Syrian Arab Republic
Ministry of Higher Education and Scientific Research
Syrian Virtual University
Master in Bioinformatics (BIS)

Designing a virtual multi-epitope vaccine against Helicobacter pylori using immunoinformatics tools

تصميم لقاح افتراضي ضد جرثومة الملوية البوابية بمساعدة أدوات المعلوماتية المناعية

A thesis submitted as a fulfillment of requirements for a Master's degree in Bioinformatics

Authored by:

Ph. Manal Al Allawi

Supervised by:

Dr. Aya Twair

Acknowledgments:

I would like to thank my supervisor, Dr. Aya Twair, for her support, guidance during my dissertation.

Special thanks to my family and friends for their support and understanding were invaluable during the period of the master's degree program.

Table of Contents

LIST OF FIGURE

List of Table

List of Abbreviation

Abbreviation	Meaning
MALT	Mucosa-Associated Lymphoid Tissue Lymphoma
WHO	World Health Organization
STT	Standard Triple Therapy
RV	Reverse Vaccinology
VacA	Vacuolating cytotoxin autotransporter
HpaA	Neuraminyllactose-binding hemagglutinin
LysA	Diaminopimelate decarboxylase
FucT	Alpha-(1,3)-fucosyltransferase FucT
AroE	Shikimate dehydrogenase (NADP(+)
Omp6	Outer membrane protein6
OipA	Outer inflammatory protein A
Omp18	Outer membrane protein 18
UreB	Urease
OMPs	Outer membrane proteins
IFN-y	Interferon-y
IL-8	Interleukin-8
HLA	Human Leukocyte Antigen
CTL	Cytotoxic T Lymphocyte
HTL	Helper T Lymphocyte
MHC	Major histocompatibility complex
APCs	surface of antigen-presenting cells
SVM	Support Vector machine
GRAVY	Grand average of hydropathicity
TLR2/TLR4	Toll-like receptor-2/-4
CAI	Codon adaptation index

Abstract

Helicobacter pylori is a gram-negative, spiral, microaerophilic bacterium that infects the stomachs of over 50% of the global human population. It is commonly acquired during childhood and, if left untreated, can persist chronically, leading to conditions such as chronic gastritis, peptic ulcer disease, gastric adenocarcinoma, and gastric B cell lymphoma. The current treatment approach involves proton-pump inhibitors and antibiotics, but it faces challenges like patient compliance, antibiotic resistance, and potential recurrence of infection. Developing an effective vaccine against H. pylori would offer significant advantages.

During this study a virtual vaccine against Helicobacter pylori was designed using immunoinformatics approaches targeting specific antigens known to be involved in the pathogenesis of the infection. The selected protective antigens include Vacuolating cytotoxin autotransporter (vacA), Neuraminyllactose-binding hemagglutinin (hpaA), Diaminopimelate decarboxylase (lysA), Alpha-(1,3)-fucosyltransferase FucT (fucT), Shikimate dehydrogenase (NADP(+)) (aroE), Outer inflammatory protein A (oipA), Outer membrane proteins(omp 6/18), Urease (ureB), and IceA2.

The epitopes underwent a thorough filtering process, including tests for antigenicity, toxicity, allergenicity, and cytokine inducibility, with the primary aim of identifying epitopes capable of triggering both T and B cell-mediated immune responses. To enhance vaccine immunogenicity, the final epitopes were fused with the Heat-labile enterotoxin B chain from E.coli (LTB) adjuvant using appropriate linkers, resulting in the development of a multi-epitope vaccine. The selected T cell epitope ensemble is expected to cover 78.09% of the global human population.

Furthermore, docking studies were conducted to assess the interaction between the vaccine and Toll-like receptor 2 (TLR2) / Toll-like receptor 4 (TLR4), demonstrating significant affinity, consistency, and stability. This research aimed to design an effective subunit vaccine capable of eliciting both humoral and cellular immune responses, with the objective of addressing antibiotic-resistant Helicobacter pylori infections.

1.Introduction

Helicobacter pylori is a Gram-negative, microaerophilic, flagellated bacteria, which usually infects more than half of the world's population [1][2]. If left untreated, it can be colonized in the stomach mucosa, persist throughout a person's life, and gradually lead to gastric diseases such as gastritis, peptic ulcer, mucosa-associated lymphoid tissue lymphoma (MALT), and stomach cancer [3]. The World Health Organization (WHO) has classified this bacterium in the first category of carcinogen agents, but currently [4], the only way to deal with this bacterium is antibiotic administration , Traditionally, H. pylori had been treated with a standard triple therapy (STT) consisting of proton pump inhibitors (PPIs), clarithromycins (CAMs), and amoxicillins (AMPCs)[5].

H. pylori resists antibiotics by various mechanisms, such as preventing antibiotic entry, changing their site of action (regarding macrolides, quinolones, and beta-lactams), antibiotic inactivation (beta-lactams), and removal after penetration into the cell (tetracycline). Thus, a therapeutic approach for reduction of the resistance likelihood is combination antibiotic therapy. Due to the overuse and constant use of antibiotics, *H. pylori* resistance to antibiotic therapy is increasing. Therefore, designing an appropriate vaccine against this bacterium is necessary [6][7].

It has been proposed to develop a new generation of vaccines using immunoinformatics, which is considered the most informative and advantageous tool for vaccine design[8]. Immunoinformatics, a subset of bioinformatics, offers a novel approach to analyzing large amounts of immunological data obtained from experimental research using various tools and databases[9]. Reverse vaccinology (RV) is a key concept in immunoinformatics, involving the in-silico screening of a pathogen's entire genome to identify genes encoding proteins suitable for vaccine targets[10]. Immunoinformatics also plays a crucial role in predicting B and T cell epitopes, reducing the time and costs associated with experimental analysis. Epitopes are selected based on their accessibility to immune system surveillance through informatics-based prediction methodologies. The use of immunoinformatics has been instrumental in

designing multi-epitope vaccines, with numerous databases and web servers available for predicting B and T cell epitopes [9].

In this study, vacA , hpaA, omp6, omp18, oipA, ureB and IceA2 have been identified as potential antigens through Reverse Vaccinology (RV). Subsequently, epitope prediction was conducted to characterize B and T cell epitopes using computational methods. The next phase involved designing optimal epitopes with adjuvants and linkers, followed by modeling and docking of peptides to ensure broad population coverage [11].

