavor to create safe, effective, and targeted immunotherapeutic options for the management of breast cancer, this thesis focuses on developing and evaluating a computationally designed breast cancer vaccine utilizing in-silico methodologies.

Chapter 2

Methodology



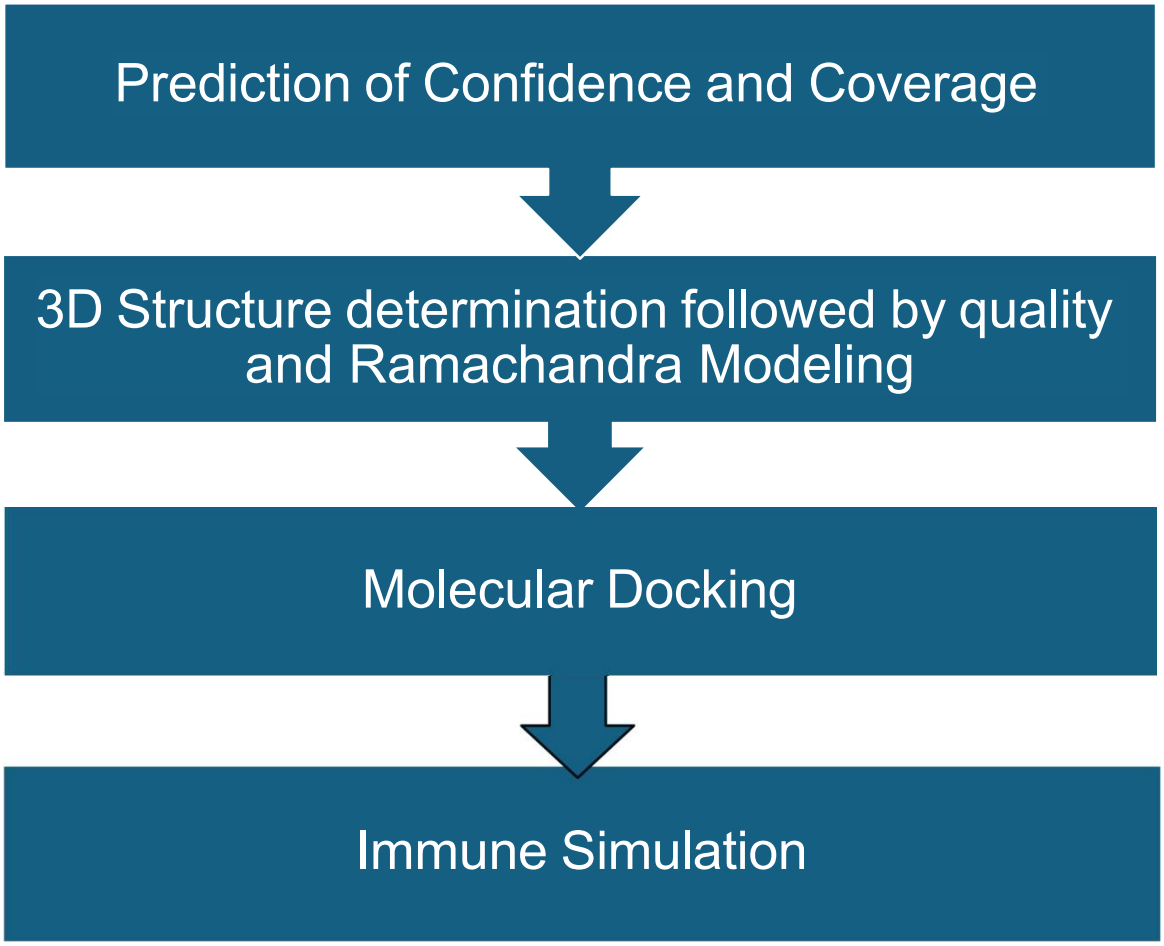


Figure 1: Schematic overview of In Silico Vaccine Design

2.1 Selection of a suitable protein sequence

A suitable protein sequence was selected using the Uniprot protein database. There were 373 proteins in the Uniprot database that were reviewed in the Swiss-Prot section. After reviewing among them, this protein was selected according to the antigenicity analysis. The sequences were downloaded in FASTA format. Antigenicity evaluation was done by the VaxiJen v.2.0 server, and tumors were set as the target organism. Definitely, a threshold value of 0.5 was applied to minimize the risk of false positivity (Doytchinova & Flower, 2007).

2.2 Selection of CTL (CytotoxicT lymphocyte)

The selection of CTL epitopes is a crucial step in the design of peptide-based or subunit vaccines, as these epitopes are responsible for inducing cell-mediated immune responses capable of eliminating tumor or infected cells. Potential CTL determinant is frequently identified by in silico anticipation methods based on their immunogenicity, antigenicity, and binding affinity to human leukocyte antigen (HLA) molecules. To predict the best CTL epitopes, algorithms like NetCTL, NetMHCpan, and the IEDB analysis resource combine factors including proteasomal cleavage, TAP transport efficiency, and MHC class I binding scores to foretell optimal CTL epitopes. All parameters were configured to their normal settings for server-based prevision, but the sorting score was modified to the combined score. Usually, the protein sequences, essentially, were entered in FASTA format. These peptides were submitted to another tool, NETMHCpan 4.1, after obtaining the required alleles from the NetCTL 1.2 tool. Essentially, the server facilitates the prediction of how MHC-I peptides will bind to specific CTL epitopes derived from the previous server. The peptide length in this server should be set to 9- 9-mer peptide. Afterward, all alleles are to be selected, and the BA (Binding Affinity) prediction should be done. Only peptides with a high binding affinity are selected from this website. Further, these peptides are eliminated according to their toxicity, allergenicity, and antigenicity. Every peptide should be nontoxic, nonallergenic, and antigen-free. " Vaxijen v2.0 " has been used to check antigenicity, " AllerTOP v2.0 " for allergenicity, and " ToxinPred " for toxicity.

2.3 Selection of HTL (Helper T Lymphocyte) Epitope

The selection of Helper T Lymphocyte (HTL) epitopes is a critical step in the process of in silico vaccine construction, as these epitopes are essential for initiating and regulating adaptive immune responses. HTL epitopes that are presented by Major Histocompatibility Complex (MHC) class II molecules activate $CD4^+$ T cells. These activated T cells assist in antibody production, cytokine secretion, and cytotoxic T cell activation, resulting in a strong and sustained immune response (Dhanda et al., 2019).

The procedure begins by identifying a suitable antigen that is specific to the disease or tumor. The amino acid sequence of the selected antigen is obtained from reliable protein databases, that is, UniProt in FASTA format.

NetMHCIIpan 4.0 is one of the most advanced and accurate prediction servers. This tool evaluates peptide binding affinity across multiple human leukocyte antigen (HLA) class II, sort of, molecules, including HLA-DR, HLA-DP, and HLA-DQ, using an artificial neural network–based approach. Peptides with low percentile ranks or IC_{50} values (typically < 100 nM) are considered strong binders and are shortlisted for further analysis (Reynisson et al., 2020). After prediction, the selected epitopes are screened for their antigenicity, allergenicity, and toxicity to ensure they are both safe and immunogenic. Antigenicity can be assessed using VaxiJen v2.0, while AllerTOP and ToxinPred are engaged to eliminate allergenic or toxic candidates (Gupta et al., 2021). Additionally, cytokine-inducing potential, especially the ability to produce interferon-gamma (IFN-γ), is evaluated, since this cytokine plays a major role in immune modulation. Alleles showing positive results for IFN-gamma and inducers for IL4 and IL10 cytokines are chosen for an additional screening process. Finally, alleles having antigenic, non-allergenic, and nontoxic properties are screened for vaccine construction.

That's why HTL epitope selection involves antigen identification, in silico prediction using NetMHCIIpan 4.0, filtering for safety and immunogenicity, and validation through structural analyses.

2.4 Selection of B cells

The selection of B cells and their epitopes plays a vital role in the rational design of a vaccine using the in-silico method. B cells are responsible for producing antibodies that recognize specific regions, known as determinants, on the surface of antigens. Identifying these epitopes accurately is essential for developing vaccines that can induce strong humoral immune responses (Parvizpour et al., 2020). The IEDB Resource Analysis server was used to predict the B-cell epitopes, where "Bepipred Linear Prediction 2.0" was selected. Importantly, when the protein sequence was given as input where the output showed a graph along with a table based on antigenicity and the length of each epitope. The epitopes that fall within the range were selected. After that, their antigenicity, allergenicity, and toxicity were checked.

