Original Article | Iran J Med Microbiol. 2022; 16(6): 506-519

10.30699/ijmm.16.6.506

Iranian Journal of Medical Microbiology | ISSN:2345-4342

In-silico Designing of Immunogenic Construct Based on Peptide Epitopes Using Immuno-informatics Tools Against Tuberculosis

Saeed Pirmoradi*

Department of Biochemistry, Shahid Chamran University of Ahvaz, Ahvaz, Iran

ABSTRACT

Background and Aim: *Mycobacterium tuberculosis* is a health problem in countries. Despite the global prevalence of tuberculosis and the lack of appropriate drugs, further progress is still needed with the help of modern methods of preparing epitope-based vaccines for tuberculosis.

Materials and Methods: In this study, specific T and B cell epitopes required for producing chimeric vaccines with the help of servers such as IEDB were determined. The antigenicity, allergenicity, and toxicity of the selected epitopes by various other servers such as VaxiJenv2.0, AllerTOP, and Toxinpred were determined, and the vaccine of the epitope was then configured with the help of special linkers. Then, analyze the structure vaccine by some other bioinformatics servers such as PRABI, SWISS-MODEL, PROCHECK, and PEPCALC, was investigated and finally detected using docking techniques to evaluate the interaction with the epitope through MVD software.

Results: The results showed that the vaccine, in terms of in silico evaluations of two-dimensional and three-dimensional structures, has a good condition. Also, the percentage of optimal placement of amino acids and bonds by PROCHECK server, % 99 percent of optimal placement of amino acids in the chimer structure was established. Also, it was non-toxic and nonallergenic and had the desired antigenicity.

Conclusion: In general, the vaccine that was able to have a favorable interaction with some components of the immune system (HLA) in the docking process, which indicates the optimal identification of this structure by the humoral and cellular immune system, of course, more reliable proof of it requires clinical phase processes.

Keywords: Immunoinformatics, Mycobacterium tuberculosis, Molecular docking, Multi-epitope vaccine

Received: 2021/12/24; **Accepted**: 2022/07/11; **Published Online:** 2022/09/09

Corresponding Information: **Saeed Pirmoradi**, *PhD in Biochemistry, Department of Biochemistry, Shahid Chamran University of Ahvaz, Ahvaz, Iran*
Email: pirmoradi150@gmail.com

Use a device to scan and read the article online

pirmoradi S. In-silico designing of immunogenic construct based on peptide epitopes using Immuno-informatics tools against tuberculosis. Iran J Med Microbiol. 2022; 16 (6):506-19.

Download citation: BibTeX | RIS | EndNote | Medlars | ProCite | Reference Manager | RefWorks
Send citation to: Mendeley Zotero RefWorks

1. Introduction

Tuberculosis is caused by the bacterium *Mycobacterium tuberculosis*, an immobile, spore-free, acidic bacillus. This disease is a major global health problem in underdeveloped and developing countries. This bacterium produces a unique compound in its wall called mycolic acid, which plays an important role in the function of the wall. This Wax wall has unique properties such as hydrophobicity, acid fasting, acid resistance, base, dryness, and many antibiotics Give (1). Around one-third of the population has TB germ or is at risk for TB. According to the DALY standard, tuberculosis is the tenth most common disease in the world and is expected to maintain its current position or reach the seventh place by 2020. The bacterium infects about 1.5 billion people and kills 1.3 million annually.

According to the World Health Organization (WHO), in 2018, about 10.0 million people with tuberculosis suffered from this disease. In Iran, according to the Ministry of Health, 14.4 cases of tuberculosis occur annually. Since 1923, the attenuated form of *Mycobacterium bovis* has become the only

established tuberculosis vaccine used worldwide for prevention. This live attenuated vaccine, BCG, has low immunity due to the risks associated with its use in immunocompromised individuals and the possibility of the pathogen returning to malignancy. Various studies have shown this vaccine to have a protective immunity of up to 80% **(2)**. In addition, WHO survey data from around the world show that 0.6 million cases of multidrug-resistant tuberculosis have been reported, of which approximately 0.24 million have occurred **(3, 4)**.

On the other hand, the growth of MDR-TB strains leads to limitations in common tuberculosis treatments. Anti-tuberculosis drugs are expensive and have serious side effects and drug resistance **(5)**. However, it is impossible to cure the disease without creating the right, dominant immunity in the human population against tuberculosis, so safer treatment processes such as vaccine design can greatly help the disease process. Therefore various methods have been used today to develop new and more effective vaccines for tuberculosis, and among these methods, subunit vaccines have shown great promise for treating this disease due to their higher diversity **(6)**. Many researchers worldwide study bioinformatics on immunogenic antigens to design new vaccines **(7)**. An ideal vaccine should detect antigens related to different stages of tuberculosis, including antigens that enter the system in the early stages of infection and enter the latent stage of infection (delay) **(8)**. Therefore, in this research process, we intend to use a new immunoinformatics method to identify peptide fragments obtained in tuberculosis clinical trials in the IEDB server, to create a new chimeric peptide construct of the vaccine that has the desired ability to activate and create immunity in the host so that we can use it in clinical evaluations.

2. Materials and Methods

Preparation of Suitable Epitopes

We used the experimental epitope method based on peptide epitopes to perform the vaccine design process. This method's error rate is greatly reduced compared to other methods because we use screening for peptides that have already been laboratory-approved and are in the IEDB database (IEDB; http://www.iedb.org/). In this database, by identifying the relevant elements, selecting the desired microorganism, and submitting them, we access many epitopes based on cell B and, MHCI, MHCII. We will pay more ASSAY, reference them, and select some of them as examples.

Then, during screening, each of the selected epitopes was examined for antigenicity with the help of the VAXIJEN server. Also, in the continuation, they were checked by the ALLERTOP server for allergens in the host body. In the next step, we have to examine the studied epitopes for the degree of toxicity. For this, we used the TOXINPRED server. Now, we select the epitopes related to T and B, which have no problem in all three cases.

Construction of a Multiepitope Vaccine

In the next step, we have to create the chimer protein P from the epitopes we screened. To do this, we put the B and T epitopes together and connect them with compounds known as linkers and called KKs. KKs are flexible linkers. After completing the sequences, we add a cholera B subunit adjuvant (CTxB) to the two ends of the chimeric sequence generated from the selected epitopes using the PAPAP linker to increase the efficiency of the immune structure in creating immunity in the host body. The NCBI database was used to prepare the desired adjuvant sequence.

2D and 3D Structure Investigation of Chimer Vaccine

We should examine our vaccine for alpha-helix, beta-sheets, and random coil for the second protein structure in the first step. In these cases, we must examine the secondary structure of the protein through the PRABI server and the GORIV method. In the next step, we examine its three-dimensional structure because a peptide vaccine must have a stable three-dimensional structure to function well. We used two Swiss models and I-Tasser servers at this stage.

Model Validation and Refinement

If our selected model needs to be refined from another server called 3Drefine, we will perform the refining process on it to obtain refined models that are closer to the natural structure than their previous structure. Now, the obtained model should be examined in terms of bands and amino acids, which have been spatially examined in our proteins, using the PROCHECK server and the Ramachandran diagram.

Physicochemical and instability evaluation of model vaccines

Next, the obtained protein model should be examined in terms of physicochemical properties to create a better-quality model. Therefore, for this type of study, PEPCALC, and protParam servers to evaluate the percentage and load of amino acids, molecular weight and polarity, extinction coefficient, half-life, water instability, and solubility index, aliphatic index and other indicators to evaluate the efficiency, durability and power, Favorable bonds with receptors

and the ability of the vaccine to move in body fluids were used.

B cell Epitope Prediction

At this stage, epitopes related to B cells were predicted using BIPRED-2.0 and LBTOPE, IEDB servers

T cell Epitopes Prediction.

Using two different online tools, ProPred-1 and MHCH2PRED, CTL epitopes of MHC class I and MHC class II target protein, respectively, were predicted. The results of these tools are very important because they use a large number of HLA alleles (human-leukocyte antigens) during the calculation. The sequence was presented in a simple format and all alleles were selected for prediction. For proteasome and immunoproteasome filters, ProPred-1 was kept with a threshold value of 5%.