2.Dataset

The reference strain Helicobacter pylori (formerly known as Campylobacter pylori) was chosen for this study. After reviewing numerous medical literature sources, several proteins were identified as potential candidates for inclusion in Helicobacter pylori vaccines. These proteins were further evaluated for their antigenicity, virulence, and subcellular localization using various bioinformatics tools and online servers. The Fasta formats of these proteins were retrieved from the UniProt database to facilitate the assessment process. Table (1) shows the chosen proteins with their references and accession number in Uniprot:

Table 1 chosen proteins from multiple medical literature

	Protein	**Reference**	**Accession Number in Uniprot**
1	Vacuolating cytotoxin autotransporter (vacA)	[12]	Q48245
2	Neuraminyllactose-binding hemagglutinin (hpaA)	[13]	Q48244
3	Diaminopimelate decarboxylase (lysA)	[14]	B4XMC6
4	Alpha-(1,3)-fucosyltransferase FucT(fucT)	[15]	O30511
5	Shikimate dehydrogenase (NADP(+)) (aroE)	[16]	Q56S04
6	Outer membrane protein Omp6 (Omp6)	[17]	A0A448UFR2
7	Outer inflammatory protein A (oipA)	[18]	U3GLE7
8	Outer membrane protein Omp18	[19]	A0A7K1N3E4
9	Urease (ureB)	[20]	Q9ZNC3
10	IceA2	[21]	L7SY90

The mentioned proteins were checked by several tools in order to know their potentials and characters. The first tool is ANTIGENpro which is a sequence-based, alignment-free and pathogen-independent predictor of protein antigenicity. ANTIGENpro is the first predictor of the whole protein antigenicity trained using reactivity data obtained by protein microarray analysis for five pathogens. Users typically submit protein sequences through the web interface, and the tool generates predictions based on its algorithms. They were also checked by VaxiJen which is a computational tool used for predicting the antigenicity of protein sequences. It is designed specifically to identify potential vaccine candidates by assessing the immunogenicity of proteins.The model used in this research was Vaxijen v2.0 and the threshold for this model was 0.5 to all listed proteins. Finally, the proteins were checked by Psortb server, which is a tool used for bacterial protein subcellular localization prediction.

Table 2 Result of many tools that reflects proteins specification

NO	Protein	ANTIGENpro Predicted Probability of Antigenicity”	VaxiJen Overall Prediction for the Protective Antigen =	Psortb
1	vacA	0.957928	0.5746 (Probable ANTIGEN).	EXTRACELULAR
2	hpaA	0.842659	0.5615 (Probable ANTIGEN).	Unknown
3	lysA	0.079965	0.4526 (Probable NON-ANTIGEN).	Cytoplasmic
4	fucT	0.597999	0.4804 (Probable NON-ANTIGEN)	Cytoplasmic
5	aroE	0.129312	0.3141 (Probable NON-ANTIGEN)	Cytoplasmic
6	Omp6	0.860049	0.5026 (Probable ANTIGEN)	Extracellular
7	oipA	0.220659	0.5617 (Probable ANTIGEN).	Unknown
8	Omp18	0.893562	0.7351 (Probable ANTIGEN).	Outer Membrane
9	UreB	0.891255	0.8285 (Probable ANTIGEN).	Cytoplasmic
10	IceA2	0.619319	0.5852 (Probable ANTIGEN).	Unknown

3.Antigen selection:

Helicobacter pylori is a Gram-negative bacterium that infects the stomach mucosa of more than half of the world's population, leading to a range of gastric diseases from gastritis to cancer. Despite the majority of infected individuals being asymptomatic, this bacterium is classified as a class I human carcinogen by the World Health Organization. Current treatments for H. pylori eradication are not completely effective, highlighting the urgent need for a vaccine. Developing a rational and well-thought-out vaccine design against H. pylori is crucial for global public health[22].

The design of an effective vaccine against H. pylori must consider the interference with the mechanisms of pathogenic bacteria. Therefore, vaccine design should contain molecules derived from different stages (multistage) of H. pylori pathogenesis. Recently, significant progress has been made in understanding H. pylori pathogenesis and the role of its virulence factors in gastric diseases. Several studies have revealed that at least three distinct and

sequential stages are required for H. pylori to exert its virulence on the colonized stomach: adhering to and colonizing the surface of gastric epithelial cells, evading and attenuating the host defense, and invading and damaging gastric mucosa [23].

Among the bacterial virulence factors is the vacuolating cytotoxin A (VacA), which represents an important determinant of pathogenicity. Intensive research on VacA has uncovered a wide range of mechanisms that contribute to interactions between the host and the pathogen. This toxin is a key focus of ongoing studies for several reasons: i) VacA possesses unique structural properties and functions that distinguish it from other bacterial toxins; ii) its role in disease progression and stomach colonization by H. pylori is well-established; iii) VacA is a promising candidate for inclusion as an antigen in a vaccine targeting H. pylori; and iv) the high allelic diversity of the vacA gene means that different variants contribute to H. pylori's pathogenicity in varying ways[24].

Outer membrane proteins (OMPs) are essential components of the outer plasma membrane of Helicobacter pylori. They play crucial roles in functions such as ion transport, adherence, structural and osmotic stability, and bacterial virulence. Additionally, due to their exposure on the bacterial surface, OMPs may also serve as antigens [25]. Omp18 of H. pylori is a powerful antigen that can induce significant interferon-γ (IFN-γ) levels [26].Omp6 is involved in helicobacter pylori biofilm formation , it play a crucial role in the pathogenesis and survival of the bacterium[27].

HpaA is a member of the outer membrane protein group in H. pylori. It has been identified as an adherence factor for blood cells [28] and is considered a putative neuraminyllactose-binding hemagglutinin. Additionally, HpaA has been characterized as a flagellar sheath protein [29] and serves as a surface-localized antigen recognized by human antibodies [30].

OipA (outer inflammatory protein A) is a crucial virulence factor that contributes to the elevation of interleukin-8 (IL-8) secretion, the production of proinflammatory factors leading to neutrophil infiltration, exacerbation of stomach inflammation, and interaction with the host cell membrane to facilitate colonization [31].

UreB is a subunit of the urease enzyme produced by H. pylori. Urease plays a crucial role in the pathogenesis of H. pylori by neutralizing the acidity of the local environment and promoting bacterial survival. This is achieved through the decomposition of urea, which is abundant in the stomach, into ammonia and carbon dioxide, thereby helping H. pylori to survive in the acidic gastric environment [32]. Due to its immunogenic properties, UreB is considered an ideal vaccine antigen against H. pylori infection [33].

IceA2 is one of the two main allelic variants of the iceA gene, with IceA1 being the other variant. Studies have suggested that the presence of IceA2 may be associated with an

increased risk of developing peptic ulcer disease and gastric cancer in individuals infected with H. pylori [34].

4.Problems & Aims

Problem:

Helicobacter pylori is a type of antibiotic-resistant germs, making it a significant life-threatening risk.

Aim:

Finding a safe and effective vaccine against the *Helicobacter pylori* using bioinformatics tools because it is a type of antibiotic-resistant germs, making it a significant life threatening risk.

5.Methodology:

❖ Workflow:

1- Retrieve protein sequences in Fasta format for the selected proteins.
2- Utilize bioinformatics tools to predict epitopes, considering Human Leukocyte Antigen (HLA) Supertypes. Focus on Cytotoxic T Lymphocyte (CTL) and Helper T Lymphocyte (HTL) epitope prediction, while also including B Lymphocyte epitope prediction.
3- Calculate population coverage to assess the reach and effectiveness of the designed vaccine.
4- Assemble the vaccine based on the predicted epitopes.
5- Conduct a comprehensive analysis, evaluating allergenicity, antigenicity, solubility, and physicochemical characteristics of the designed vaccine.
6- Perform a thorough two-dimensional structure analysis of the vaccine.
7- Employ three-dimensional structure modeling and refinement techniques to enhance the accuracy of the vaccine structure.
8- Identify discontinuous B-cell epitopes, specifically focusing on conformational epitopes.
9- Stimulate immune responses to analyze the effectiveness and potential reactions to the designed vaccine.
10- Utilize molecular docking studies to explore the interactions between the vaccine and relevant target molecules.

5.1. Retrieve protein sequences from Uniprot

Retrieve protein sequence in Fasta format for the selected proteins, excluding lysA , fucT and aroE due to its weak antigenicity.