2.5 Construction of vaccine with linkers

In multiepitope vaccine construction, different linkers were utilized to prepare an in silico multi-peptide vaccine. While maintaining their structural integrity and immunogenicity, the linkers play a crucial role in joining selected epitopes into a single constructed sequence. As AAY, GPGPG, KK, and EAAAK are commonly used to ensure flexibility, enhance epitope presentation, antigenicity, and promote efficient antigen processing (Pandey et al., 2018). Notably, these peptide linkers prevent the formation of junctional epitopes and aid proper folding of the vaccine construction. Indeed, the rational use of linkers for in silico vaccine construction enhances stability, solubility, and overall immune response, making them essential for effective vaccine development (Khatoon et al., 2017). Ultimately, this approach improves the efficacy of vaccines.

2.6 Validating the Vaccine's Allergenicity, Antigenicity, and Toxicity

After the construction of a new vaccine, the first and foremost thing is to assay the allergenicity, antigenicity, and toxicity of the constructed vaccine. "Vaxijen v.2.0" is utilized to check the antigenicity of the vaccine, where the threshold was set to 0.5, and the target organism was set as tumor. Again, to check the allergenicity, "Allergen Online" is used. Finally, to assess the perniciousness of the vaccine "T3DB (The Toxin and Toxin Target Database)" server is utilized. And this database will give information about the toxic entity and the nontoxic entities.

2.7 Biochemical Analysis of the Constructed Vaccine

The biochemical analysis of the constructed vaccine involves evaluating its physicochemical stability, solubility, and molecular interactions to ensure antigenic integrity and immunogenic potential (Gasteiger et al., 2005).

The biochemical properties of the construct vaccine are checked by using the server Protparam, and Expasy servers are used to assess molecular weight, theoretical pI, and hydropathicity, stability, and solubility for biological applications. What's more, these servers can provide information about the number of amino acids, molecular weight, instability index, Gravy, and so on.

2.8 3D Structural model of the constructed Vaccine

The 3D structural modeling of the constructed vaccine was done to predict its conformation, stability, and proper folding. Now, the tertiary structure was generated using homology modeling tools such as Phyre2, followed by structural validation through Ramachandran plot analysis. Moreover, this server shows the result based on confidence and insurance coverage in percentage. Confidence refers to the reliance on predicting the right model, and coverage means the ability of the server to predict the right framework. Phyre2 server provides the result as a pdB file through email.

2.9 Evaluation of Ramachandran plots

Evaluation of Ramachandran plots generated by SWISS-MODEL Expasy assesses the stereochemical validity of the modeled vaccine. A high percentage of residues in favored regions confirms reliability and strengthens confidence in the predicted 3D model (Waterhouse et al., 2018). On top of that, the pdB file, which was sent through email, should be uploaded. Moreover, the angles Phi and psi were plotted against each other. It helps to determine the polypeptide chain (Schweder et al., 2003).

2.10 Assessment of Z-Score

With the aim of getting the z-score of the constructed vaccine, the server called ProSA-web was used. Evaluation of the Z score helps to evaluate the overall structural quality of the modeled vaccine through comparison of its energy deviation (Wiederstein & Sippl et al., 2007).

Moreover, the pdB file was uploaded; basically, the server and the chart were obtained. Notably, the z-score vs number of residue graph, a knowledge-based energy vs sequence position graph, was seen as the outcome.

2.11 Molecular docking of the constructed vaccine

Molecular docking plays a central role in evaluating the binding compatibility between the designed vaccine constructs and their target immune receptor. However, in the vaccine construction process, this approach allows for predicting how multi-epitope constructs interact with receptors such as TRL2, TRL3, and TRL4 (Comeau et al.,2004). The ClusPro server was used to check the binding affinity. The pdB file for TLR was downloaded from the RCSB database and uploaded along with the main pdB file as a ligand. Now, the result page will show 10 docked complexes, and one will be selected among them.

2.12 Immune Response Stimulation

Immune stimulation plays a vital role in vaccine development as it is used to establish the efficiency of the construct in the stimulation of strong humoral and cell-mediated responses. Proper engagement of receptors like Toll-like receptors, which enhances antigen processing and promotes strong immunogenicity, making immune stimulation a critical component in evaluating vaccine efficacy (Kawai & Akira et al., 2011). The C-ImmSim server was used, where graphical information about the levels of antigen, lymphocytes, immunoglobulins, and so on was provided. Importantly, here the vaccine sequence was provided, and three injections were added where the simulation step 300 was selected and the time steps of the injections were 1,82 and 168, respectively.

Chapter 3

Results

3.1 Antigenicity and protein selection

Name of protein	Lysosome membrane protein 2
Gene	SCARB2
Amino acid	478
Organism	Homo sapiens (Human)
Status	UniprotKB reviewed (Swiss-Prot)

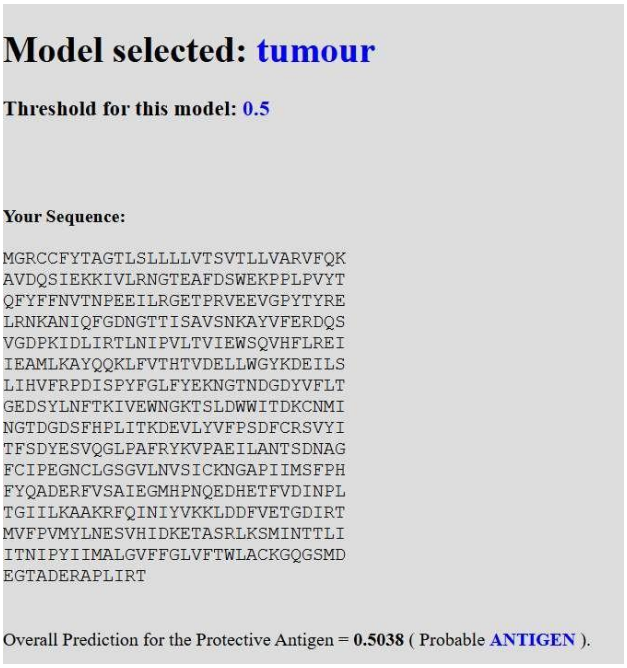
Protein sequence:

MGRCCFYTAGTLSLLLLVTSVTLLVARVFQK
AVDQSIEKKIVLRNGTEAFDSWEKPPLPVYT
QFYFFNVTNPEEILRGETPRVEEVGPYTYRE
LRNKANIQFGDNGTTISAVSNKAYVFERDQS
VGDPKIDLIRTLNIPVLTVIEWSQVHFLREI
IEAMLKAYQQKLFVTHTVDELLWGYKDEILS
LIHVFRPDISPYFGLFYEKNGTNDGDYVFLT
GEDSYLNFTKIVEWNGKTSLDWWITDKCNMI
NGTDGDSFHPLITKDEVLYVFPSDFCRSVYI

TFSDYESVQGLPAFRYKVPAEILANTSDNAG
FCIPEGNCLGSGVLNVSICKNGAPIIMSFPH
FYQADERFVSAIEGMHPNQEDHETFVDINPL
TGIILKAAKRFQINIYVKKLDDFVETGDIRT
MVFPVMYLNESVHIDKETASRLKSMINTTLI
ITNIPYIIMALGVFFGLVFTWLACKGQGSMD
EGTADERAPLIRT

Antigenicity

The VaxiJen v2.0 server was used to find out the antigenicity of the selected protein.



Model selected: tumour

Threshold for this model: 0.5

Your Sequence:

MGRCCFYTAGTLSLLLLVTSVTLLVARVFQK
AVDQSIEKKIVLRNGTEAFDSWEKPPLPVYT
QFYFFNVTNPEEILRGETPRVEEVGPYTYRE
LRNKANIQFGDNGTTISAVSNKAYVFERDQS
VGDPKIDLIRTLNIPVLTVIEWSQVHFLREI
IEAMLKAYQQKLFVTHTVDELLWGYKDEILS
LIHVFRPDISPYFGLFYEKNGTNDGDYVFLT
GEDSYLNFTKIVEWNGKTSLDWWITDKCNMI
NGTDGDSFHPLITKDEVLYVFPSDFCRSVYI
TFSDYESVQGLPAFRYKVPAEILANTSDNAG
FCIPEGNCLGSGVLNVSICKNGAPIIMSFPH
FYQADERFVSAIEGMHPNQEDHETFVDINPL
TGIILKAAKRFQINIYVKKLDDFVETGDIRT
MVFPVMYLNESVHIDKETASRLKSMINTTLI
ITNIPYIIMALGVFFGLVFTWLACKGQGSMD
EGTADERAPLIRT

Overall Prediction for the Protective Antigen = **0.5038** (Probable **ANTIGEN**).