Molecular Docking

The molecular binding process was performed by special software such as the Molegro Virtual Docker to predict the affinity between the vaccine structure and some MHCIs, human MHCIIs with access codes to the HLA-DRB1-0101 (PDB=5ni9) and A0201 (PDB = 4uq3) databases. Molecular to predict affinity between vaccine structure and some human MHCI, MHCII was performed with access code from RCSB database and after minimizing energy and recipient and vaccine structural modifications

Normal State Analysis for Structural Stability Analysis of Vaccine/receptor Complex

Normal analysis was performed to analyze the structural stability of the selected docking complex via the iMODS web server. The server shows NMA mobility with arrows indicating motion dynamics. Eigenvalues indicate the strength of structures and deformation diagrams of non-reinforced sections of the structure. The server also provides factor B, variance, covariance map and link matrix of the desired structure.

Immunological Simulation

In silico safety, simulation was performed using a C-ImmSim server to confirm immunization and immune response against the selected vaccine. This server uses a machine learning base to predict epitopes and related security interactions. The server automatically simulates three anatomical sections: **(1)** bone, where hematopoietic stem cells are stimulated, and myeloid cells are produced **(2)** lymphatic organs, and **(3)** Thymus, in which simple T cells are selected to prevent autoimmunity

Codon optimization

The candidate vaccine codon was optimized for maximum vaccine expression in the host. In this process, reverse translation and codon optimization were performed to predict the expression of the vaccinated model in an expression vector using the Java Codon Compatibility Tool (JCat) to increase the vaccine translation rate in the prokaryotic E. coli K12 system. The server provides a codon matching index (CAI) that indicates codon usage bias and GC content. CAI values between 0.8 and 1 are adaptable, and GC content between 30% to 70% is useful for translation activities.

3. Results

Results of antigenicity, Allergenicity and Toxicity of Selected Epitopes

In the first step, after using the IEDB server to obtain the antigens obtained in experimental and laboratory processes for the pathogen Mycobacterium tuberculosis, each of these peptide epitopes was examined by VAXIJEN server to determine the antigenicity. We gave and isolated fifteen of them that had more desirable antigenicity **(9)**. After examining the allergenicity of the selected epitopes, it is time to examine their antigenicity, which we have isolated as having no allergenic properties for further studies **(10, 11)**. At this stage, we have to evaluate each epitope in terms of toxicity in the two previous stages of antigenicity and allergy. Also, with the Toxinpred server, three stages of evaluation of a large number of epitopes have resulted. Fifteen produced chimeric proteins, the primary composition of our vaccine, were isolated **(12)**.

Results of Configuration and Design of the Initial Structure of the Chimer Vaccine Protein

In this study, we obtained several peptide sequences of different sizes from the IEDB database as candidates for producing chimeric vaccine proteins. End of N Peptide Sequences Using the Hard PAPAP Linker. Rigid linkers function to prohibit unwanted interactions between the discrete domains. We added the cholera B-subunit adjuvant sequence (CTxB) obtained from the NCBI database to form the initial protein structure of our Chimer vaccine. We tried to put the peptides related to B and T cells together and bind them together (Figure 1).

Figure 1. Structure of chimeric structure vaccine (red=adjuant, Purple=rigid linker,blue=flexible linker,green=epitopes)

At this stage of the vaccine design, we need to bind the peptide epitopes that we isolated in the previous steps of the analysis to form the sequence of the original structure of our complete protein. Therefore, in this step, we first put together the epitopes associated with B and then the epitopes associated with T, and between them, we used lysine/lysine linkers to connect. Through the NCBI database, the required adjuvant sequence (cholera toxin B subunit) and with the help of the PAPAP hard linker, we connected to the N end and C end of our chimeric protein sequence (of course, in obtaining the adjuvant sequence, care must be taken that it is not partial samples), thus doing the basic structure of protein sequences. Our chimer was created for further vaccine design studies.

Result Codon optimization.

Nucleotide sequences for VACCINE were obtained after reverse translation and codon optimization on the JCat server. The CAI values of the optimized sequences were 1, and the content CG was 48.247%, which indicated the proper state of the vaccine sequence after optimization in the prokaryotic host E. coli for the cloning process.

Results of Evaluation of the 2D and 3D Structure of the Chimer Vaccine

At this stage of evaluating the composition of the produced chimera in terms of two-dimensional and three-dimensional structure, we examined that if undesirable folds are created in its structure, we should remove the disruptive epitope parts of the chimera structure from its original structure at the very beginning. With the help of Prabi server and using the GORIV method, we examined the second structure of the chimera, and the results of this study were as follows (Tabale 1).

Table 1. Evaluation of two-dimensional structure of the primary chimeric protein of vaccine by prabi server by GORIV method

GOR4	
Alpha helix(Hh)	208 is 39.77%
3_{10} helix(Gg)	0 is 0.00%
Pi helix(Ii)	0 is 0.00%
Beta bridge(Bb)	0 is 0.00%
Extended strand(Ee)	95 is 18.16%
Beta turn(Tt)	0 is 0.00%
Bend region(Ss)	0 is 0.00%
Random coil(Cc)	220 is 42.07%
Ambiguous states(?)	0 is 0.00%
Other states	0 is 0.00%

Then we have to get the 3D structure, and because there is no structure from our chimer composition, we used the modeling homology method using I-TASSER servers and entered the chimer protein sequence in this server as a template. Moreover, we found a combination with a peptide sequence similar to our chimeric protein. By examining several indicators in these models, such as sequence identity, Local Quality Estimate, z-score, and comparison, the similarity of the model to the natural compounds in this server Gives the seq comparison index with the number 83.33% gives us an acceptable.

Also in the image below, Local Quality Estimate, z-score, comparison indicators show the similarity of the three-dimensional model combination (model_02) provided by the SWISS-MODEL server based on the chimeric protein sequence pattern of our designed vaccine **(13, 14)**. If during the review of the desired 3D model, it is determined that the relevant model is not of good quality in terms of the mentioned indicators, we must modify it to get (Figure 2).

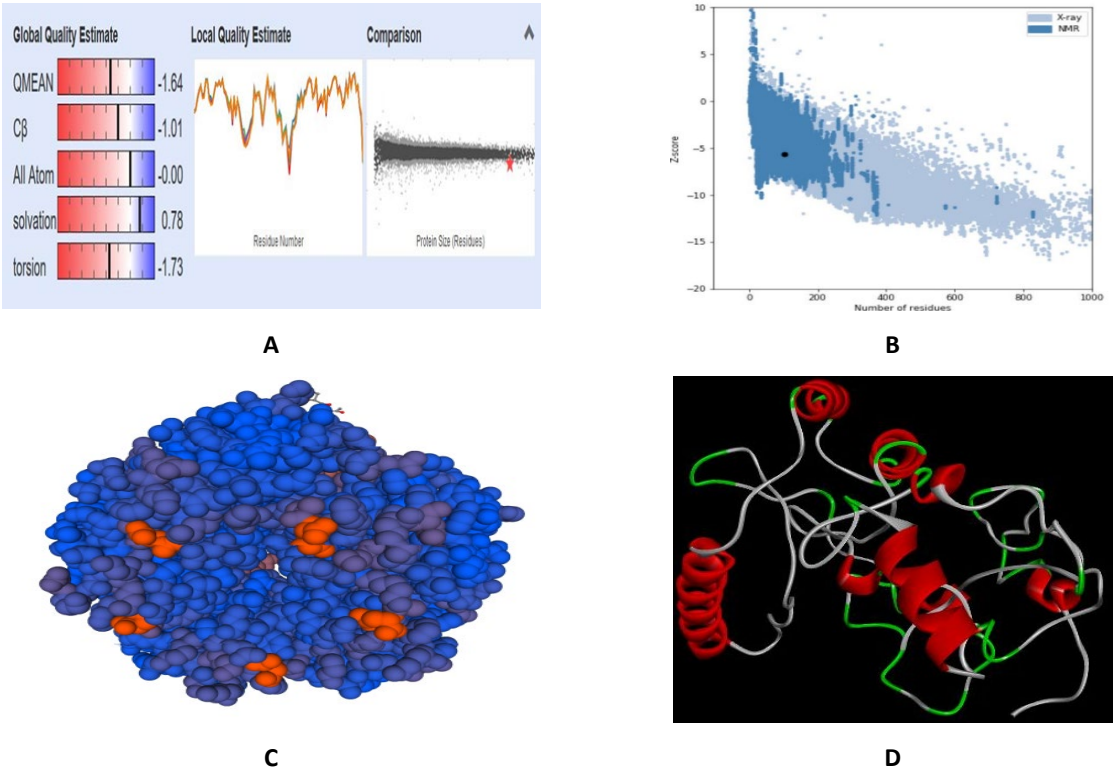


Figure 2. A: Image obtained from evaluating Local Quality Estimate, z-score, and comparison indicators. The third structure of the model vaccine with the help of SWISS-MODEL server B; The image is the result of evaluating the three-dimensional structure of the vaccine modeled by the Prosa server, which shows a Z score of -5.65.C: Three-dimensional shape of the SWISS model vaccine. D: Three-dimensional shape of the I-TASSER model vaccine.