Table 3 List of chosen proteins that reflects proteins specifications

	Protein	Accession number in Uniprot	Refernce
1	Vacuolating cytotoxin autotransporter (vacA)	Q48245	[12]
2	Neuraminyllactose-binding hemagglutinin (hpaA)	Q48244	[13]
3	Outer membrane protein Omp6 (Omp6)	0A448UFR2	[17]
4	Outer inflammatory protein A (oipA)	U3GLE7	[18]
5	Outer membrane protein Omp18	A0A7K1N3E4	[19]
6	Urease (ureB)	Q9ZNC3	[20]
7	IceA2	L7SY90	[21]

5.2- Epitopes Predictions:

The human leukocyte antigen (HLA) system,also known as the major histocompatibility complex[MHC] in human ,is a crucial component of the immune system located on chromosome 6, encodes cell surface molecules that present antigenic peptides to T cells[35]. It consists of two classes: class I (HLA-A, HLA-B, HLA-C) and class II (HLA-DR, HLA-DQ, HLA-DP) [36].

5.2.1- Cytotoxic T Lymphocyte (CTL) predictions:

We performed epitope prediction for Cytotoxic T Lymphocyte (CTL) for all proteins using the Immune Epitope Database Analysis Resource website, which is an adjunct to the Immune Epitope Database (IEDB). Specifically, we utilized the netmhcpan 4.1el method to assess the ability of each subsequence within an amino acid sequence (or set of sequences) to bind to a specific MHC class I molecule.

5.2.2-Helper T lymphocytes (HTL) prediction:

CD4+ T cells, also known as helper T cells, play a vital role in the immune system by assisting other immune cells in their functions. These cells recognize specific molecular structures called epitopes on the surface of antigens, which are typically proteins or peptides derived from pathogens such as bacteria, viruses, or other foreign substances. Epitopes are often peptides that are presented on the surface of antigen-presenting cells (APCs), such as dendritic cells, macrophages, or B cells. These APCs digest antigens into smaller peptides and display them on their surface bound to molecules called major histocompatibility complex class II (MHC-II). Helper T cells recognize these epitopes when they are presented by MHC-II molecules on APCs. This interaction triggers a series of immune responses, including the activation of B cells (resulting in antibody production) and cytotoxic T cells that can directly kill infected cells[37].

The majority of prediction algorithms for human MHC class II to date have focused on HLA molecules encoded in the DR locus, which represents a significant gap in knowledge as HLA DP and DQ molecules are likely equally important but have been studied less due to experimental challenges. In the Syrian population, the most frequent alleles in the DRB1 locus were DRB1*11 (26.4%), DRB1*04 (14%), and DRB1*07 (12%). Conversely, the most common observed alleles at the DQB1 locus were DQB1*03 (40.9%) and DQB1*05 (25.1%)[38].

5.2.3- B Lymphocyte epitope prediction:

B-cell epitopes are crucial for the development of peptide vaccines, and the ABCpred server, which employs an artificial neural network, was utilized to predict linear B-cell epitope regions. This tool aids in the identification of epitope regions that are valuable for selecting synthetic vaccine candidates[39][40].

5.2.4-Epitope Screening and Characterization:

We utilized VaxiJen v2.0 to predict the antigenicity of the epitope sequences, with a prediction parameter accuracy set to 0.5. Additionally, the predicted epitopes underwent assessment on the AllerTop V 2.0 server to eliminate allergic proteins. AllerTop employs an alignment-free technique based on key physicochemical features of proteins, demonstrating a server sensitivity of approximately 94% in predicting allergens. To evaluate the toxicity of selected epitopes, we employed ToxinPred, utilizing the support vector machine (SVM) approach on the server, with all parameters set to their default values. SVM, a widely adopted machine-learning technique for toxicity prediction, effectively distinguishes between toxic and non-toxic epitopes. Other tests run on the epitopes based on their types will be mentioned with the results.

5.3 population coverage :

This study was conducted to investigate the HLA genetic diversity in the Syrian population. Human Leukocyte Antigen (HLA) alleles demonstrate significant variability among different ethnic groups, and play crucial roles in disease susceptibility and resistance. The frequency of HLA class I alleles (A, B, and C) was examined in 105 unrelated, healthy Syrian individuals from various regions, revealing a total of 58 observed alleles. The most common alleles for HLA class I A locus were A*02, A*24, A*01, and A*03; for B locus were B*35, B*51, B*44, and B*52; and for C locus were C*04, C*07, C*12, and C*06 [36]. The test was conducted using the population coverage tool on the IEDB website.

5-4-Multi-Epitope Vaccine Construction :

Because of their small size, peptides do not trigger an immune response when used alone as vaccines. To address this, they require a carrier with a strong immune-stimulating adjuvant to activate both the innate and adaptive immune systems. Linkers also play a crucial role in mimicking the vaccine's ability to function as an independent immunogen, helping to generate higher antibody concentrations than those achieved with a single immunogen. To construct a multi-epitope vaccine , CTL, HTL and B cell were joined in orderly fashion with appropriate linker. The vaccine construct started with Heat-labile enterotoxin B chain from E.coli (LTB) adjuvant [44] at the N- terminal end followed by EAAAK linker. AAY linkers were added to joine each CTL epitope, whereas GPGPG and KK linkers were used to joine HTL and B cell, respectively. The resulting vaccine was checked for both antigenic and non-allergenic properties.

LTB sequence was retrieved from Uniprot (Uniprot accession number: **A0A2S1ZC12).**

5-5- Analysis of physicochemical Attributes , Solubility Prediction , Secondary Structure for the Vaccine Construct :

The physical and chemical characteristics of the multi-epitope vaccines were analyzed using the ProtParam tool. This analysis included an assessment of amino acid composition, estimated half-life, instability index, extinction coefficient, theoretical pI, atomic composition, molecular weight, and the grand average of hydropathicity (GRAVY) to facilitate subsequent experimental investigations. The stability index, a crucial parameter, played a pivotal role in identifying and eliminating protein candidates with an instability index greater than 40. [45]

Solubility prediction were assisted by both Protein-sol web tool [46] and SOLpro Solubility predictor[47].

Protein-sol is a simple and free, web-based suite of theoretical calculations
and predictive algorithms for understanding protein solubility and stability. The scaled solubility value is the predicted solubility[46].

The SOLpro server was utilized to predict protein solubility, specifically determining the likelihood of a protein being soluble upon overexpression in E. coli through a two-stage SVM architecture. Each classifier in the initial layer received distinct sets of attributes to characterize the sequence, and the final SVM classifier synthesized the data to predict the protein's solubility and the associated probability [47].

Secondary structural attributes, including α-helix, β-strand, and random coils, were evaluated using the

Sopma server.

5-6 Tertiary Structure Prediction and Validation of the Multi-Epitope Vaccine

The 3D structure prediction and validation of the multi-epitope vaccine involved modeling the potential configuration of the combined protein vaccine using the PHYRE2 Server[49]. Phyre2 utilizes advanced remote homology detection techniques to construct 3D models, predict ligand binding sites, and assess the impact of amino acid variants in a user's protein sequence.