Figure 2: Antigenicity score of the protein sequence (Doytchinova & Flower, 2007)

3.2 Selection of CTL epitopes

The CTL epitopes are identified from the NetCTL1.2 server, which has strong binding with the MHC class I peptides. From the server, 12 epitopes were found with strong bond. And these epitopes were selected on the basis of their antigenicity, allergenicity, and toxicity.

Table 1: Selected epitopes with Antigenicity, allergenicity, and toxicity.

Peptide	Antigenicity	Allergenicity	Toxicity
TVDELLWGY	Non-Antigen	Non allergen	Non toxin
ISPYFGLFY	Antigen	Non allergen	Non toxin
PSDFCRSVY	Non antigen	Non allergen	Non toxin
RTMVFPVMY	Non antigen	Non allergen	Non toxin
ISAVSNKAY	Antigen	Allergen	Non toxin
IIMSFPHFY	Antigen	Allergen	Non toxin
RTMVFPVMY	Non antigen	Non allergen	Non toxin
TVDELLWGY	Non antigen	Non allergen	Non toxin
SVYITFSDY	Non antigen	Allergen	Non toxin
YTAGTLSLL	Non antigen	Allergen	Non toxin
TLIITNIPY	Non antigen	Allergen	Non toxin
DISPYFGLF	Antigen	Non allergen	Non toxin

3.3 Selection of HTL epitopes

From the server, NetMHCIIpan 4.0, the strong binding epitopes were selected by MHC class II alleles. The HTL epitopes were selected on the basis of IFN, IL4, and IL10, along with antigenicity, allergenicity, and toxicity. The selected epitopes should match the criteria of antigen, non-allergenic, and non-toxic. Moreover, IFN should be positive, IL4 should be an inducer, and IL10 should be an inducer as well.

Table 2: Selected HTL epitopes

Peptide	Antigenicity	Allergenicity	Toxicity	IFN	IL4	IL10
VSNKAYVFERDQSVG	Antigen	Non allergen	Non toxin	Negative	Inducer	Inducer
SNKAYVFERDQSVGD	Antigen	Non allergen	Non toxin	Positive	Inducer	Inducer
NKAYVFERDQSVGDP	Antigen	Allergen	Non toxin	Positive	Inducer	Inducer
KAYVFERDQSVGDPK	Antigen	Non allergen	Non toxin	Positive	Inducer	Inducer
TNDGDYVFLTGEDSY	Antigen	Allergen	Non toxin	Negative	Non inducer	Inducer
NDGDYVFLTGEDSYL	Antigen	Allergen	Non toxin	Negative	Non inducer	Inducer
DGDYVFLTGEDSYLN	Antigen	Allergen	Non	Negative	Non	Inducer
GDYVFLTGEDSYLNF	Antigen	Allergen	Non toxin	Negative	Non inducer	Inducer
ITFSDYESVQGLPAF	Non antigen	Allergen	Non toxin	Negative	Inducer	Non inducer
TFSDYESVQGLPAFR	Non antigen	Allergen	Non toxin	Negative	Inducer	Non inducer
FSDYESVQGLPAFRY	Non antigen	Allergen	Non toxin	Positive	Inducer	Inducer

SDYESVQGLPAFRYK	Non antigen	Allergen	Non toxin	Negative	Inducer	Non-Inducer
ETFVDINPLTGIILK	Non antigen	Non allergen	Non toxin	Negative	Inducer	Non inducer
TFVDINPLTGIILKA	Non antigen	Non allergen	Non toxin	Negative	Inducer	Non inducer
FVDINPLTGIILKAA	Non antigen	Non allergen	Non toxin	Negative	Non inducer	Non inducer
VDINPLTGIILKAAK	Non antigen	Non allergen	Non toxin	Negative	Non inducer	Non inducer
ESVHIDKETASRLKS	Antigen	Non allergen	Non toxin	Negative	Inducer	Inducer
SVHIDKETASRLKSM	Antigen	Non allergen	Non toxin	Negative	Inducer	Inducer
EKKIVLRNGTEAFDS	Antigen	Non allergen	Non toxin	Negative	Non-Inducer	Non-Inducer
GLPAFRYKVPAEILA	Non antigen	Non-Allergen	Non toxin	Negative	Inducer	Non inducer
LPAFRYKVPAEILAN	Non antigen	Non allergen	Non toxin	Negative	Inducer	Non inducer
PAFRYKVPAEILANT	Non antigen	Allergen	Non toxin	Positive	Inducer	Non inducer
FHPLITKDEVLYVFP	Antigen	Non allergen	Non toxin	Negative	Non inducer	Non inducer
HPLITKDEVLYVFPS	Antigen	Non allergen	Non toxin	Negative	Non inducer	Non inducer
YLNESVHIDKETASR	Antigen	Non allergen	Non toxin	Negative	Inducer	Inducer
LNESVHIDKETASRL	Antigen	Non allergen	Non toxin	Positive	Inducer	Inducer
NESVHIDKETASRLK	Antigen	Non allergen	Non toxin	Negative	Inducer	Inducer
AYVFERDQSVGDPKI	Non antigen	Non allergen	Non toxin	Positive	Inducer	Inducer

SLIHVFRPDISPYFG	Non antigen	Non allergen	Non toxin	Negative	Inducer	Inducer
LIHVFRPDISPYFGL	Antigen	Non allergen	Non toxin	Positive	Inducer	Inducer
IHVFRPDISPYFGLF	Antigen	Allergen	Non toxin	Positive	Inducer	Inducer
HVFRPDISPYFGLFY	Antigen	Allergen	Non toxin	Positive	Inducer	Inducer
SFPHFYQADERFVSA	Non antigen	Non allergen	Non toxin	Positive	Inducer	Non inducer
FPHFYQADERFVSAI	Non antigen	Non allergen	Non toxin	Positive	Inducer	Non inducer
PHFYQADERFVSAIE	Non antigen	Non allergen	Non toxin	Negative	Inducer	Non inducer
HFYQADERFVSAIEG	Non antigen	Non allergen	Non toxin	Positive	Inducer	Non inducer
AYVFERDQSVGDPKI	Non antigen	Non allergen	Non toxin	Negative	Inducer	Inducer
YTQFYFFNVTNPEEI	Antigen	Non allergen	Non toxin	Positive	Inducer	Non inducer
TQFYFFNVTNPEEIL	Antigen	Non allergen	Non toxin	Positive	Non inducer	Non inducer
QFYFFNVTNPEEILR	Antigen	Allergen	Non toxin	Positive	Non inducer	Non inducer
VPAEILANTSDNAGF	Antigen	Non allergen	Non toxin	Negative	Inducer	Non inducer
VSAIEGMHPNQEDHE	Non antigen	Non allergen	Non toxin	Negative	Inducer	Non inducer
FPVMYLNESVHIDKE	Non antigen	Non allergen	Non toxin	Negative	Inducer	Inducer
PVMYLNESVHIDKET	Antigen	Non allergen	Non toxin	Negative	Inducer	Inducer
VYTQFYFFNVTNPEE	Antigen	Non allergen	Non toxin	Positive	Non inducer	Non inducer
FYFFNVTNPEEILRG	Antigen	Allergen	Non toxin	Positive	Non inducer	Non inducer