In the continuation of the examination process, we should examine the model we obtained during the modeling homology process in terms of placing amino acids in the desired and high-risk position of our vaccine model, which we do with the Procheck web server or I-TASSER by examining the diagram. We performed Ramachandran and found that about 96% of the amino acids in the vaccine model we studied are in the desired regions (Figure 3) **(15)**.

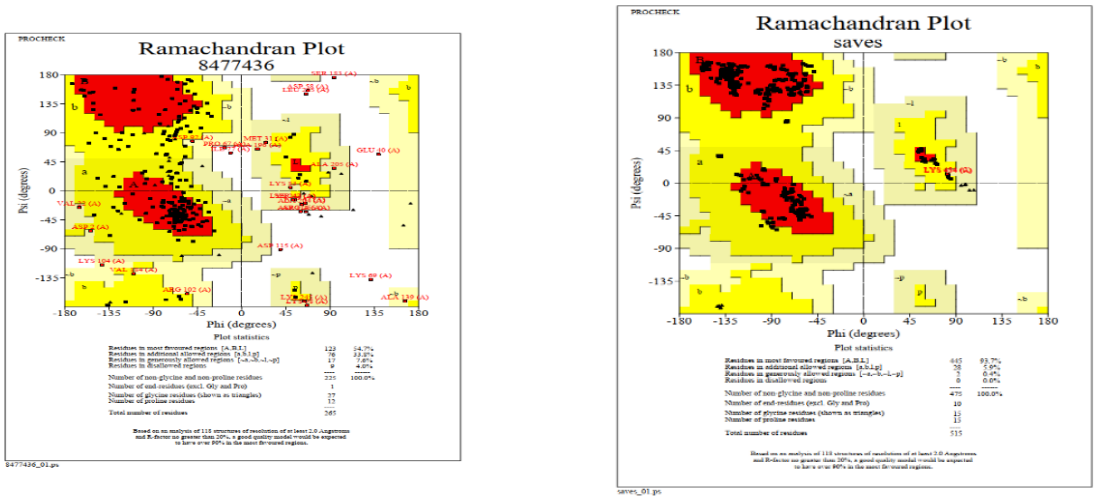


Figure 3. The image obtained from Ramachandran data shows the number of amino acid residues in the desired and almost desirable region with the help of PROcheck server, which includes 96- 99%. (A=I-TASSE, B=SWISS MODEL)

Results of Physical and Chemical Properties

At this stage of the evaluation of the vaccine, we evaluated many physical and chemical properties in it by servers such as Protparam and Pepcalc, which resulted in a combination of the vaccine with a stability coefficient of 39.88 was a stable combination in the host body and half It has a suitable life and 30 hours in the body of mammals. Also, the composition of the case has a molecular weight of 56100.14, and gravity of -0.278, and an aliphatic coefficient of 80.67. It is also in good condition in terms of hydrophobicity and extinction coefficient. PHI = 9.75 and, according to the results of the Pepcalc server, had good solubility in water, which allows us to enter it in the laboratory phase processes (Table 2) **(16, 17)**.

Table 2. Table from evaluation of physicochemical results of vaccines modeled by Pepcalc server

Physiochemical properties	
Number of residues	513
Molecular weight	56100.14g/mol
Extinction coefficient	33710$m^{-1}cm^{-1}$
Iso-electric point	9.75
Net charge at pH7	16.6
Estimated solubility	Good water solubility

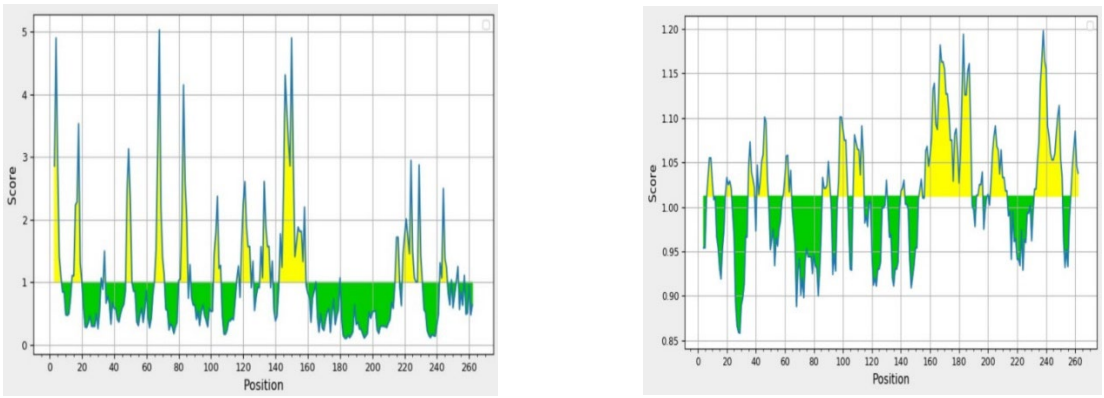
Results of Predicting Interaction with B cells

The interaction of epitopes with immune B cells was done with the help of BIPRED-2.0 server, which is one of the best methods for predicting B cell epitopes because it shows significant results based on the IEDB database and crystallography-derived epitopes **(18, 19)**. However, the other server is LBTOPE, which is obtained based on experimentally synthesized epitopes in laboratories in the IEDB database and has selective parameters such as 15 amino acid length and specific scoring threshold in this server in Table 3 **(6, 20)**.

Table 3. Results from the evaluation of predicting the interaction of epitopes with immune B cells using the LBTOPE server

Original peptide						
Peptide	**% probability**	**Hydrophobicity**	**Hydropathicity**	**Hydrophilicity**	**charge**	**Mol wt**
ADEEQQQALSSQMGF	57.36	-0.21	-0.96	0.25	-3.00	1668.97
EKASVPGGGDMGGMDF	83.34	-0.05	-0.39	0.33	-2.00	1554.94
LDVAEGDTVIYSKYG	83.32	-0.04	-0.05	0.01	-2.00	1630.00
MFSGFDPWLPSLGNP	47.57	-0.08	-0.00	-0.63	-1.00	1665.11
DNIGNANIGFGN	33.12	-0.04	-0.43	-0.23	-1.00	1205.43
AERAPVEADAGGGQKVLVRN	63.32	-0.22	-0.48	0.50	0.00	2037.55
AGADVVGSDDLIE	33.23	0.01	-0.36	0.36	-4.00	1260.50
AEMKTDAATL	27.23	-0.13	-0.12	0.36	-1.00	1050.32
AEMKTDAATLA	21.91	-0.09	-0.05	0.28	-1.00	1121.41
RADEEQQQAL	35.63	-0.48	-0.81	0.98	-2.00	1187.37
DQVHFQPLPPAVVKLSDALI	90.55	0.01	0.41	-0.30	-0.50	2187.87
VTGSVVCTTAAGNVNIAIGG	66.02	0.15	1.11	-0.63	0.00	1804.35
LFAALPSFAGLRPTFDTRLM	50.32	0.01	0.58	-0.49	1.00	2258.94
DDIKATYDKGILTVSVAVSE	32.43	-0.08	0.12	0.23	-2.00	2124.66
LQVPSPSMGRDIKVQFQSGG	69.00	-.14	-0.39	0.00	1.00	2131.74