Subsequently, the protein's structure underwent validation using the PROCHECK online server[50], which provides a Ramachandran plot. This plot details the percentage of residues in favored, allowed, and outlier regions, allowing for the evaluation of the modeled tertiary structure's quality. To enhance the model, GalaxyRefine was utilized, and both the pre- and post-refinement models were subjected to ProSA-web protein structure analysis [51]. ProSA is a widely-used tool that assesses 3D models for potential errors using the zscore, indicating overall model quality by measuring the deviation of the structure's total energy from an energy distribution derived from random conformations. Z-scores outside the range characteristic for native proteins indicate potential structural errors.

5-7 Defining discontinuous B. cell epitope (conformation) :

B-cell epitope play an important role in humoral response , were identified with the refined 3D protein

Vaccine model. The prediction of conformational B-cell epitope was conducted using the Ellipro Server [52].

5-8 Immune response Simulation:

The probability of the designed vaccine inducing both humoral and cellular immune responses was further assessed using the C-IMMSIM server[53] .

5-9 Molecular docking :

For the receptor-vaccine docking analysis, we modeled the 3D structures of Toll-like receptor 2 (TLR2) from the Uniprot (Uniprot ID: O60603) and Toll-like receptor 4 (TLR4) (Uniprot:O00206) using phyre2. Subsequently, protein-receptor docking was performed

using the ClusPro 2.0 server. ClusPro employs docking approaches such as DOT and ZDOCK, both utilizing FFT correlation techniques[54], This method categorizes docked conformations based on pairwise RMSD histograms, predicting a range of docking poses that span stable (deep energy minima) to transient (shallow minima) interactions.

The intermolecular energy landscape was systematically mapped to identify favorable docking poses using ClusPro. These poses were further analyzed and validated using Chimera, an interactive visualization and analysis tool for molecular structures. Chimera supports the examination of density maps, molecular trajectories, and sequence alignments, providing insights into the structural details and interactions between the TLR receptors and the vaccine candidate[55].

This integrated approach ensures comprehensive exploration and validation of potential binding configurations between the receptor proteins (TLR2 and TLR4) and the vaccine candidate, leveraging computational tools to elucidate molecular interactions crucial for vaccine design and development.

5-10 Codon adaptation and in-silico cloning:

The multi-epitope vaccine sequence underwent reverse translation and subsequent optimization for Escherichia coli codon usage, enhancing the expression efficiency when cloned into the specified expression system. VectorBuilder server facilitated this process. The GC content and codon adaptation index (CAI) of the cloned sequences were evaluated to gauge their expression levels. An optimal CAI ratio falls within the range of 0.8 to 1, ensuring favorable transcriptional and translational efficiency. Simultaneously, the desirable GC content should range between 30% and 70%. The engineered construct was cloned into the pET-28a (+) expression vector using SnapGene.

Results & Discussion

6-Results:

6-1 Proteins sequences retrieval in Fasta format Uniprot.

6-2 Epitope prediction:

6-2-1 Cytotoxic T Lymphocyte (CTL) predictions:

The most common alleles for HLA class I A locus in the Syrian population were: A*02, A*24, A*01, A*03; for B locus were: B*35, B*51, B*44, B*52 and for C locus were: C*04, C*07, C*12, C*06 [36]

We employed the VaxiJen 2.0 server, specifically designed for predicting protective antigens, tumor antigens, and subunit vaccines, to evaluate the immunogenic potential of each epitope sequence. A threshold of 0.5 was set for this assessment, aligning with the optimal accuracy observed in many models at this threshold. To identify highly immunogenic peptides, the selected epitopes underwent analysis using the Class I Immunogenicity tool from IEDB Analysis Resource, with scores above 0.9 indicating significant immunogenicity [41].
To ensure the safety of the vaccine components, accepted epitopes were screened for allergenicity using the AllerTOP v. 2.0 tool. It is crucial that the vaccine components do not trigger allergic responses [42] . Additionally, we utilized the ToxinPred tool to assess the toxicity of accepted epitopes. This in-silico method specializes in predicting the toxicity of peptides and proteins, aiding in designing peptides with minimal toxicity and identifying potential toxic regions in proteins [43] .The best 5 result protein (according to the fact that hige score =good binder) was taken and it will undergo more tests to choose the best ones ever.

Table 4 CTLs prediction of vacA peptides

peptide	score	Toxicity	Allergenticity	Immunogenicity score	overall Prediction for the Protective Antigen
YLAPSYSTI	0.96218	Non-Toxin	**allergen**	-0.27998	0.2515 (Probable NON-ANTIGEN).
FYYSPWNYF	0.975024	Non-Toxin	**allergen**	0.02242	1.1762 (Probable ANTIGEN)
LIDSHDAGY	0.906738	Non-Toxin	non-allergen	-0.05407	0.2140 (Probable NON-ANTIGEN).
KVWRIQAGK	0.910978	Non-Toxin	non-allergen	0.19736	**0.6791** (Probable **ANTIGEN**).
SLYDGATLNL	0.904618	Non-Toxin	non-allergen	0.11562	**0.8078** (Probable **ANTIGEN**).

Table 5 CTLs prediction of hpaA

peptide	score	Toxicity	Allergenticity	Immunogenicity score	overall Prediction for the Protective Antigen
ALDEKILLL	0.989044	Non-Toxin	non-allergen	0.00739	0.2826 (Probable NON-ANTIGEN).

VLIPAGFVK	0.783863	Non-Toxin	allergen	0.22496	-0.2989 (Probable NON-ANTIGEN).
SSDKDDLSF	0.640548	Non-Toxin	non-allergen	-0.27334	**0.6799** (Probable **ANTIGEN**)
VEQILQNQGY	0.683623	Non-Toxin	allergen	-0.11133	-0.7108 (Probable NON-ANTIGEN
HPASEKVQAL	0.587228	Non-Toxin	allergen	-0.30231	0.2896 (Probable NON-ANTIGEN).

Table 6 CTLs prediction of omp6

peptide	**score**	**Toxicity**	**Allergenticity**	**Immunogenicity score**	**overall Prediction for the Protective Antigen**
IPTINTNYY	0.955319	Non-Toxin	non-allergen	0.16914	**0.6062** (Probable **ANTIGEN**
RLYSLYLNY	0.949455	Non-Toxin	allergen	-0.19509	**1.4922** (Probable **ANTIGEN**).
AYMSVGYQI	0.941776	Non-Toxin	allergen	-0.22217	**0.7204** (Probable **ANTIGEN**).
GEIQKVSNAY	0.927164	Non-Toxin	allergen	-0.38149	**0.4624** (Probable **NON-ANTIGEN**).
SVINDTISY	0.914103	Non-Toxin	allergen	0.11049	**0.3573** (Probable **NON-ANTIGEN**)

Table 7 CTLs prediction of OipA

peptide	score	Toxicity	Allergenticity	Immunogenicity score	overall Prediction for the Protective Antigen
NELSFGYKY	0.983979	Non-Toxin	allergen	-0.15329	**0.8610** (Probable **ANTIGEN**).
NELSFGYKYFNE	0.626755	Non-Toxin	allergen	-0.21071	**0.6591** (Probable **ANTIGEN**).
SLFSEQNTK	0.8819	Non-Toxin	non-allergen	-0.12279	**0.2970** (Probable **NON-ANTIGEN**).
KSLFSEQNTK	0.840898	Non-Toxin	non-allergen	-0.04122	**0.1715** (Probable **NON-ANTIGEN**).
ALNQAINNA	0.703354	Non-Toxin	allergen	0.03548	**0.2220** (Probable **NON-ANTIGEN**).