RSVYITFSDYESVQG	Antigen	Allergen	Non toxin	Negative	Inducer	Non inducer
VFPVMYLNESVHIDK	Non antigen	Non-Allergen	Non toxin	Negative	Inducer	Inducer
FPVMYLNESVHIDKE	Non antigen	Non-Allergen	Non toxin	Negative	Inducer	Inducer
VARVFQKAVDQSIEK	Antigen	Non allergen	Non toxin	Negative	Inducer	Non inducer
ARVFQKAVDQSIEKK	Antigen	Non allergen	Non toxin	Positive	Inducer	Inducer
LREIIEAMLKAYQQK	Antigen	Allergen	Non toxin	Positive	Inducer	Inducer
TGEDSYLNFTKIVEW	Antigen	Allergen	Non toxin	Negative	Inducer	Inducer
GEDSYLNFTKIVEWN	Antigen	Allergen	Non toxin	Negative	Inducer	Inducer
EDSYLNFTKIVEWNG	Antigen	Allergen	Non toxin	Negative	Inducer	Inducer
DSYLNFTKIVEWNGK	Antigen	Allergen	Non toxin	Negative	Inducer	Non
DHETFVDINPLTGII	Non antigen	Non allergen	Non toxin	Negative	Inducer	Non
FLREIIEAMLKAYQQ	Non antigen	Allergen	Non toxin	Positive	Inducer	Inducer
REIIEAMLKAYQQKL	Non antigen	Allergen	Non toxin	Positive	Inducer	Inducer
EIIEAMLKAYQQKLF	Non antigen	Allergen	Non toxin	Positive	Inducer	Inducer
PLTGIILKAAKRFQI	Non antigen	Non allergen	Non toxin	Negative	Non inducer	Inducer
LTGIILKAAKRFQIN	Non antigen	Allergen	Non toxin	Negative	Non inducer	Inducer
TGIILKAAKRFQINI	Non antigen	Allergen	Non toxin	Negative	Non inducer	Inducer
LVTSVTLLVARVFQK	Non antigen	Non allergen	Non toxin	Negative	Non inducer	Inducer

VTSVTLLVARVFQKA	Non antigen	Allergen	Non toxin	Negative	Non inducer	Inducer
TSVTLLVARVFQKAV	Non antigen	Allergen	Non toxin	Negative	Non inducer	Inducer
DNGTTISAVSNKAYV	Antigen	Non allergen	Non toxin	Positive	Inducer	Inducer
NGTTISAVSNKAYVF	Antigen	Allergen	Non toxin	Positive	Inducer	Inducer
GTTISAVSNKAYVFE	Antigen	Non allergen	Non toxin	Positive	Inducer	Inducer
TTISAVSNKAYVFER	Antigen	Non allergen	Non toxin	Positive	Inducer	Inducer
VGDPKIDLIRTLNIP	Antigen	Allergen	Non toxin	Negative	Inducer	Non inducer
GDPKIDLIRTLNIPV	Antigen	Non allergen	Non toxin	Negative	Inducer	Non inducer
DPKIDLIRTLNIPVL	Antigen	Non allergen	Non toxin	Negative	Non inducer	Non inducer
FSDYESVQGLPAFRY	Non antigen	Allergen	Non toxin	Positive	Inducer	Inducer
DYESVQGLPAFRYKV	Non antigen	Allergen	Non toxin	Negative	Inducer	Inducer
YESVQGLPAFRYKVP	Non antigen	Allergen	Non toxin	Positive	Inducer	Non inducer
ESVQGLPAFRYKVPA	Non antigen	Non allergen	Non toxin	Negative	Inducer	Non-Inducer
SVQGLPAFRYKVPAE	Non antigen	Allergen	Non toxin	Negative	Inducer	Non inducer
NGAPIIMSFPHFYQA	Antigen	Allergen	Non toxin	Negative	Inducer	Non inducer
GAPIIMSFPHFYQAD	Non antigen	Allergen	Non toxin	Negative	Inducer	Non inducer
ETFVDINPLTGIILK	Non antigen	Non	Non	Negative	Inducer	Non inducer

APIIMSFPHFYQADE	Antigen	Allergen	Non toxin	Positive	Inducer
TGIILKAAKRFQINI	Non antigen	Allergen	Non toxin	Negative	Non inducer
GIILKAAKRFQINIY	Non	Non-Allergen	Non toxin	Negative	Non inducer
AKRFQINIYVKKLDD	Antigen	Non-Allergen	Non toxin	Negative	Non inducer
KRFQINIYVKKLDDF	Antigen	Allergen	Non	Negative	Inducer
RFQINIYVKKLDDFV	Antigen	Allergen	Non	Negative	Inducer
EKKIVLRNGTEAFDS	Antigen	Non allergen	Non toxin	Negative	Inducer
SFPHFYQADERFVSA	Non	Non	Non toxin	Negative	Inducer
GDPKIDLIRTLNIPV	Antigen	Non	Non toxin	Negative	Inducer
REIIEAMLKAYQQKL	Non-Antigen	Allergen	Non toxin	Positive	Inducer
LTGIILKAAKRFQIN	Non-Antigen	Allergen	Non toxin	Negative	Non inducer
FRPDISPYFGLFYEK	Antigen	Allergen	Non toxin	Negative	Inducer
RPDISPYFGLFYEKN	Antigen	Non	Non toxin	Negative	Inducer
VSNKAYVFERDQSVG	Antigen	Non allergen	Non toxin	Negative	Inducer
LSLIHVFRPDISPYF	Non-Antigen	Non	Non	Negative	Inducer
LIHVFRPDISPYFGL	Antigen	Non allergen	Non toxin	Positive	Inducer
SLDWWITDKCNMING	Antigen	Allergen	Non toxin	Positive	Inducer
LDWWITDKCNMINGT	Antigen	Allergen	Non toxin	Positive	Inducer

DWWITDKCNMINGTD	Antigen	Allergen	Non toxin	Positive	Inducer	Non inducer
TQFYFFNVTNPEEIL	Antigen	Non	Non toxin	Positive	Non	Non inducer
FGLFYEKNGTNDGDY	Antigen	Non	Non toxin	Negative	Non	Non inducer
GLFYEKNGTNDGDYV	Antigen	Non	Non toxin	Negative	Non	Non inducer
KVPAEILANTSDNAG	Antigen	Non	Non toxin	Negative	Inducer	Non inducer
PAEILANTSDNAGFC	Antigen	Allergen	Non toxin	Negative	Inducer	Non inducer
LVARVFQKAVDQSIE	Non antigen	Non allergen	Non toxin	Negative	Inducer	Non inducer
EILSLIHVFRPDISP	Non antigen	Non allergen	Non toxin	Positive	Inducer	Inducer
ILSLIHVFRPDISPY	Non antigen	Allergen	Non toxin	Positive	Inducer	Inducer
SLIHVFRPDISPYFG	Non antigen	Non allergen	Non toxin	Negative	Inducer	Inducer
FRPDISPYFGLFYEK	Antigen	Allergen	Non toxin	Negative	Inducer	Inducer

3.4 Selection of B-cell epitopes

To determine B-cell epitopes, the IEDB Analysis Resource tool was used. Here, the Bepipred Linear Epitope Prediction 2.0 method was chosen as it's more advanced. A graph was obtained here, which was score vs position, where the threshold was set to 0.5. In the graph yellow bar indicates the most desired B epitopes, and the green bar indicates those not desired.

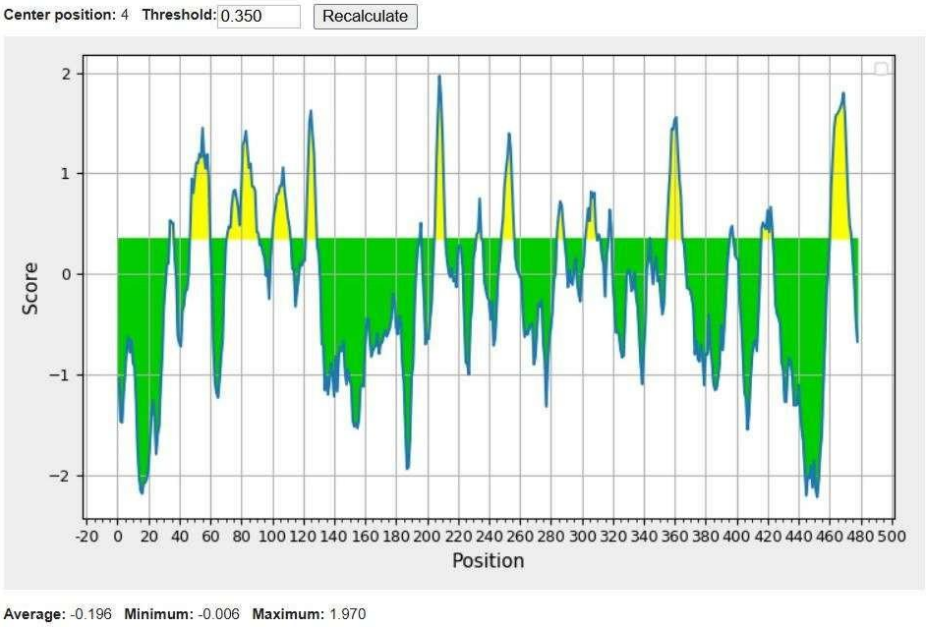


Figure 3: B-cell epitope graph (Jespersen et al., 2017)

B-cell epitopes were selected based on the criteria of allergenicity, antigenicity, and toxicity. The length of the peptides must be in the range of 7-30. Only the epitope that fulfills all the requirements should be selected for the construction of a vaccine.