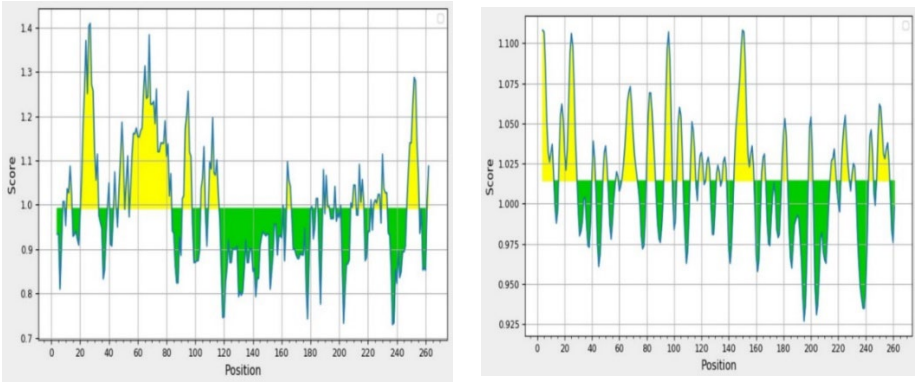
At this stage, amino acid-based methods were used to identify potential predictive B cell epitopes. Predicting the interaction of epitopes with immune B cells was done with the help of IEDB server, which is one of the best methods for predicting B cell epitopes **(21, 22)**. During this stage, different methods of IEDB server analysis were used to predict the B-cell epitope. In Kolaskar and Tongaonkar32 antigen prediction methods, antigens were examined based on physiological characteristics of amino acids and abundance in known experimental epitopes. The average protein antigen orientation is 1.10, maximum 1.20, and minimum 1.05. The antigen threshold for protein was also 1.02. All values greater than 1.00 determined their antigenic potential. Therefore, it was concluded that some of them have a predetermined threshold value and can express the B cell response. For an epitope to be a powerful stimulus for B cells, it must have a good level of access. The prediction of access to the level of Emini was used to detect these features. It is shown in Figure 4.



A B

Figure 4. A: Prediction of Kolashkar and Tongaonkar antigens from immunogenic constructs (Notes: X-axis and Y-axis show antigen sequence position and orientation, respectively. Threshold value is 1.02. Areas are higher than antigen threshold. Shown in yellow.). B: Predicting the access to the Emini surface of the immunogenic structure (Notes: The X-axis and the Y-axis indicate the sequence position and the surface probability, respectively. The threshold value is 1,000. The regions above the antigen threshold are shown in yellow.

Areas containing beta-screws are usually available and naturally hydrophobic. These two properties are antigenic regions of a protein that can be investigated based on the Chou and Fasman evaluation model and by examining beta-screws **(23)**. Areas 10-25, 60-70, 80-85, 90-97, 140-160 and 245-265 were considered as a beta-rotation region compared to other sections. Empirical evidence has also shown that the flexibility of a peptide is related to its antigenicity. Therefore, the Karplus and Schulz flexibility prediction method was formed on this basis and used to study the flexibility of immunogenic structural epitopes. This forecasting method found that areas 20-33, 45-83, 90-95, 105-120, and 205-215 are the most flexible parts of the immunogenic structure under study (Figure 5).



A B

Figure 5. A: Prediction of flexibility by the Karplus and Schulz method of immunogenic antigenic structures (flexible regions are yellow. The X-axis and the Y-axis indicate position and score, respectively. The threshold is 21.0). B: Chou and Fasman's method predict the beta screw of safety structure (X-axis and Y-axis show position and score, respectively. The threshold is 0.95. Areas with beta screw are yellow, above the threshold value shown.

Finally, the Bepipred linear epitope prediction tool was used, based on a Hidden Markov model, the only method for predicting linear B-cell epitopes. Based on all the data, it was predicted that the immunogenic peptide sequences of amino acids 0-30, 60-100, 120-155, and 245-255 could induce the desired immune response as B cell epitopes (Figure 6).

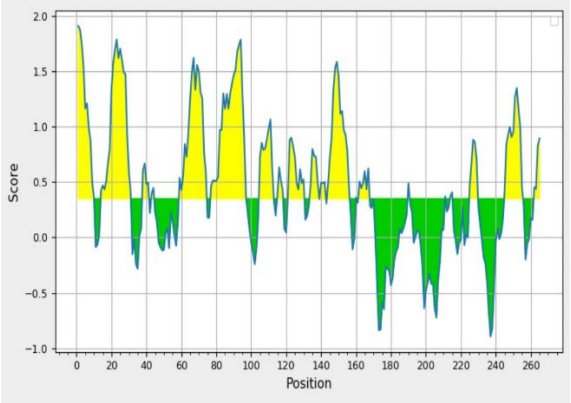


Figure 6. Bepipred linear epitope prediction in the immunogenic structure (Notes: x-axis and y-axis show position and score, respectively. The threshold is 0.50. Areas with beta screws are shown in yellow. The highest peak region represents the most powerful B cell epitope.

Results of predicting interaction prediction with MHCI, MHCII

At this stage of the work, Propred-I and MCH2PRED servers were used to predict epitopes by T cells. ProPred-I used a basic matrix method to scan and predict peptides against the library of MHCI class alleles and MCH2PRED for MHCII class alleles. The sequence of peptides in FASTA format was performed on these servers. At the same time, all alleles were selected with a higher score peptide with a threshold of 5% and proteasome filtration and the immune proteasome **(21, 23, 24)**. In addition, screening for immunogenicity, toxicity, and antigenicity of epitopes was performed, and a number of selected epitopes were selected for subsequent processing based on their antigen score. Then we selected some of them that had a good interaction with MHCI, MHCII, and B and classified them in a separate table because they were the epitopes that play the most important role in stimulating and creating immunity in the body. The hosts played a role due to their vaccine (Table 4).

Table 4. Results of prediction on the interaction of epitopes used in vaccine production with MHCII, MHCI, B cells using Propred-I, MHC2pred, and bcpred server

VAXIJEN SCORE	ALLERGENICITY	TOXICITY	B cell epitopes	MHCII HLAS	MHCI HLAS
0.8533	NON ALLERGIC	NON TOXIC	1AGADVVGSDDLIE13	HLA-B*2702, HLA-B*2705 HLA-B*3501,HLA-B*3701,HLA-B*3902,HLA-B40,HLA-B*4403,HLA-B*5201,HLA-B*5801,HLA-B60,HLA-B61,HLA-B62,HLA-Cw*0301	HLA-DR9, HLA-DR13, HLA-DR53, HLA-DQ4, HLA-DQ7, HLA-DQ8,HLA-DQA1*0501, HLA-DQB1*02, HLA-DRB4*0101, HLA-DRB1*0405, HLA-DRB1*0802 , HLA-DRB1*0901, HLA-DRB1*1101,HLA-DRB1*1302, RT1.B
1.4873	NON ALLERGIC	NON TOXIC	LQVPSPSMGRDIKVQFQSGG20	HLA-A*1101,HLA-A3, HLA-A*3101,HLA-A*3302,HLA-A68.1, HLA-A20 Cattle,HLA-B*2702,HLA-B*3501, HLA-B*3701,HLA-B40, HLA-B*4403,HLA-B*5301,HLA-B*51,HLA-B*5801,HLA-B62,HLA-B*0702,	HLA-DR9, HLA-DR2, HLA-DR3, I-Ag7,HLA-DR13, HLA-DR15,HLA-DR51, HLA-DR52,HLA-DR53, HLA-DQ4, HLA-DQ6,HLA-DQ7, HLA-DQ8, HLA-DQB1*0301,HLA-DRB1*0101,HLA-DRB4*0101,HLA-DRB1*0404, HLA-DRB1*0405,HLA-DRB1*0901, HLA-DRB1*1101, HLA-DRB1*1302 , HLA-DRB1*1501,
0.6500	NON ALLERGIC	NON TOXIC	1LDVAEGDTVIYSKYG15	HLA-A1,HLA-A*3302, HLA-A68.1,HLA-B*2702, HLA-B*3501,HLA-B*3701,HLA-B40, HLA-B*4403,HLA-B*5101, HLA-B*5103,HLA-B*5201,HLA-B*51,HLA-B*5801,HLA-B61,	HLA-DR4, HLA-DR9, HLA-DR2, HLA-DR15, HLA-DR52,HLA-DQ4, HLA-DQ7, HLA-DQ8, HLA-DQB1*0301,HLA-DRB1*0405, HLA-DRB1*0901
0.4903	NON ALLERGIC	NON TOXIC	1DDIKATYDKGILTVSVAVSE20	HLA-A2, HLA-A*0201, HLA-A*0205,HLA-A24, HLA-A20 Cattle,HLA-A2.1,HLA-B*2705,HLA-B*3501,HLA-B*3701,HLA-B*3902, HLA-B40,HLA-B*4403, HLA-B*5101,HLA-B*5102,HLA-B*5103, HLA-B*5201,HLA-B*5301,HLA-B*5401, HLA-B*51,HLA-B*5801, HLA-B60,HLA-B61,HLA-B7,HLA-B*0702, HLA-Cw*0401,	HLA-DR4, HLA-DR9, HLA-DR2, HLA-DR13, HLA-DR52,HLA-DQ4, HLA-DQ6, HLA-DQ7, HLA-DQ8, HLA-DQB1*0301, HLA-DRB1*0101,HLA-DRB5*0101, HLA-DRB1*0301, HLA-DRB1*0401,HLA-DRB1*0405, HLA-DRB1*0802, HLA-DRB1*0901, ,HLA-DRB1*1101, HLA-DRB1*1302,