Table 8 CTLs prediction of OMP18

peptide	score	Toxicity	Allergenticity	Immunogenicity score	overall Prediction for the Protective Antigen
NTDEFGSSEY	0.992247	Non-Toxin	non-allergen	0.0181	**0.8977** (Probable **ANTIGEN**
TLDEIVQKA	0.86191	Non-Toxin	allergen	0.05265	0.3022 (Probable NON-ANTIGEN)
SVKNALVIK	0.850218	Non-Toxin	allergen	0.06375	**0.8885** (Probable **ANTIGEN**).
KENHMQVLL	0.701687	Non-Toxin	non-allergen	-0.22123	0.4324 (Probable NON-ANTIGEN)

KAKENHMQV	0.212238	Non-Toxin	non-allergen	-0.16098	**1.5358** (Probable **ANTIGEN**).

Table 9 CTLs prediction of UreB

peptide	score	Toxicity	Allergenticity	Immunogenicity score	overall Prediction for the Protective Antigen
ATNPGPPAK	0.95027	Toxin	non-allergen	0.0228	-0.1241 (Probable NON-ANTIGEN).
GIDTHIHFI	0.680631	Non-Toxin	allergen	0.29874	**1.4037** (Probable **ANTIGEN**
FASGVTTMI	0.675023	Non-Toxin	allergen	-0.0127	0.0295 (Probable NON-ANTIGEN
HALDVADKY	0.971193	Non-Toxin	allergen	-0.01153	0.1865 (Probable NON-ANTIGEN).
NPNFDGVGF	0.87227	Non-Toxin	allergen	0.22384	**1.0496** (Probable **ANTIGEN**).

Table 10 CTLs prediction of IceA2

peptide	score	Toxicity	Allergenticity	Immunogenicity score	overall Prediction for the Protective Antigen
YENGIHKRTY	0.958286	Non-Toxin	non-allergen	0.05646	**1.3094** (Probable **ANTIGEN**).
GVVAVTTSK	0.828025	Non-Toxin	allergen	0.06561	0.1782 (Probable NON-ANTIGEN)
QVSGGVVAV	0.714036	Non-Toxin	non-allergen	0.10996	**1.2094** (Probable **ANTIGEN**
TTSKGKVEEY	0.695095	Non-Toxin	non-allergen	-0.26584	**1.2054** (Probable **ANTIGEN**).
VVIKVVNGK	0.784178	Non-Toxin	allergen	-0.0804	

					0.8171 (Probable **ANTIGEN**).

The accepted epitopes were the epitopes which met the required conditions and whose binding score was higher than 0.9. The accepted epitopes are shown in the table (11) below:

Table 11 The acceptable CTLs epitope

	Epitope	Score	Toxicity	Allergenicity	Immunogenicity Score	**overall Prediction for the Protective Antigen**
VacA	KVWRIQAGK	0.910978	Non-Toxin	non-allergen	0.19736	**0.6791** (Probable **ANTIGEN**).
	SLYDGATLNL	0.904618	Non-Toxin	non-allergen	0.11562	**0.8078** (Probable **ANTIGEN**).
Omp6	IPTINTNYY	0.955319	Non-Toxin	non-allergen	0.16914	**0.6062** (Probable **ANTIGEN**
Omp18	NTDEFGSSEY	0.992247	Non-Toxin	non-allergen	0.0181	**0.8977** (Probable **ANTIGEN**
IceA2	YENGIHKRTY	0.958286	Non-Toxin	non-allergen	0.05646	**1.3094** (Probable **ANTIGEN**).

6-2-2- Helper T Lymphocyte (HTL) epitope prediction :

Most prediction algorithms for human MHC class II have focused on HLA molecules in the DR locus, leaving a gap in our understanding of the equally important HLA DP and DQ molecules. This is because DP and DQ molecules are more challenging to study experimentally.

The most common alleles in the Syrian population for the DRB1 locus were DRB1*11 (26.4%), DRB1*04 (14%), and DRB1*07 (12%). However, the most frequent alleles for the DQB1 locus were DQB1*03 (40.9%) and DQB1*05 (25.1%) [36] .

BA predictions evaluate the ability of a peptide to bind an MHC II molecule based on Prediction method: netmhciipan_ba 4.1 | Low ic50 = good binders.

Table 12 HTLs prediction of vacA

Table 13 HTLs

Epitope	IC50	overall Prediction for the Protective Antigen	Toxicity	Allergenicity	IFN-Yprediction
QRFASLESAAEVLYQ	15.05	**0.0462** (Probable **NON-ANTIGEN**)	Non-Toxin	non-allergen	0.013334961
AIVGGIATGTAVGTV	12.01	**0.5805** (Probable **ANTIGEN**) .	Non-Toxin	non-allergen	0.13514127
IVGGIATGTAVGTVS	13.45	**0.6500** (Probable **ANTIGEN**) .	Non-Toxin	non-allergen	0.24199452
IPAIVGGIATGTAVG	14.41	**0.5658** (Probable **ANTIGEN**) .	Non-Toxin	non-allergen	0.007911662 9
PAIVGGIATGTAVGT	13.02	**0.7967** (Probable **ANTIGEN**	Non-Toxin	non-allergen	0.17282877

prediction of hpaA

Epitope	IC50	overall Prediction for the Protective Antigen	Toxicity	Allerrgenicity	IFN-Yprediction
ENKFKNQTTLKVEQI	26.69	**0.6947** (Probable **ANTIGEN**) .	Non-Toxin	non-allergen	-0.242059955
LDEKILLLRPAFQYR	45.97	**1.0706** (Probable **ANTIGEN**) .	Non-Toxin	Allergen	0.45982546
NKFKNQTTLKVEQIL	28.82	**0.3608** (Probable **NON-ANTIGEN**	Non-Toxin	non-allergen	-0.50060577

YENKFKNQTTLKVEQ	32.08	**0.7354** (Probable **ANTIGEN**).	Non-Toxin	non-allergen	-0.30903841
EYENKFKNQTTLKVE	38.32	**0.9156** (Probable **ANTIGEN**).	Non-Toxin	non-allergen	-0.27930636

Table 14 HTLs prediction of Omp6

Epitope	**IC50**	**overall Prediction for the Protective Antigen**	**Toxicity**	**Allerrgenicity**	**IFN-Yprediction**
FQFLFNTGLRLQGIH	23.34	**1.0795** (Probable **ANTIGEN**).	Non-Toxin	non-allergen	-0.11848833
NTFNAITSMIDSAKK	18.69	**-0.4028** (Probable **NON-ANTIGEN**).	Non-Toxin	Allergen	-1.0903248
YIKSNFFNSASNVFT	11.49	**-0.0269** (Probable **NON-ANTIGEN**).	Non-Toxin	non-allergen	0.3574998
FQFLFNTGLRLQGIH	27.14	**1.0795** (Probable **ANTIGEN**).	Non-Toxin	non-allergen	-0.41253792
PTINTNYYSFMGAKL	33.68	**0.6186** (Probable **ANTIGEN**).	Non-Toxin	non-allergen	-0.34626002