Table 3: Selected B-cell epitopes.

Peptide	Length	Antigenicity	Allergenicity	Toxicity
TEAFDSWEKPPLP	13	Non-antigen	Non-allergen	Non-toxin
NPEEILRGETPRVEEVGPYTY	21	Antigen	Allergen	Non-toxin
IQFGDNGTTISA	12	Antigen	Non-allergen	Non-toxin
DQSVGDP	7	Non-antigen	Non-allergen	Non-toxin
KNGTNDG	7	Antigen	Non-allergen	Non-toxin
INGTDGDSF	9	Antigen	Non-allergen	Non-toxin
ANTSDNA	7	Antigen	Allergen	Non-toxin
GMHPNQEDHE	10	Non-antigen	Non-allergen	Non-toxin
HIDKETAS	8	Non-antigen	Non-allergen	Non-toxin
QGSMDEGTADERAP	14	Antigen	Allergen	Non-toxin

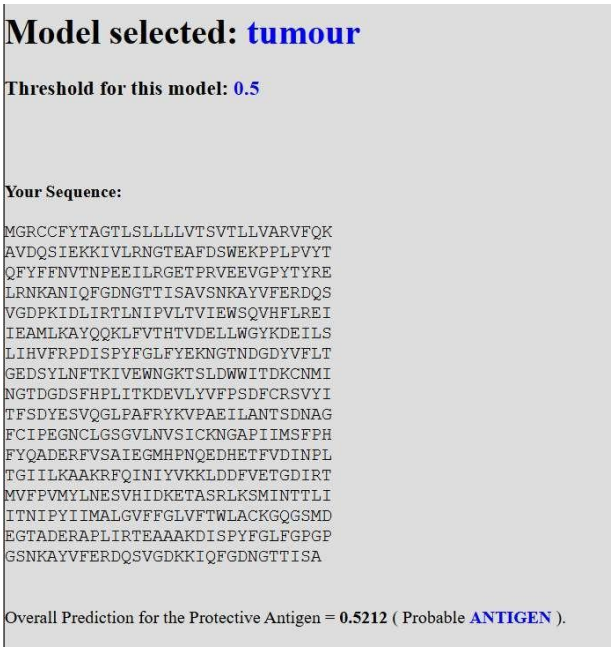
3.5 Construction of the final vaccine

The newly constructed vaccine is combined with the eligible and criteria-matched peptides. Linkers are used in multi-peptide vaccine construction to ensure the structural stability and immunological effectiveness. The linkers that are used in the construction of the final vaccine are EAAAK, GPGPG, and KK. Along with the linkers, criteria-matched CTL, HTL, and B cells are included in the sequence.

MGRCCFYTAGTLSLLLLVTSVTLLVARVFQKAVDQSIEKKIVLRNGTEAFDSWEKPPLPV
YTQFYFFNVTNPEEILRGETPRVEEVGPYTYRELRNKANIQFGDNGTTISAVSNKAYVFE
RDQSVGDPKIDLIRTLNIPVLTVIEWSQVHFLREIIEAMLKAYQQKLFVTHTVDELLWGY
KDEILSLIHVFRPDISPYFGLFYEKNGTNDGDYVFLTGEDSYLNFTKIVEWNGKTSLDWW
ITDKCNMINGTDGDSFHPLITKDEVLYVFPSDFCRSVYITFSDYESVQGLPAFRYKVPAE
ILANTSDNAGFCIPEGNCLGSGVLNVSICKNGAPIIMSFPHFYQADERFVSAIEGMHPNQ
EDHETFVDINPLTGIILKAAKRFQINIYVKKLDDFVETGDIRTMVFPVMYLNESVHIDKE
TASRLKSMINTTLIITNIPYIIMALGVFFGLVFTWLACKGQGSMDEGTADERAPLIRTEAAAK
DISPYFGLFGPGPGSNKAYVFERDQSVGDKKIQFGDNGTTISA

3.6 Antigenicity of the constructed vaccine

The antigenicity of the constructed vaccine is evaluated from the server VaxiJen v2.0. Here, the antigenic score is 0.5212 with the threshold set to 0.5.



Model selected: tumour

Threshold for this model: 0.5

Your Sequence:

MGRCCFYTAGTLSLLLLVTSVTLLVARVFQK
AVDQSIEKKIVLRNGTEAFDSWEKPPLPVYT
QFYFFNVTNPEEILRGETPRVEEVGPYTYRE
LRNKANIQFGDNGTTISAVSNKAYVFERDQS
VGDPKIDLIRTLNIPVLTVIEWSQVHFLREI
IEAMLKAYQQKLFVTHTVDELLWGYKDEILS
LIHVFRPDISPYFGLFYEKNGTNDGDYVFLT
GEDSYLNFTKIVEWNGKTSLDWWITDKCNMI
NGTDGDSFHPLITKDEVLYVFPSDFCRSVYI
TFSDYESVQGLPAFRYKVPAEILANTSDNAG
FCIPEGNCLGSGVLNVSICKNGAPIIMSFPH
FYQADERFVSAIEGMHPNQEDHETFVDINPL
TGIILKAAKRFQINIYVKKLDDFVETGDIRT
MVFPVMYLNESVHIDKETASRLKSMINTTLI
ITNIPYIIMALGVFFGLVFTWLACKGQGSMD
EGTADERAPLIRTEAAAKDISPYFGLFGPGP
GSNKAYVFERDQSVGDKKIQFGDNGTTISA

Overall Prediction for the Protective Antigen = **0.5212** (Probable **ANTIGEN**).

Figure 4: Antigenicity of the constructed vaccine.

3.7 Biochemical analysis of the constructed vaccine

The biochemical properties of the vaccine are evaluated by using the Protparam server. The results, such as GRAVY, instability index, molecular weight, and molecular formula, are provided by this server. Moreover, it provides information about the therapeutic pI, molecular weight, and amino acid. These features are evaluated for an eligible vaccine candidate.

Number of amino acids: 526

Theoretical pI: 4.99

Molecular weight: 59324.69

Instability index:

The instability index (II) is computed to be 33.60.

This classifies the protein as stable.

Aliphatic index: 89.11

Grand average of hydropathicity (GRAVY): -0.081

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).

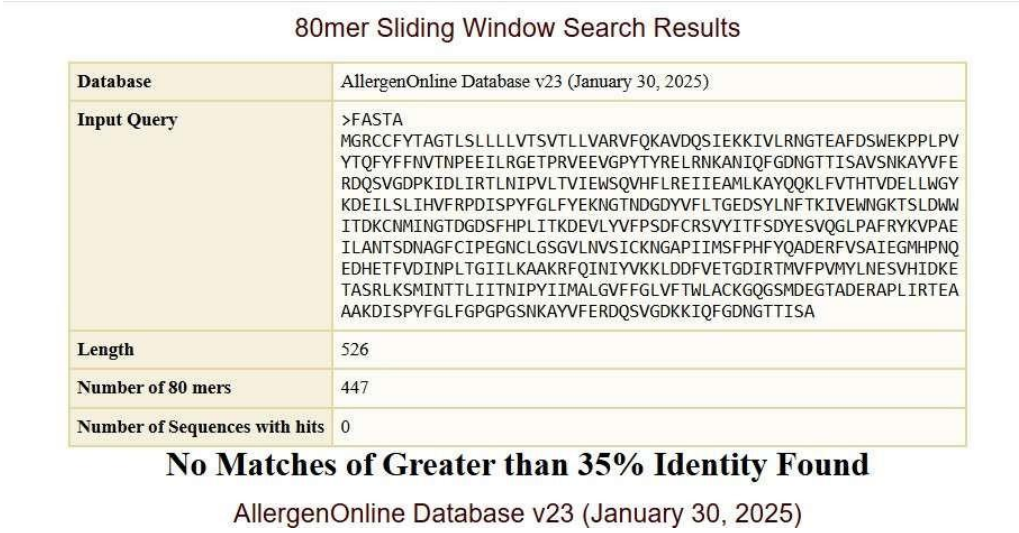
Instability index:
The instability index (II) is computed to be 33.60
This classifies the protein as stable.