VAXIJEN SCORE	ALLERGENICITY	TOXICITY	B cell epitopes	MHCII HLAS	MHCI HLAS
0.9820	NON ALLERGIC	NON TOXIC	VTGSVVCTTAAGNVNIAIGG	HLA-B*3701,HLA-B40, HLA-B*5101,HLA-B*5102,HLA-B*5103, HLA-B*5301,HLA-B*51, HLA-B*5801,HLA-B61,	HLA-DR9, HLA-DR2, HLA-DR3, HLA-DR13, HLA-DQ4,HLA-DQ7, HLA-DQ8, HLA-DQB1*03, HLA-DQB1*0301,HLA-DRB1*0101, HLA-DRB5*0101, HLA-DRB1*0401, HLA-DRB1*0405,HLA-DRB1*0901, HLA-DRB1*1101, HLA-DRB1*1302, I-Ad, I-Ag7
0.7500	NON ALLERGIC	NON TOXIC	DQVHFQPLPPAVVKLSDALI	HLA-A1, HLA-A2, HLA-A*0201, HLA-A*0205, HLA-A*1101, HLA-A24, HLA-A3, HLA-A68.1, HLA-A2.1, HLA-B*2702, HLA-B*2705, HLA-B*3501, HLA-B*3701, HLA-B*3801, HLA-B*3901, HLA-B*3902, HLA-B*5101, HLA-B*5201, HLA-B*5301, HLA-B*5401, HLA-B*51, ,HLA-B62, HLA-B7, HLA-B*0702, HLA-B8, HLA-Cw*0301, HLA-Cw*0401, HLA-Cw*0602, HLA-Cw*0702,	HLA-DR4, HLA-DR9, HLA-DR7, HLA-DR13, HLA-DR15,HLA-DR52, HLA-DR53, HLA-DQ4, HLA-DQ6,HLA-DQ7, HLA-DQ8, HLA-DQB1*03,HLA-DQB1*0301, HLA-DRB1*0101, HLA-DRB4*0101 ,HLA-DRB5*0101, HLA-DRB1*0401, HLA-DRB1*0404, HLA-DRB1*0405, HLA-DRB1*0802, HLA-DRB1*0901 ,HLA-DRB1*1302, HLA-DRB1*1501
1.3524	NON ALLERGIC	NON TOXIC	AERAPVEADAGGGQKVLVRN20	HLA-A1, HLA-A*3302, HLA-A68.1, HLA-B*2702, HLA-B*3501, HLA-B*3701, HLA-B*3801, HLA-B*3902, HLA-B40, HLA-B*4403, HLA-B*5101, HLA-B*5102, HLA-B*5103, HLA-B60, HLA-B61, HLA-B7, HLA-B8,	HLA-DR9, HLA-DR2, HLA-DR13, HLA-DR15, HLA-DR52, HLA-DQ7, HLA-DQ8, HLA-DRB1*0101, HLA-DRB4*0101, HLA-DRB1*0401, HLA-DRB1*0405,HLA-DRB1*0901,
1.2080	NON ALLERGIC	NON TOXIC	1AGADVVGSDDLIE13	HLA-A*3302 ,HLA-A68.1 ,HLA-B*3701 ,HLA-B40 ,HLA-B*4403 ,HLA-B*5101 ,HLA-B*5102,HLA-B60 ,HLA-B61 ,HLA-Cw*0301 ,	HLA-DR9, HLA-DR2, HLA-DQ4, HLA-DQ6, HLA-DQ7, HLA-DQ8, HLA-DQB1*0301, HLA-DRB1*0405, HLA-DRB1*0901,

Molecular Docking Results

At this stage, after examining the connectivity of the model vaccine designed with MHCI, MHCII, and B, we must now show the interaction of the epitopes in the structure of the vaccine produced by performing the molecular docking process using Molegro virtual docker software. Imagine this Identification of epitopes with MHCI and MHCII molecules after evaluation and optimization. The docking process's main goal was to create a stable complex by creating the correct orientations between the vaccine in question and the MHCI and MHCII compounds. We prepare PDB data. After reviewing some articles, we found two combinations of HLA as the best type of HLA, which were named HLA-A-0201 and HLA-DRB1-01: 01 for MHCI, and MHCII, respectively, then the PDB file. We prepared them for docking. While docking the PBD file of MHCI protein, MHCII is separately entered into MOLEGRO VIRTUAL DOCKER software and used. We remove its ligands. Then we enter the vaccine ligands we designed in the previous step into the software. Now, before the docking process on the protein, we have made the necessary preparations by correcting the MHCI and MHCII protein errors separately for each of them so that in the molecular preparation stage, we first select a state from which the bands and orientations of hydrogen bonds and charges Flexible bonds are also identified in the ligand. Then, in the protein preparation step, we identify and eliminate the errors of its residues and optimize that their replacement residues are in the best orientation with the surrounding residues in the protein structure. Next, we need to find the holes in the protein structure to determine which parts of the protein are most likely to interact between the ligand and the protein, and by doing so, we make pre-docking preparations. Then select the ligands, then set the SCORING mode to MolDock scoring (GRID) mode, and select the connection location based on the holes. Then select the algorithm based on Mol Dock SE, select the number of thighs and the POSES storage format, and start the docking process. From the designed vaccine ligands, the ligands that had the highest negative energy were selected and used to investigate the interaction with MHCI and MHCII proteins, and the amino acids involved in this interaction were identified (Figure 7).