Table 15 HTLs prediction of OIPA

Epitope	**IC50**	**overall Prediction for the Protective Antigen**	**Toxicity**	**Allerrgenicity**	**IFN-Yprediction**
RNGFYLGLNFAEGSY	35.85	**0.5969** (Probable **ANTIGEN**).	Non-Toxin	non-allergen	0.033493408
FGYKYFLGKKRIIGF	10.56	**-0.4696** (Probable **NON-ANTIGEN**).	Non-Toxin	non-allergen	-0.32154094

SFGYKYFLGKKRIIG HLA-DRB1*11:01	15.84	**-0.3476** (Probable **NON-ANTIGEN**).	Non-Toxin	non-allergen	-0.31271631
ERNGFYLGLNFAEGS HLA-DRB1*04:05	40.44	**0.7028** (Probable **ANTIGEN**).	Non-Toxin	non-allergen	-0.090006614
AKKSLFSEQNTKAIR HLA-DRB1*04:01	37.28	**0.2770** (Probable **NON-ANTIGEN**).	Non-Toxin	non-allergen	-0.358012

Table 16: HTLs prediction of Omp18

Epitope	**IC50**	**overall Prediction for the Protective Antigen**	**Toxicity**	**Allerrgenicity**	**IFN-Yprediction**
LGVKRTLSVKNALVI	34.72	**0.5363** (Probable **ANTIGEN**	Non-Toxin	non-allergen	0.49550167
GVKRTLSVKNALVIK	35.92	**0.7662** (Probable **ANTIGEN**).	Non-Toxin	non-allergen	0.20219355
ALGVKRTLSVKNALV	40.32	**0.5440** (Probable **ANTIGEN**).	Non-Toxin	non-allergen	0.43746888
KPAIESGTIIASIYF	12.82	**0.5426** (Probable **ANTIGEN**	Non-Toxin	non-allergen	-0.61508315
EKPAIESGTIIASIY	13.81	**0.1486** (Probable **NON-ANTIGEN**).	Non-Toxin	non-allergen	-0.50612797

Table 17: HTLs prediction of UreB

Epitope	**IC50**	**overall Prediction for the Protective Antigen**	**Toxicity**	**Allerrgenicity**	**IFN-Yprediction**
VEDTKAAIAGRTMHT	26.48	**0.8617** (Probable **ANTIGEN**)	Non-Toxin	Allergen	-0.46348677

RVEDTKAAIAGRTMH	30.18	**1.1256** (Probable **ANTIGEN**	Non-Toxin	Allergen	-0.49363784
EDTKAAIAGRTMHTR	31.95	**1.2131** (Probable **ANTIGEN**).	Non-Toxin	Allergen	-0.55168684
GRVEDTKAAIAGRTM	36.21	**1.0418** (Probable **ANTIGEN**).	Non-Toxin	Allergen	-0.035228983
DTKAAIAGRTMHTRH	39.34	**1.1254** (Probable **ANTIGEN**).	Non-Toxin	Allergen	-0.46169504

Table 18 HTLs prediction of IceA

Epitope	**IC50**	**overall Prediction for the Protective Antigen**	**Toxicity**	**Allerrgenicity**	**IFN-Yprediction**
VQVSGGVVAVTTSKG DQB1*03:01	9.10	**0.8828** (Probable **ANTIGEN**).	Non-Toxin	Allergen	-0.65563287
NVQVSGGVVAVTTSK DQB1*03:01	9.47	**1.0697** (Probable **ANTIGEN**)	Non-Toxin	Allergen	-0.73094903
INVQVSGGVVAVTTS DQB1*03:01	9.68	**1.2203** (Probable **ANTIGEN**)	Non-Toxin	Allergen	-0.90985254
AINVQVSGGVVAVTT DQB1*03:01	10.05	**1.2182** (Probable **ANTIGEN**).	Non-Toxin	Allergen	-0.57634869
QVSGGVVAVTTSKGK DQB1*03:01	10.22	**1.1527** (Probable **ANTIGEN**).	Non-Toxin	Allergen	-0.53236651

The accepted epitopes were entered into IL4pred which is an In-Silico platform for designing and discovering of Interleukin-4 inducing peptides. IL4pred allows users to predict whether their peptide has the ability to induce IL4 or not IL4 inducing prediction results are shown in the Table (19).

Table 19 The Acceptable HTLs epitope

	Epitope	IC50	overall Prediction for the Protective Antigen	Toxicity	Allerrgenicity	IFN-Yprediction	IL4 inducing prediction
VagA	AIVGGIATGTAVGTV	12.01	**0.5805** (Probable **ANTIGEN**).	Non-Toxin	non-allergen	0.13514127	Non-IL4-inducer
	IVGGIATGTAVGTVS	13.45	**0.6500** (Probable **ANTIGEN**).	Non-Toxin	non-allergen	0.24199452	Non-IL4-inducer
	IPAIVGGIATGTAVG	14.41	**0.5658** (Probable **ANTIGEN**).	Non-Toxin	non-allergen	0.0079116629	Non-IL4-inducer
	PAIVGGIATGTAVGT	13.02	**0.7967** (Probable **ANTIGEN**	Non-Toxin	non-allergen	0.17282877	IL4-inducer
OIPA	RNGFYLGLNFAEGSY	35.85	**0.5969** (Probable **ANTIGEN**).	Non-Toxin	non-allergen	0.033493408	Non-IL4-inducer
Omp18	LGVKRTLSVKNALVI	34.72	**0.5363** (Probable **ANTIGEN**	Non-Toxin	non-allergen	0.49550167	Non-IL4-inducer
	GVKRTLSVKNALVIK	35.92	**0.7662** (Probable **ANTIGEN**).	Non-Toxin	non-allergen	0.20219355	Non-IL4-inducer
	ALGVKRTLSVKNALV	40.32	**0.5440** (Probable **ANTIGEN**).	Non-Toxin	non-allergen	0.43746888	Non-IL4-inducer

6-2-3- Linear B Lymphocyte (LBL) epitope prediction:

The predicted B cell epitopes are ranked according to their score obtained by trained recurrent neural networks. Higher score of the peptide means the higher probability to be an epitope. All the peptides shown here are above the threshold value chosen which is 0.51

Table 20 LBL Prediction vacA

Peptide	Score	overall Prediction for the Protective Antigen	Toxicity	Allerrgenicity
QGTINYLVRGGKVATL	0.96	**0.3099** (Probable **NON-ANTIGEN**	Non-Toxin	non-allergen
HYWIKGGQWNKLEVDM	0.93	**0.5909** (Probable **ANTIGEN**).	Non-Toxin	non-allergen
GLNIIAPPEGGYKDKP	0.92	**0.6386** (Probable **ANTIGEN**).	Non-Toxin	non-allergen
EVLYQFAPKYEKPTNV	0.92	**0.8571** (Probable **ANTIGEN**	Non-Toxin	non-allergen
YYLGNSTPTENGGNTT	0.91	**1.2027** (Probable **ANTIGEN**).	Non-Toxin	non-allergen

Table 21 LBL Prediction of hpaA

Peptide	Score	overall Prediction for the Protective Antigen	Toxicity	Allerrgenicity
KRTIQKKSEPGLLFST	0.91	**0.4827** (Probable **NON-ANTIGEN**).	Non-Toxin	Allergen
AGMENTSHELICACTE	0.84	**0.3706** (Probable **NON-ANTIGEN**).	Non-Toxin	non-allergen
AMNGEIVLRPDPKRTI	0.82	**0.7841** (Probable **ANTIGEN**	Non-Toxin	non-allergen
YRDNIAKEYENKFKNQ	0.82	**0.5595** (Probable **ANTIGEN**	Non-Toxin	non-allergen
PESVNYHPASEKVQAL	0.81	**0.5353** (Probable **ANTIGEN**	Non-Toxin	non-allergen