Aliphatic index: 89.11
Grand average of hydropathicity (GRAVY):-0.081

Figure 5: Biochemical prediction of the constructed vaccine (Gasteiger et al.,2005.)

3.8 Evaluating the allergenicity and toxicity of the constructed vaccine

The allergenicity of the constructed vaccine is evaluated using the Allergen Online server. The sequence is provided in FASTA format using an 80-mer sliding window. The number of sequences with hits was observed to be 0.



80mer Sliding Window Search Results

Database	AllergenOnline Database v23 (January 30, 2025)
Input Query	>FASTA MGRCCFYTAGTLSLLLLVTSVTLLVARVFQKAVDQSIEKKIVLRNGTEAFDSWEKPPLPV YTQFYFFNVTNPEEILRGETPRVEEVGPYTYRELRNKANIQFGDNGTTISAVSNKAYVFE RDQSVGDPKIDLIRTLNIPVLTVIEWSQVHFLREIIEAMLKAYQQKLFVTHTVDELLWGY KDEILSLIHVFRPDISPYFGLFYEKNGTNDGDYVFLTGEDSYLNFTKIVEWNGKTSLDWW ITDKCNMINGTDGDSFHPLITKDEVLYVFPSDFCRSVYITFSDYESVQGLPAFRYKVPAE ILANTSDNAGFCIPEGNCLGSGVLNVSICKNGAPIIMSFPHFYQADERFVSAIEGMHPNQ EDHETFVDINPLTGIILKAAKRFQINIYVKKLDDFVETGDIRTMVFPVMYLNESVHIDKE TASRLKSMINTTLIITNIPYIIMALGVFFGLVFTWLACKGQGSMDEGTADERAPLIRTEA AAKDISPYFGLFGPGPGSNKAYVFERDQSVGDKKIQFGDNGTTISA
Length	526
Number of 80 mers	447
Number of Sequences with hits	0

No Matches of Greater than 35% Identity Found

AllergenOnline Database v23 (January 30, 2025)

Figure 6: Allergenicity identification of the constructed vaccine (Goodman et al., 2016)

Additionally, the toxicity of the constructed vaccine was evaluated on the server using T3DB. This server helps to determine the presence of harmful metabolites or particles. Thus, the result returns with no toxicity.

Figure 7: Toxicity of the constructed vaccine (Wishart et al., 2015)

3.9 3D model of the constructed vaccine

The 3D model of the vaccine was constructed using the server Phyre2. The server shows that the constructed vaccine has coverage of 75% and confidence is 100%. In addition, 392 residues have been modelled with 100% confidence, as viewed by the 3D model. The coverage of the constructed vaccine came out 75% instead of 100% because not all the selected epitopes match the HLA allele, some of these limit the compatibility. Thus, the alleles are not covered by the chosen epitopes, which results in reducing the coverage of the vaccine.

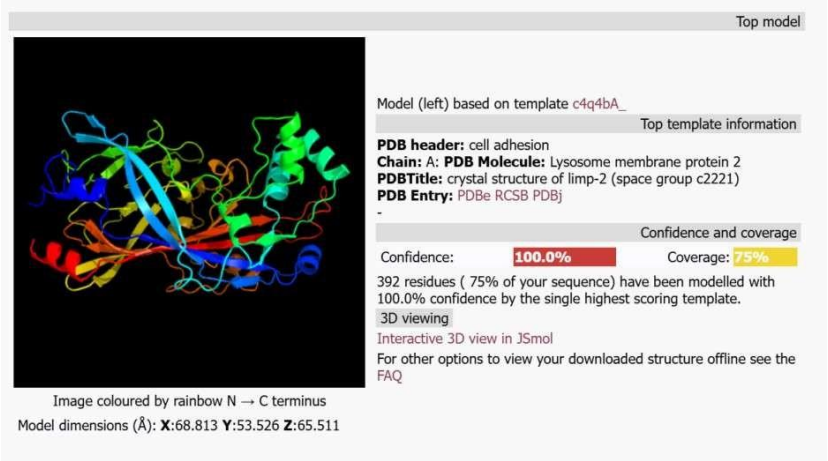


Figure 8: 3D model of the constructed vaccine with confidence and coverage (Kelley et al., 2015)

3.10 Ramachandran Plotting

Ramachandran plot was generated by the Swiss Model Expasy server that provides an assessment of the stereochemical quality of the vaccine. From the MolProbity results, we find that Ramachandran Favored was 95.64%, Ramachandran Outliers was 0.26%. The acceptable range for the Ramachandran favored region is greater than 90% and the outlier range is less than 2%. Thus, the result obtained is within the desired range.

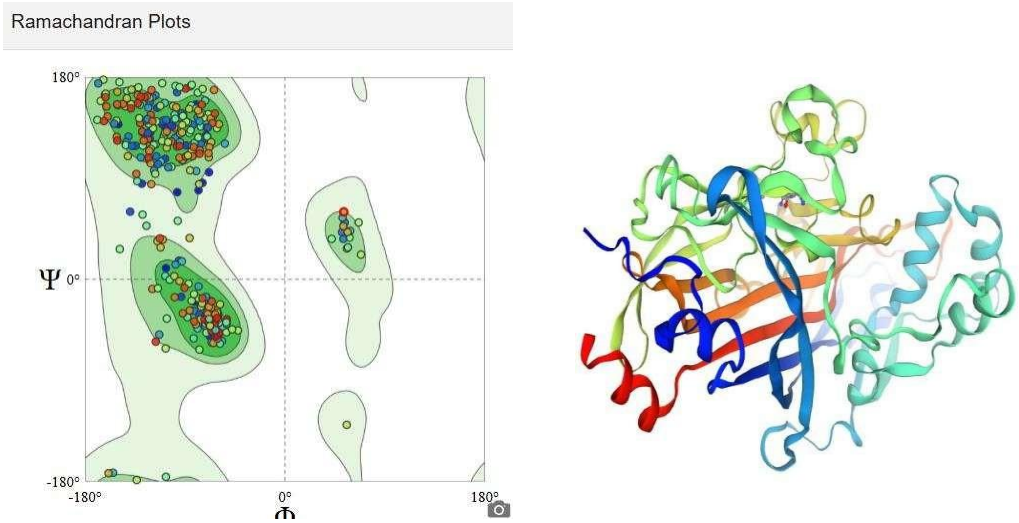


Figure 9: Ramachandran plot on Swiss Model (Waterhouse et al., 2018)

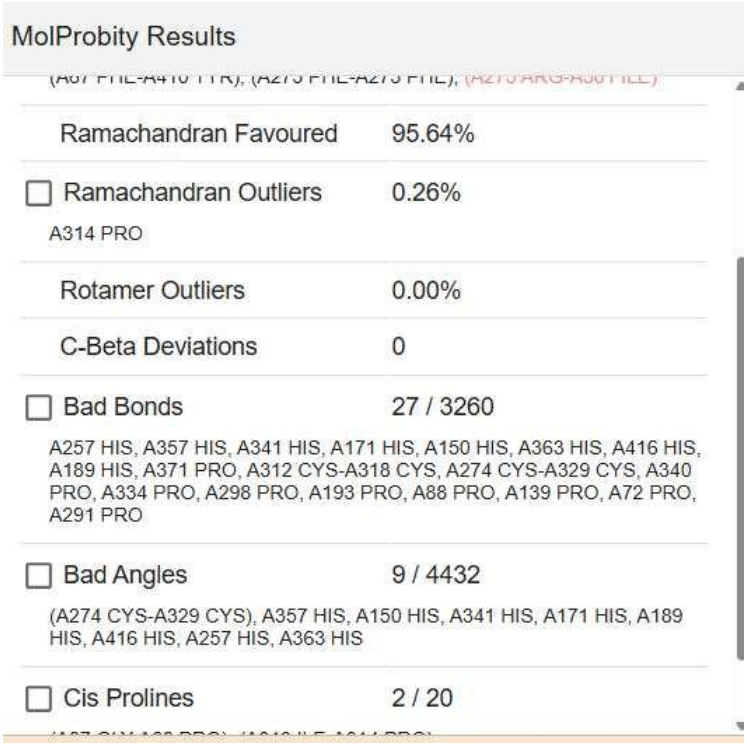


Figure 10: MolProbility results found from the Swiss-Model tool (Schwede et al. 2003)

3.11 Analysis of Z-score

To determine the Z-score of the constructed vaccine ProSA-web server was used. The server shows a graph of Z-score vs Number of residues, where the score was -9.27. The structures from different regions, such as X-ray, NMR, are observed in distinct colors. Moreover, in the second graph, the energy demonstrates local model quality, where most of the residues are in negative form except the N-terminal region. Here, in the graph, the darker region indicates the normal range of negative proteins, which represents a good model, and the lighter region indicates deviation from the negative range, meaning a poor or misfolded model. However, the Z-score was within the desired range, thus the proteins are good and correctly folded (Wiederstein & Sippl, 2007).