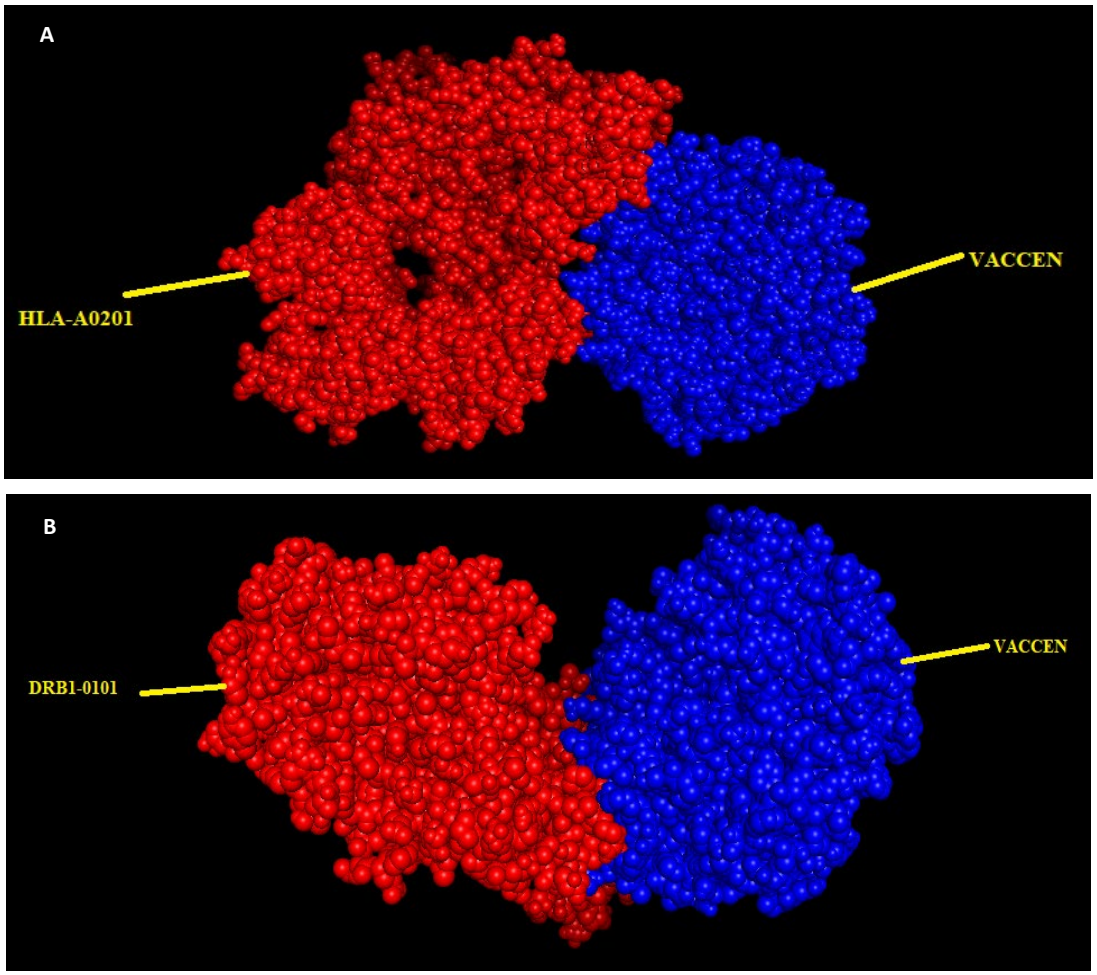


Figure 7. Results of molecular docking between vaccines modeled with MHCI (A) and MHCII (B) proteins

Results of Normal State Analysis for Structural Stability of Vaccine/receptor Complex

Then, to analyze the biophysical stability and changes of the vaccine-receptor complex, molecular dynamic simulations were performed through the iMODs server. The resulting iMOD original chain deformation is shown in Figure 8A. The area where the hinges are located has a strong tendency to deform. The values of factor B calculated by normal state analysis are proportional to the square root of the mean. Factor B shows the atomic position fluctuations and measures each atom's unpredictability, as shown in Figure 8B.

Figure 8C shows the eigenvalues closely correlated with the energy required to smooth the structure, and in general, this figure reflects the stability of the dock complex by the eigenvalue shown. The eigenvalue of the vaccine-receptor complex is 2.569466e-05, which indicates less energy required for the deformation of the structure and the complex's stability.

Activates immune cascades to remove antigens. The variance is inversely related to the eigenvalue. In Figure 8D, the individual variance with red and green represents cumulative variance. The covariance map shows the relationship between the pairs of residues because the colors red, blue, and white indicate the correlated, anti-correlated, and unrelated pairs of residues shown in Figure 8E. The elastic lattice diagram shows the pair of atoms attached to the spring (springs), and each point on the diagram represents a spring's spring between the corresponding pair of atoms. In the diagram, the darker gray color indicates the stiffer springs shown in Figure 8F

Immunological Simulation Results

The results provided by the C-ImmSim server were consistent with the actual safety responses, as evidenced by the increased production of secondary responses. The initial response to the vaccine was demonstrated by increasing high IgM levels. In the secondary and tertiary responses, the increase in the B cell population as an increase in IgG1+IgG2, IgM, and IgG+IgM levels is shown in Figure 7. In addition, we found an increase in helper T cell populations and cytotoxic T cell populations with memory cell proliferation (Figure 7D). We also noticed higher levels of cytokines such as IFN-γ, IL-10, TGF-B, and IL-12, as shown. These observations showed that the designed vaccine elicited promising anti-Quid-19 immunogenic reactions (Figure 9).

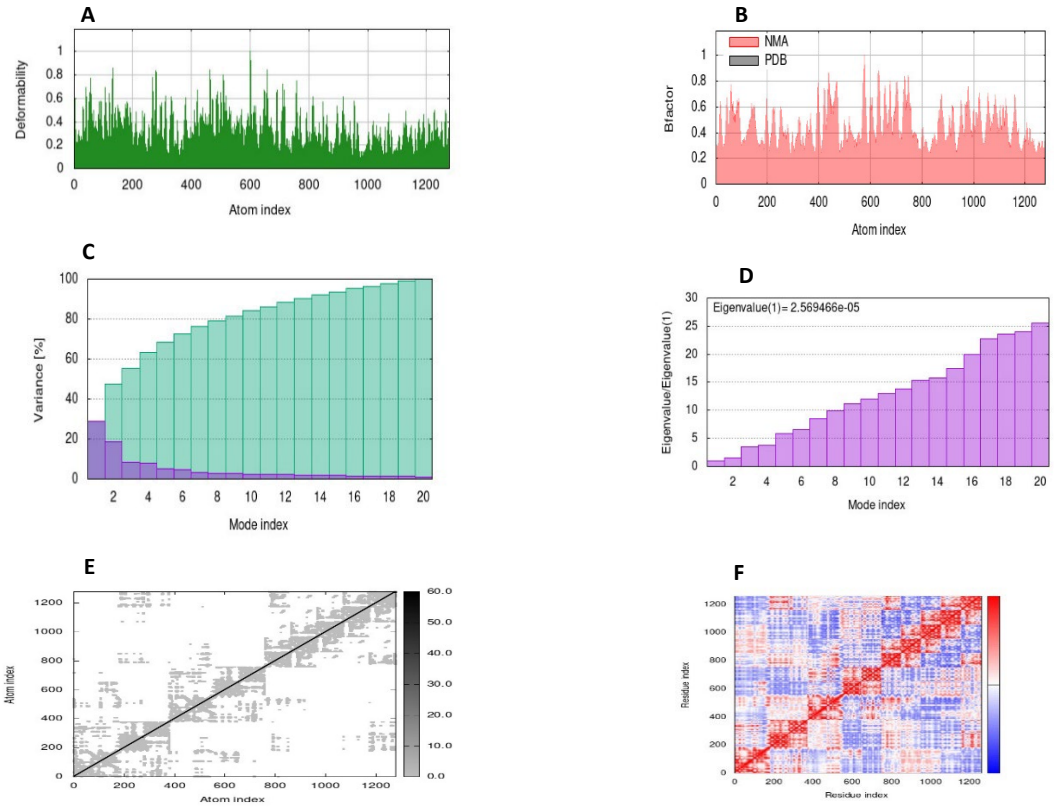


Figure 8. Molecular dynamics simulation of the VACCEN complex. The stability of the protein-protein complex was investigated through deformation (A), factor B (B), eigenvalue (C), variance (D), covariance matrix (E) and elastic lattice (F).