Table 22 LBL Prediction of Omp6

Peptide	Score	overall Prediction for the Protective Antigen	Toxicity	Allerrgenicity
ERMEMRANEPRTEINS	0.95	**1.3947** (Probable **ANTIGEN**).	Non-Toxin	Allergen
SSVINDTISYLKGDNL	0.93	**0.1806** (Probable **NON-ANTIGEN**).	Non-Toxin	non-allergen
EARRTDPESPNQQSTF	0.91	**0.7793** (Probable **ANTIGEN**).	Non-Toxin	Allergen
NGKPWGINASGNACNI	0.89	**0.8800** (Probable **ANTIGEN**).	Non-Toxin	non-allergen
MWQVIASNLANNSLST	0.89	**0.6718** (Probable **ANTIGEN**).	Non-Toxin	non-allergen

Table 23 LBL Prediction of OIPA

Peptide	Score	overall Prediction for the Protective Antigen	Toxicity	Allerrgenicity
QGSIGEKASAQNALNQ	0.90	**1.0562** (Probable **ANTIGEN**).	Non-Toxin	Allergen
AGMENTSHELICACTE	0.84	**0.3706** (Probable **NON-ANTIGEN**).	Non-Toxin	non-allergen
LGLNFAEGSYIKGQGS	0.83	**1.0231** (Probable **ANTIGEN**)	Non-Toxin	non-allergen
FWLHAERNGFYLGLNF	0.81	**0.6714** (Probable **ANTIGEN**).	Non-Toxin	Allergen
TKAIRDAQNALNKVKD	0.77	**-0.1933** (Probable **NON-ANTIGEN**).	Non-Toxin	Allergen

Table 24 : LBL Prediction of Omp18

Peptide	Score	overall Prediction for the Protective Antigen	Toxicity	Allerrgenicity

SGTIIASIYFDFDKYE	0.96	**0.4639** (Probable **NON-ANTIGEN**).	Non-Toxin	Allergen
ERMEMRANEPRTEINM	0.93	**1.4382** (Probable **ANTIGEN**)	Non-Toxin	Allergen
AGDVSAKMVQTAPVTT	0.86	**1.1553** (Probable **ANTIGEN**).	Non-Toxin	non-allergen
IKTISFGETKPKCAQK	0.86	**0.6467** (Probable **ANTIGEN**	Non-Toxin	Allergen
RTEINMPSHELICACT	0.85	**0.1725** (Probable **NON-ANTIGEN**	Non-Toxin	non-allergen

Table 25 LBL Prediction UreB

Peptide	**Score**	**overall Prediction for the Protective Antigen**	**Toxicity**	**Allerrgenicity**
IPPQQTPTAFASGVTT	0.96	**0.3343** (Probable **NON-ANTIGEN**).	Non-Toxin	Allergen
AGGIDTHIHFIPPQQT	0.96	**1.0036** (Probable **ANTIGEN**	Non-Toxin	non-allergen
EEYATNPGPPAKGNTP	0.95	**0.3126** (Probable **NON-ANTIGEN**	Non-Toxin	Allergen
AGRVEDTKAAIAGRTM	0.93	**1.0067** (Probable **ANTIGEN**).	Non-Toxin	non-allergen
TTMIGGGTGPADGTNA	0.92	**1.0996** (Probable **ANTIGEN**	Non-Toxin	non-allergen

Table 26 LBL Prediction of IceA2

Peptide	**Score**	**overall Prediction for the Protective Antigen**	**Toxicity**	**Allerrgenicity**
HELICACTERPYLRIX	0.88	**1.0755** (Probable **ANTIGEN**).	Non-Toxin	Allergen
YENGIHKRTYGSNAIN	0.83	**0.9606** (Probable **ANTIGEN**).	Non-Toxin	non-allergen

GNICEAPESVMAVVIK	0.82	**0.1737** (Probable **NON-ANTIGEN**).	Non-Toxin	non-allergen
KGKVEEYKNGIYKRTY	0.78	**0.6175** (Probable **ANTIGEN**).	Non-Toxin	non-allergen
SNAINVQVSGGVVAVT	0.58	**1.2286** (Probable **ANTIGEN**).	Non-Toxin	Allergen

Table 27 : Acceptable LBLs epitope

	Peptide	**Score**	**overall Prediction for the Protective Antigen**	**Toxicity**	**Allerrgenicity**
VacA	HYWIKGGQWNKLEVDM	0.93	**0.5909** (Probable **ANTIGEN**).	Non-Toxin	non-allergen
	GLNIIAPPEGGYKDKP	0.92	**0.6386** (Probable **ANTIGEN**).	Non-Toxin	non-allergen
	EVLYQFAPKYEKPTNV	0.92	**0.8571** (Probable **ANTIGEN**	Non-Toxin	non-allergen
	YYLGNSTPTENGGNTT	0.91	**1.2027** (Probable **ANTIGEN**).	Non-Toxin	non-allergen
HpaA	AMNGEIVLRPDPKRTI	0.82	**0.7841** (Probable **ANTIGEN**	Non-Toxin	non-allergen
	YRDNIAKEYENKFKNQ	0.82	**0.5595** (Probable **ANTIGEN**	Non-Toxin	non-allergen
	PESVNYHPASEKVQAL	0.81	**0.5353** (Probable **ANTIGEN**	Non-Toxin	non-allergen
Omp6	NGKPWGINASGNACNI	0.89	**0.8800** (Probable **ANTIGEN**).	Non-Toxin	non-allergen
	MWQVIASNLANNSLST	0.89	**0.6718** (Probable **ANTIGEN**).	Non-Toxin	non-allergen
OIPA	LGLNFAEGSYIKGQGS	0.83	**1.0231** (Probable **ANTIGEN**)	Non-Toxin	non-allergen

Omp18	AGDVSAKMVQTAPVTT	0.86	**1.1553** (Probable **ANTIGEN**).	Non-Toxin	non-allergen
UreB	AGGIDTHIHFIPPQQT	0.96	**1.0036** (Probable **ANTIGEN**	Non-Toxin	non-allergen
	AGRVEDTKAAIAGRTM	0.93	**1.0067** (Probable **ANTIGEN**).	Non-Toxin	non-allergen
	TTMIGGGTGPADGTNA	0.92	**1.0996** (Probable **ANTIGEN**	Non-Toxin	non-allergen
IceA2	YENGIHKRTYGSNAIN	0.83	**0.9606** (Probable **ANTIGEN**).	Non-Toxin	non-allergen
	KGKVEEYKNGIYKRTY	0.78	**0.6175** (Probable **ANTIGEN**).	Non-Toxin	non-allergen

Table 28 Final chosen (CTL,HTL,LBL) epitope

CTLS

VacA		Omp6	Omp18	IceA2
KVWRIQAGK	SLYDGATLNL	IPTINTNYY	NTDEFGSSEY	YENGIHKRTY
0.910978	0.904618	0.955319	0.992247	0.958286
HLA-A*03:01	HLA-A*02:03	HLA-B*35:01	HLA-A*01:01	HLA-B*44:03