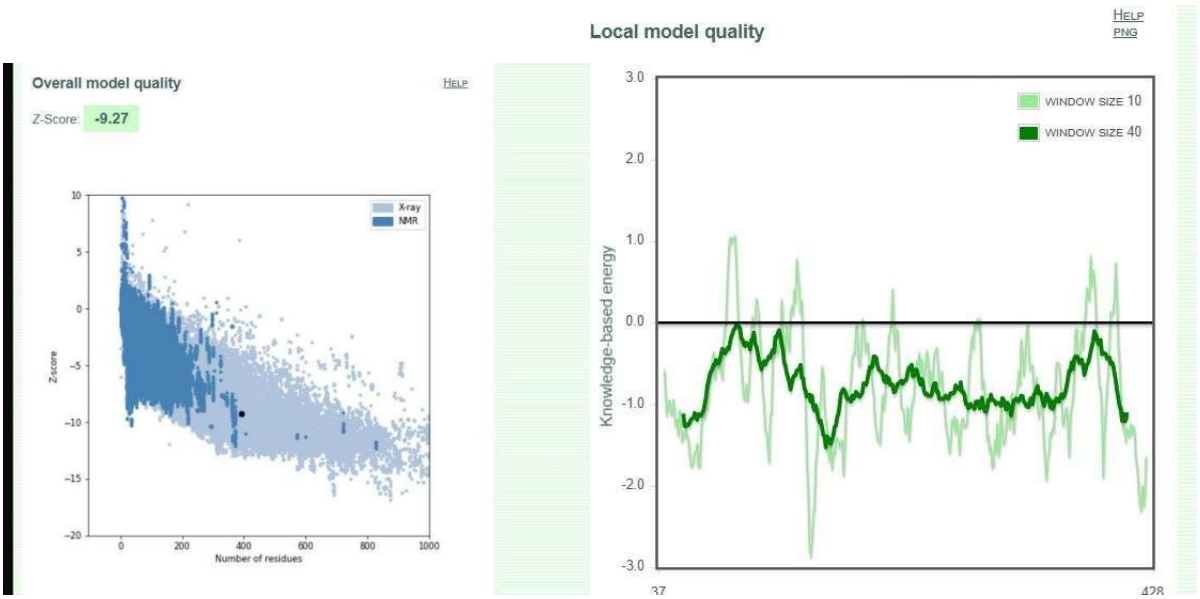
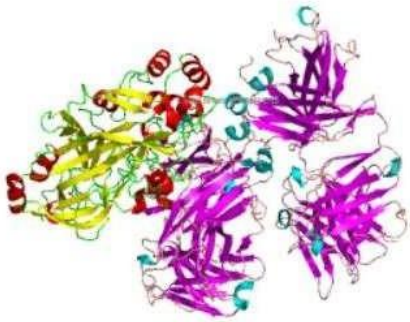


Figure 11: Z-score of the constructed vaccine (Wiederstein & Sippl et al., 2007)

3.12 Molecular Docking of the Constructed Vaccine

Molecular docking was done by using the server "ClusPro2.0: Protein-Protein docking". Here, human Toll-Like Receptor 3 was used to bind with the ligand, and the pdB file obtained was inserted into the receptor folder. After completion of docking, several docked results were seen. Among them, the lowest energy with the most stable dock was chosen as the docked model.

0



2

Cluster Scores

We strongly encourage you to read the FAQ related to these scores before using them.

Cluster	Members	Representative	Weighted Score
0	75	Center	-856.1
		Lowest Energy	-856.1
1	42	Center	-641.3
		Lowest Energy	-769.0

Figure 12: Molecular docking and Cluster score from ClusPro (Comeau et al.,2004)

3.13 Immune Simulation of the Constructed Vaccine

Immune simulation of the constructed vaccine is done by the C-ImmSim browser to check the antibody production ability in the body. It predicts the immune response over time. In this server, injections are provided in various intervals, and the result is seen through antigen count per mL Vs number of days on the graph. The black line of the graph indicates the antigen, and the others are the antibody. Firstly, for injection 1, the simulation steps were 300, and the time step of injection 1 was given as the input, along with the sequence of the newly constructed vaccine in FASTA format. Secondly, for injection 2, the time step was 84, and lastly, for injection 3, the time step was 168.Thus, the output resulted in a graph where the green line indicates the antigen concentration refers to the peak after each injection.

Again, the red line indicates the antibody (IgM, IgG) that shows the strength and types of antibody response. Lastly, the blue line indicates the B cell population, which is B cell activation and proliferation over time. An ideal vaccine response has high IgG after the second or third dose, persistent memory B and T cells, and antigen clearance, that is decrease in the number of green lines. Comparing the ideal criteria with the newly contrasted vaccine response, we select this as a vaccine candidate.

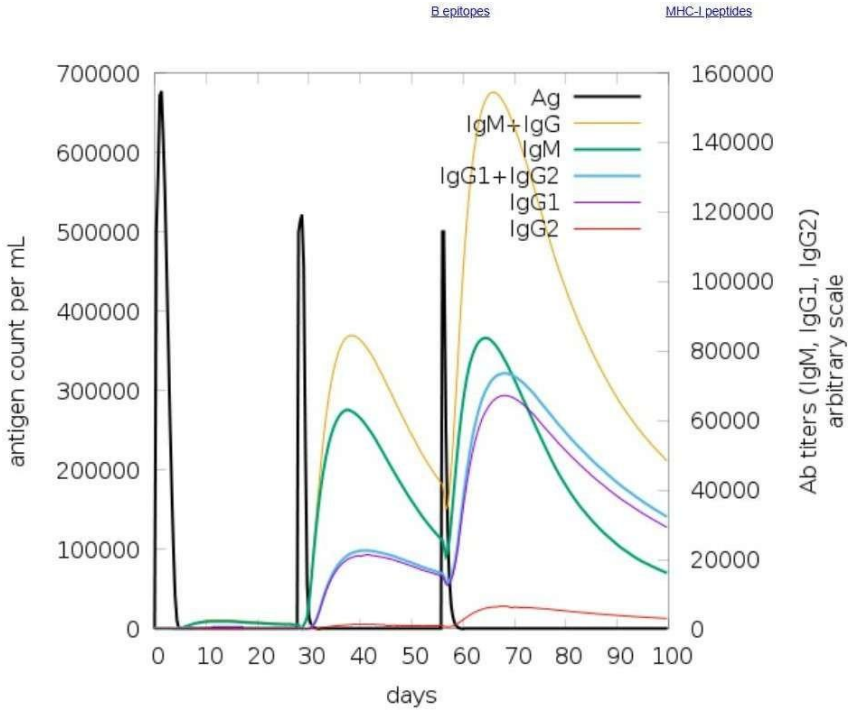


Figure 13: Immune simulation of the constructed vaccine using C-Immsim (Rapin et al., 2010)