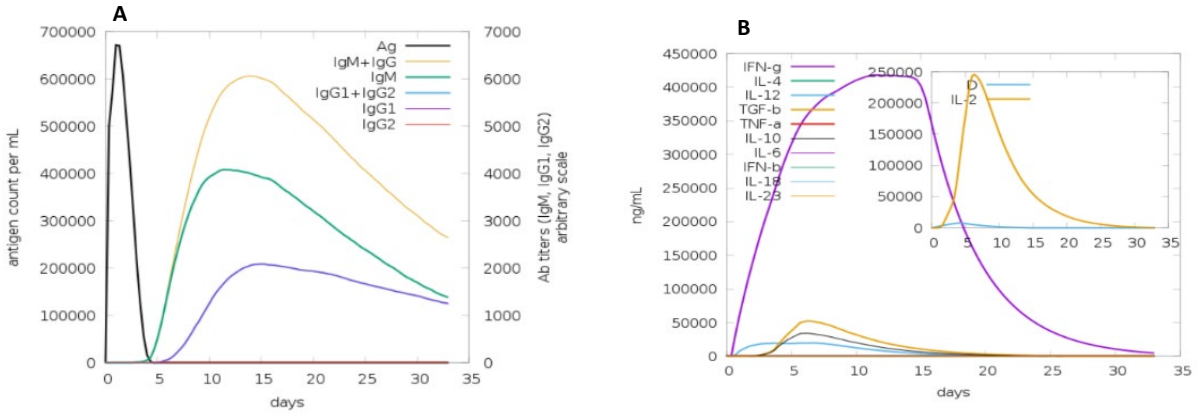


Figure 9. Immunological simulation results provided by the C-ImmSim server

4. Discussion

Tuberculosis (tuberculosis) is a serious disease that varies in mortality and various variables in infected people and is estimated at two million people a year. If left untreated, the mortality rate can increase by 60% **(22, 25, 26)**. This process mainly endangers Third World people due to poverty, low health standards, and inadequate medical care **(27-29)**. The most important vaccine used in most human societies is the BCG vaccine, which is a live and weakened structure that has weaknesses against this disease in children and adults **(30)**. Therefore, new vaccine structures obtained with the help of new methods are very much needed today to deal with this disease. Vaccine design has become more accurate and efficient, based on new technologies and access to data and genetic and protein databases. In addition, the use of bioinformatics tools is much more beneficial than the old methods of vaccine production and design, both in terms of efficiency, time, and economy **(31)**. Therefore, today the development of new specific

vaccines to prevent tuberculosis is becoming a top priority for researchers in the field of tuberculosis. Preparing the sequence of antigens capable of stimulating the immune system is an important step in vaccine design, and this process can be easily done by insilico processes **(32)**. n this work, a series of linear epitopes of B and T cells were predicted with the help of bioinformatics servers, which can stimulate different parts of cellular and humoral immunity **(33)**. Prediction and preparation of recognizable epitopes by B cells is an important factor in the design of vaccines, which determines the location of antigen-antibody interactions **(19, 34, 35)**. In this article, these epitopes were predicted through different servers. Also, T cell epitopes are essential for stimulating adaptive immunity through the presentation by MHC molecules **(36)**. Therefore, selecting epitopes that can interact with MHCs is an essential aspect of T cell stimulation (37), and detecting vaccines designed by CD4+ and CD8+ T cells is vital in creating a more favorable immunity in the host body **(38, 39)**. After predicting every epitope of B and T cells through the sequence of antigens, we connected them with the help of kk and PAPAP linkers to create the main and overall structure of the intended vaccine. Further investigations should be prepared **(40)**. PAPAP linkers were used as interfaces between adjuvant sequences and vaccine epitopes to better express vaccine efficacy.

Also, the lack of allergenic properties, the lack of toxicity, and the favorable antigenicity score in the epitopes of the designed vaccine have increased its efficiency as a desirable vaccine **(40)**. The mentioned vaccine needs a suitable host for proper expression and the most suitable host for expressing such recombinant proteins in E. coli expression systems **(41, 42)**. Next, codon optimization was carried out to prepare and obtain a high expression level of the recombinant vaccine in the E.coli system (K12 strain). That is, having a good GC content (48.277%) and CAI score (1.0) of the investigated vaccine sequences was desirable for this work. After this process, immunological simulation analyses were performed on it to confirm these cloning results, and favorable results of the role in stimulating various immunological parts were observed, which was done by some immunoinformatics servers. In general, in this study, various software tools were used to evaluate the different properties of the vaccine. In addition, the evaluation of the epitope-based subunit vaccine predicted by the immunological simulation process and biophysical stability analysis by the iMODs and C-ImmSim servers was highly acceptable. These results confirm the combination of the designed vaccine as a potential immunization vaccine against tuberculosis. However, final and complementary confirmation of all these potential immunoinformatics-based results requires laboratory and clinical processes.

5. Conclusion

In this article, a series of new methods based on immunoinformatics have been adopted to identify peptides derived from antigens of some structures exposed to the surface instead of focusing on the entire pathogen, which is a more efficient solution. This method can reduce costs in terms of economy and time. In this method, only parts of the antigenic regions of the immunogenicity stimulus of Mycobacterium tuberculosis protein were screened to prepare the epitope for possible vaccine development. Finally, after the design of the vaccine structure, evaluation of structural bioinformatics, investigation of the clone, immunogenicity, and investigation of allergenicity, the toxicity of the results indicated the confirmation of its favorable conditions. However, the final confirmation of its efficiency will require use in laboratory and clinical processes and obtaining confirmation through them.

Acknowledgment

The authors would like to thank all the people who helped in the various stages of this research.

Funding

This research did not receive any specific grant from any agency in the public, commercial, or not-for-profit sector.

Ethical Approval

This article is an Insilico work and does not require Ethical Approval.

Availability of Data and Materials

The datasets used and/or analyzed during the current study would be available from the corresponding author on reasonable request.

Author Contributions

All authors approved the final version of the article for submission.

Conflict of Interest

The authors declared no competing interests.

Reference

1. Bennett JE, Dolin R, Blaser MJ. Mandell, douglas, and bennett's principles and practice of infectious diseases: 2-volume set: Elsevier Health Sciences. Philadelphia, Saunders, Elsevier; 2014.

2. DruSzczyńSka M, Kowalewicz-Kulbat M, Fol M, WłOdarczyk M, RuDnIcka W. Latent M. tuberculosis infection--pathogenesis, diagnosis, treatment and prevention strategies. Pol J Microbiol. 2012;61(1):3-10. [DOI:10.33073/pjm-2012-001] [PMID]

3. Maji A, Misra R, Dhakan DB, Gupta V, Mahato NK, Saxena R, et al. Gut microbiome contributes to impairment of immunity in pulmonary tuberculosis patients by alteration of butyrate and propionate producers. Environ Microbiol. 2018;20(1):402-19. [DOI:10.1111/1462-2920.14015] [PMID]

4. Singhvi N, Singh Y, Shukla P. Computational approaches in epitope design using DNA binding proteins as vaccine candidate in Mycobacterium tuberculosis. Infect Genet Evol. 2020;83:104357. [DOI:10.1016/j.meegid.2020.104357] [PMID]

5. Gholami A, Moosavi L. Incidence Rate of ExtraPulmonary Tuberculosis in Urmia city during 2004-2007. Nurs midwifery res j. 2010;8(2):0.

6. Ong E, He Y, Yang Z. Epitope promiscuity and population coverage of Mycobacterium tuberculosis protein antigens in current subunit vaccines under development. Infect Genet Evol. 2020;80:104186. [DOI:10.1016/j.meegid.2020.104186] [PMID]

7. Barker LF, Brennan MJ, Rosenstein PK, Sadoff JC. Tuberculosis vaccine research: the impact of immunology. Curr Opin Immunol. 2009;21(3): 331-8. [DOI:10.1016/j.coi.2009.05.017] [PMID]

8. Daniel T, Bates J, Downes K. History of tuberculosis. Tuberculosis: pathogenesis, protection, and control. 1994 (16):13-24. [PMCID] [DOI:10.1128/9781555818357.ch2] [PMID]

9. Dimitrov I, Bangov I, Flower DR, Doytchinova I. AllerTOP v. 2-a server for in silico prediction of allergens. J Mol Model. 2014;20(6):1-6. [DOI:10.1007/s00894-014-2278-5] [PMID]

10. Dimitrov I, Naneva L, Doytchinova I, Bangov I. AllergenFP: allergenicity prediction by descriptor fingerprints. Bioinformatics. 2014;30(6):846-51. [DOI:10.1093/bioinformatics/btt619] [PMID]

11. Gupta S, Kapoor P, Chaudhary K, Gautam A, Kumar R, Open Source Drug Discovery C, et al. In silico approach for predicting toxicity of peptides and proteins. PloS One. 2013;8(9):e73957. [DOI:10.1371/journal.pone.0073957] [PMID] [PMCID]

12. Thevenet P, Shen Y, Maupetit J, Guyon F, Derreumaux P, Tuffery P. PEP-FOLD: an updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides. Nucleic Acids Res. 2012;40(W1):W288-W93. [DOI:10.1093/nar/gks419] [PMID] [PMCID]

13. Shen Y, Maupetit J, Derreumaux P, Tufféry P. Improved PEP-FOLD approach for peptide and miniprotein structure prediction. J Chem Theory Comput. 2014;10(10):4745-58. [DOI:10.1021/ct500592m] [PMID]

14. Zhang Y, Skolnick J. Scoring function for automated assessment of protein structure template quality. Proteins. 2004;57(4):702-10. [DOI:10.1002/prot.20264] [PMID]

15. Kumar V, Kancharla S, Kolli P, Jena M. Reverse vaccinology approach towards the in-silico multiepitope vaccine development against SARS-CoV-2. F1000Research. 2021;10. [PMID] [PMCID] [DOI:10.12688/f1000research.36371.1]

16. Ikai A. Thermostability and aliphatic index of globular proteins. J Biochem. 1980;88(6):1895-8.

17. Jespersen MC, Peters B, Nielsen M, Marcatili P. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. Nucleic Acids Res. 2017;45(W1):W24-W9. [DOI:10.1093/nar/gkx346] [PMID] [PMCID]

18. Singh H, Ansari HR, Raghava GPS. Improved method for linear B-cell epitope prediction using antigen's primary sequence. PloS One. 2013;8(5):e62216. [PMID] [PMCID] [DOI:10.1371/journal.pone.0062216]

19. El-Manzalawy Y, Dobbs D, Honavar V. Predicting linear B-cell epitopes using string kernels. J Mol Recognit: An Interdisciplinary Journal. 2008;21 (4):243-55. [DOI:10.1002/jmr.893] [PMID] [PMCID]

20. Faria AR, Costa MM, Giusta MS, Grimaldi Jr G, Penido MLO, Gazzinelli RT, et al. High-throughput analysis of synthetic peptides for the immunodiagnosis of canine visceral leishmaniasis. PLoS Negl Trop Dis. 2011;5(9): e1310. [DOI:10.1371/journal.pntd.0001310] [PMID] [PMCID]

21. Kaye K, Frieden TR. Tuberculosis control: the relevance of classic principles in an era of acquired immunodeficiency syndrome and multidrug resistance. Epidemiol Rev. 1996;18 (1): 52-63. [DOI:10.1093/oxfordjournals.epirev.a017916] [PMID]

22. Cosivi O, Grange JM, Daborn CJ, Raviglione MC, Fujikura T, Cousins D, et al. Zoonotic tuberculosis

due to Mycobacterium bovis in developing countries. Emerging Infect Dis. 1998;4(1):59. [DOI:10.3201/eid0401.980108] [PMID] [PMCID]

23. Singh H, Raghava GPS. ProPred: prediction of HLA-DR binding sites. Bioinformatics. 2001;17 (12): 1236-7. [DOI:10.1093/bioinformatics/17.12.1236] [PMID]
24. Yang X, Yu X. An introduction to epitope prediction methods and software. Rev Med Virol. 2009;19(2):77-96. [DOI:10.1002/rmv.602] [PMID]
25. Waaler H. Tuberculosis and poverty. Int J Tuberc Lung Dis. 2002 6(9):745-6.
26. Elhag M, Sati A, Saadaldin M, Hassan M. Immunoinformatics prediction of epitope-based peptide vaccine against Mycobacterium tuberculosis PPE65 family protein. bioRxiv2019 p. 755983. [DOI:10.1101/755983]
27. Colditz GA, Brewer TF, Berkey CS, Wilson ME, Burdick E, Fineberg HV, et al. Efficacy of BCG vaccine in the prevention of tuberculosis: meta-analysis of the published literature. Jama. 1994;271(9):698-702. [PMID] [DOI:10.1001/jama.1994.03510330076038]
28. Iseman MD. Evolution of drug-resistant tuberculosis: a tale of two species. Proc Natl Acad Sci. 1994;91(7):2428-9. [DOI:10.1073/pnas.91.7.2428] [PMID] [PMCID]
29. Schaaf HS, Marais BJ, Whitelaw A, Hesseling AC, Eley B, Hussey GD, et al. Culture-confirmed childhood tuberculosis in Cape Town, South Africa: a review of 596 cases. BMC Infect Dis. 2007;7(1):1-8. [DOI:10.1186/1471-2334-7-140] [PMID] [PMCID]
30. Riccomi A, Piccaro G, Christensen D, Palma C, Andersen P, Vendetti S. Parenteral vaccination with a tuberculosis subunit vaccine in presence of retinoic acid provides early but transient protection to M. tuberculosis infection. Front Immunol. 2019;10:934. [PMID] [PMCID] [DOI:10.3389/fimmu.2019.00934]
31. Reginald K, Chan Y, Plebanski M, Poh CL. Development of peptide vaccines in dengue. Curr Pharm Des. 2018;24(11):1157-73. [PMID] [PMCID] [DOI:10.2174/1381612823666170913163904]
32. Bahrami AA, Payandeh Z, Khalili S, Zakeri A, Bandehpour M. Immunoinformatics: in silico approaches and computational design of a multi-epitope, immunogenic protein. Int Rev Immunol. 2019;38(6):307-22. [DOI:10.1080/08830185.2019.1657426] [PMID]
33. Watt J, Liu J. Preclinical progress of subunit and live attenuated Mycobacterium tuberculosis vaccines: A review following the first in human efficacy trial. Pharmaceutics. 2020;12(9):848. [DOI:10.3390/pharmaceutics12090848] [PMID] [PMCID]
34. Lu LL, Suscovich TJ, Fortune SM, Alter G. Beyond binding: antibody effector functions in infectious diseases. Nat Rev Immunol. 2018;18(1):46-61. [DOI:10.1038/nri.2017.106] [PMID] [PMCID]
35. Krocova Z, Plzakova L, Pavkova I, Kubelkova K, Macela A, Ozanic M, et al. The role of B cells in an early immune response to Mycobacterium bovis. Microb Pathog. 2020;140:103937. [DOI:10.1016/j.micpath.2019.103937] [PMID]
36. Singhvi N, Gupta V, Gaur M, Sharma V, Puri A, Singh Y, et al. Interplay of human gut microbiome in health and wellness. Indian J Microbiol. 2020;60(1):26-36. [DOI:10.1007/s12088-019-00825-x] [PMID] [PMCID]
37. Eickhoff CS, Terry FE, Peng L, Meza KA, Sakala IG, Van Aartsen D, et al. Highly conserved influenza T cell epitopes induce broadly protective immunity. Vaccine. 2019;37(36):5371-81. [PMID] [PMCID] [DOI:10.1016/j.vaccine.2019.07.033]
38. Patankar YR, Sutiwisesak R, Boyce S, Lai R, Lindestam Arlehamn CS, Sette A, et al. Limited recognition of Mycobacterium tuberculosis-infected macrophages by polyclonal CD4 and CD8 T cells from the lungs of infected mice. Mucosal Immunol. 2020;13(1):140-8. [PMCID] [DOI:10.1038/s41385-019-0217-6] [PMID]
39. Russell SL, Lamprecht DA, Mandizvo T, Jones TT, Naidoo V, Addicott KW, et al. Compromised metabolic reprogramming is an early indicator of CD8+ T cell dysfunction during chronic Mycobacterium tuberculosis infection. Cell reports. 2019;29(11):3564-79. [PMID] [PMCID] [DOI:10.1016/j.celrep.2019.11.034]
40. Meza B, Ascencio F, Sierra-Beltrán AP, Torres J, Angulo C. A novel design of a multi-antigenic, multistage and multi-epitope vaccine against Helicobacter pylori: an in silico approach. Infect Genet Evol. 2017;49:309-17. [DOI:10.1016/j.meegid.2017.02.007] [PMID]
41. Chen R. Bacterial expression systems for recombinant protein production: E. coli and beyond. Biotechnol Adv. 2012;30(5):1102-7. [DOI:10.1016/j.biotechadv.2011.09.013] [PMID]
42. Rosano GL, Ceccarelli EA. Recombinant protein expression in Escherichia coli: advances and challenges. Front Immunol. 2014;5:172. [DOI:10.3389/fmicb.2014.00172]