Chapter 4

Discussion

Breast cancer is one of the leading causes of cancer-related mortality among women in the world, which is defined by the uncontrolled growth of the number of breast cells caused by genetic factors, hormones, and various environmental factors. The immunoinformatics advances have promoted the idea of vaccine-based approaches that are focused on eliciting selective immune responses to antigens associated with tumors. Survival can be achieved by early detection through screening, targeted therapy, surgery, radiation, and immunotherapy. In silico investigation provides a comprehensive analysis of breast cancer antigens and immunorelevant epitopes using various computational tools and databases. Suitable protein sequences from the UniProt database should be selected. Using reliable immunoinformatic servers, cytotoxic T lymphocyte (CTL), helper T lymphocyte (HTL), and B-cell epitopes are used to identify regions with strong immunogenic potential. Similar computational strategies have been widely recognized for enabling the rational design of vaccines and therapeutics (Vita et al., 2019). The 3D structural model is generated through servers such as SWISS-MODEL, showing favorable stereochemical quality. This was confirmed through Ramachandran plot analysis, where most residues were found in favored regions, and further supported by an acceptable Z-score, indicating good model stability. Such structural validation remains an essential criterion for ensuring accurate docking and immunological simulation (Waterhouse et al., 2018). To evaluate innate immune recognition potential, the final vaccine construction was docked with Toll-like receptor 3 (TLR3) using standard protein–protein docking platforms. The predicted binding orientation and docking energy suggested strong interactions with the pdB file obtained, which could support effective immune activation. Protein–protein docking has previously been shown to enhance understanding of receptor–ligand dynamics in vaccine development (Pierce et al., 2014). The analysis also included molecular docking of selected CTL and HTL epitopes with their corresponding MHC molecules. In fact, strong binding affinities and stable hydrogen-bond interactions indicated that the epitopes are likely presented effectively to T cells. In addition, the predicted toxicity of each epitope was assessed using toxicity-screening servers, ensuring that only non-toxic and biologically safe candidates were included, which is essential for developing a safe vaccine (Gupta et al., 2013)

To simulate the immune system, the multi-epitope vaccine was assessed using C-ImmSim, which generated graphs showing elevated T-cell counts, memory cell formation, and consistent B-cell responses. These patterns indicate an ability to induce long-lasting immunity, aligning with findings from prior computational immunology studies demonstrating the predictive accuracy of immune simulation (Rapin et al., 2010).

Chapter 5

Conclusion

Breast cancer is one of the worst diseases that women can experience globally. Currently, no vaccine exists on the market to mitigate the disease. There are therapies that offer hope against breast cancer, but in reality, they also carry additional burdens and adverse health effects. This research aimed to develop an in-silico vaccine for breast cancer. To form the vaccine, in silico tools were used to assess its toxicity, allergenicity, and antigenicity. Moreover, an assessment of the 3D model of the vaccine was made, and binding ability and structural analysis were done for validation. The goal is to support the quality of the created vaccine. Furthermore, in compliance with good manufacturing practices, in vitro and in vivo assessments are necessary for the quality, safety, and effectiveness of the vaccine. Developed vaccine limitations include that findings assessed do not include greater than an 80% coverage, which means all safety and effectiveness of the vaccine are not assessed, but it is a good start. Furthermore, it's a conglomeration of many findings and assessments that create the effectiveness of the developed vaccine.

When it comes to continued research for the vaccine for practical application, more studies should be conducted. Future research should focus on the in-silico findings towards preclinical and clinical application for practical use. Practical application can be enhanced by including more proteins associated with breast cancer in the development. Practical application can be compounded by including developed technologies or artificial intelligence because this would yield expedited development processes and reliable epitope predictions. Targeted immunotherapies can change the way we treat cancer.

Reference

Gupta, S., Kapoor, P., Chaudhary, K., Gautam, A., Kumar, R., & Raghava, G. P. S. (2013). In silico approach for predicting the toxicity of peptides and proteins.

Kawai, T., & Akira, S. (2011). Toll-like receptors and their crosstalk with other innate receptors in infection and immunity. Immunity.

Pierce, B. G., Wiehe, K., Hwang, H., Kim, B. H., Vreven, T., & Weng, Z. (2014). ZDOCK server: Interactive docking prediction of protein–protein complexes. Bioinformatics.

Rapin, N., Lund, O., Bernaschi, M., & Castiglione, F. (2010). Computational immunology meets bioinformatics: The use of prediction tools for molecular binding in the simulation of the immune system.

Schwede, T., Kopp, J., Guex, N., & Peitsch, M. C. (2003). SWISS-MODEL: An automated protein homology-modeling server. Nucleic Acids Research.

Vita, R., Mahajan, S., Overton, J. A., Dhanda, S. K., Martini, S., Cantrell, J. R., ... Peters, B. (2019). The Immune Epitope Database (IEDB): 2018 update. Nucleic Acids Research.

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T. A. P., Rempfer, C., Bordoli, L., Lepore, R., & Schwede, T. (2018). SWISS-MODEL: Homology modelling of protein structures and complexes. Nucleic Acids Research.

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.

Comeau, S. R., Gatchell, D. W., Vajda, S., & Camacho, C. J. (2004). ClusPro: A fully automated algorithm for protein–protein docking. Nucleic Acids Research.

Dhanda, S. K., Vir, P., & Raghava, G. P. S. (2019). Designing interferon-gamma-inducing MHC class II binders. Biology Direct.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., &Bairoch, A. (2005). Protein identification and analysis tools on the ExPASy server. In J. M. Walker (Ed.), The Proteomics Protocols Handbook.

Gupta, S., Kapoor, P., Chaudhary, K., Gautam, A., Kumar, R., & Raghava, G. P. S. (2021). In silico prediction of toxicity in peptides and proteins.

Khatoon, N., Pandey, R. K., & Prajapati, V. K. (2017). Exploring Leishmania secretory proteins to design B and T cell multi-epitope subunit vaccine using an immunoinformatic approach. Scientific Reports.

Larsen, M. V., Lundegaard, C., Lamberth, K., Buus, S., Lund, O., & Nielsen, M. (2007). Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. BMC Bioinformatics.

Pandey, R. K., Bhatt, T. K., & Prajapati, V. K. (2018). Novel immunoinformatic approaches to design a multi-epitope subunit vaccine for malaria by investigating Anopheles' salivary protein. Scientific Reports.

Parvizpour, S., Pourseif, M. M., Razmara, J., Rafi, M. A., &Omidi, Y. (2020). Epitope-based vaccine design: A comprehensive overview of bioinformatics approaches. Drug Discovery Today.

Rajendran, H., &Karuppasamy, R. (2023). Design and in-silico validation of a novel MZF-1-based multi-epitope vaccine to combat metastatic triple-negative breast cancer. Vaccines.
Reynisson, B., Alvarez, B., Paul, S., Peters, B., & Nielsen, M. (2020). NetMHCpan-4.1 and NetMHCIIpan-4.0: Improved predictions of MHC antigen presentation. Nucleic Acids Research. Zhou, Y. (2024). Therapeutic vaccines in breast cancer. Highlights in Science, Engineering and Technology.
Harbeck, N., & Gnant, M. (2017). Breast cancer. The Lancet.

Gupta, S., Desai, D., Patel, M., Jyotishi, C., Mahmoud Saleh, A., & Gupta, R. (2025). Screening of oncogenic proteins and development of a multiepitope peptide vaccine targeting AKT1 and PARP1 for breast cancer by integrating reverse vaccinology and immune-informatics approaches. Asian Pacific Journal of Cancer Prevention.

Prawiningrum, A. F., Paramita, R. I., & Panigoro, S. S. (2022). Immunoinformatics approach for epitope-based vaccine design: Key steps for breast cancer vaccine. Diagnostics.

Doytchinova, I. A., & Flower, D. R. (2007). VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinformatics.

Rapin, N., Lund, O., Bernaschi, M., & Castiglione, F. (2010). Computational immunology meets bioinformatics: The use of prediction tools for molecular binding in the simulation of the immune system.

Jespersen, M. C., Peters, B., Nielsen, M., &Marcatili, P. (2017). BepiPred-2.0: improving sequence- based B-cell epitope prediction using conformational epitopes. Nucleic Acids Research.

Goodman, R. E., Ebisawa, M., Ferreira, F., Sampson, H. A., Van Ree, R., Vieths, S., Baumert, J. L., Bohle, B., Lalithambika, S., Wise, J., & Taylor, S. L. (2016). AllergenOnline: A peer-reviewed, curated allergen database to assess novel food proteins for potential cross-reactivity. Molecular Nutrition & Food Research.

Wishart, D., Arndt, D., Pon, A., Sajed, T., Guo, A. C., Djoumbou, Y., Knox, C., Wilson, M., Liang, Y., Grant, J., Liu, Y., Goldansaz, S. A., & Rappaport, S. M. (2014). T3DB: the toxic exposome database. Nucleic Acids Research.

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.

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.