HTLs

VacA			Omp18			OIPA
AIVGGIATG TAVGTV	IVGGIATGT AVGTVS	IPAIVGGIA TGTAVG	LGVKRTLS VKNALVI	GVKRTLSV KNALVIK	ALGVKRTLS VKNALV	RNGFYLGL NFAEGSY
12.01	13.45	14.41	34.72	35.92	40.32	35.85
HLA-DQB1*03:01	HLA-DQB1*03:01	HLA-DQB1*03:01	HLA-DRB1*07:01	HLA-DRB1*07:01	HLA-DRB1*07:01	HLA-DRB1*04:05

LBLs

VacA	HYWIKGGQWNKLEVDM	0.93
	GLNIIAPPEGGYKDKP	0.92
	EVLYQFAPKYEKPTNV	0.92
	YYLGNSTPTENGGNTT	0.91
HpaA	AMNGEIVLRPDPKRTI	0.82
	YRDNIAKEYENKFKNQ	0.82
	PESVNYHPASEKVQAL	0.81
Omp 6	NGKPWGINASGNACNI	0.89
	MWQVIASNLANNSLST	0.89
OIPA	LGLNFAEGSYIKGQGS	0.83
Omp 18	AGDVSAKMVQTAPVTT	0.86
UreB	AGGIDTHIHFIPPQQT	0.96
	AGRVEDTKAAIAGRTM	0.93
	TTMIGGGTGPADGTNA	0.92
IceA2	YENGIHKRTYGSNAIN	0.83
	KGKVEEYKNGIYKRTY	0.78

Note : (5 CTL , 7 HTL, 8 LBL) will be construct the vaccine

6-3-Population Coverage Calculation :

The test was conducted using the population coverage tool in the IEDB site and due to the absence of the Syrian Arab Republic among the countries, the surrounding countries were used. The region of southwest Asia where Syria is located was selected,

specifying the most frequent alleles of MHC-I and MHC-II which were previously mentioned. The results are as in the next picture and the percentage was 87.09%.

Population Coverage Calculation Result

population/area	Class combined		
	coverage[a]	average_hit[b]	pc90[c]
Southwest Asia	87.09%	3.69	0.77
Average	**87.09**	**3.69**	**0.77**
Standard deviation	**0.0**	**0.0**	**0.0**

[a] projected population coverage
[b] average number of epitope hits / HLA combinations recognized by the population
[c] minimum number of epitope hits / HLA combinations recognized by 90% of the population

Population: Southwest Asia

MHC class	Coverage	Average hit	PC90
combined	87.09%	3.69	0.77

Figure 1 Southwest Asia population coverage calculation results.

6-4 Multi-Epitope Vaccine Construction :

A total 5 CTL ,7 HTL , 8 LBL epitopes based on their antigenic nature ,binding e and energy , The non- allergenic property were linked together by using AAY, GPGPG and KK linkers to construct the final vaccine ,linkers Are crucial for providing flexibility to amino acid residues, ensuring prolonged conformation or protein folding. Idudn the N- terminal of the vaccine, an adjuvant was added to increase the immunogenicity of the vaccine. Moreover , an EAAAK linker was used to join the adjuvant to the CTL epitopes.The final vaccine construct comprised 456 amino acid residues.

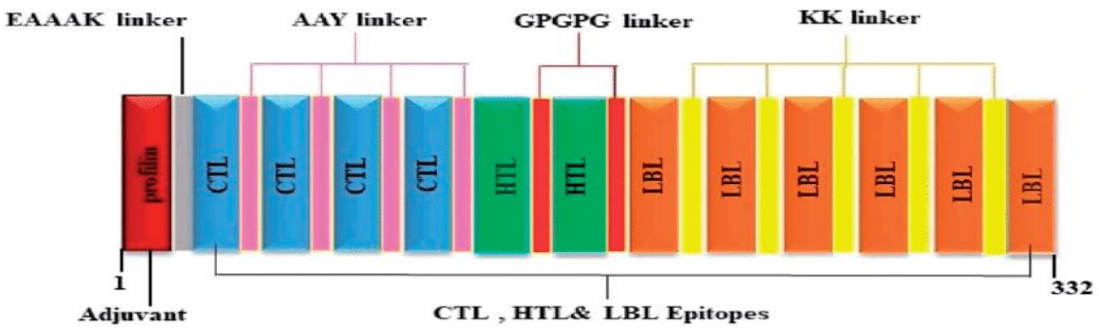

Figure 2 The structural arrangement of B and T cell epitopes along with linkers and adjuvant for the final multi-epitope vaccine construct

6-4-1- Primary structure :

LCAHGAPQSITELCSEYRNTQIYTINDKILSYTESMAGKREMVIITFKSGATFQVEVPGSQHIDSQKKAIER
MKDTLRITYLTETKIDKLCVWNNKTPNSIAAISMEAAAKKVWRIQAGKAAYSLYDGATLNLAAYSLYDGA
TLNLAAYNTDEFGSSEYAAYYENGIHKRTYGPGPGAIVGGIATGTAVGTVGPGPGIVGGIATGTAVGTVS
GPGPGIPAIVGGIATGTAVGGPGPGLGVKRTLSVKNALVIGPGPGGVKRTLSVKNALVIKGPGPGALGVK
RTLSVKNALVGPGPGRNGFYLGLNFAEGSYKKHYWIKGGQWNKLEVDMKKGLNIIAPPEGGYKDKPKK
AMNGEIVLRPDPKRTIKKYRDNIAKEYENKFKNQKKNGKPWGINASGNACNIKKLGLNFAEGSYIKGQG
SKKAGDVSAKMVQTAPVTTKKAGGIDTHIHFIPPQQT

The structure above shows that the chosen epitopes are connected together with suitable linkers and adjuvant by N-terminal position.

6-5-Specifications Prediction of the Vaccine Construct :

ProtParam results described the physiological nature of the designed vaccine
Number of amino acids: 456,**Molecular weight:** 48053.26,**Theoretical pI:** 9.76 indicating its basic nature ,**Instability index:** is computed to be 24.82,This classifies the protein as stable,
Aliphatic index: 79.85
Grand average of hydropathicity (GRAVY -0.307 (hydrophilic).
In addition , the immunological potency of the vaccine is determined by assessing the antigenicity score . The vaccine is antigenic,with a score of **0.7752**. It is neither allergen nor toxic .
The predicted solubility by Protein- Sol is 0.643which indicates is predicted to have a higher solubility than the average soluble E.coli protein from the experimental solubility dataset [48] ,The protein is predicted to be SOLUBLE with probability 0.974503 based on Solpro server. All shown in table below:

Table 29 Physicochemical properties of the vaccine

Number of amino acids	456
Molecular weight	48053.26
Theoretical Pi	9.76

Total number of negatively charged residues (Asp + Glu)	31
Total number of positively charged residues (Arg + Lys)	60
Formula	C2148H3446N596O630S11
Estimated half-life	The estimated half-life is: 5.5 hours (mammalian reticulocytes, in vitro). 3 min (yeast, in vivo). 2 min (Escherichia coli, in vivo).
Instability index	24.82 (stable)
Aliphatic index	79.85
Grand average of hydropathicity (GRAVY)	-0.307 (hydrophilic)
Antigenicity (vaxijen)	**0.7752** (Probable **ANTIGEN**).
Allergenicity (Allertop)	Non-Allergen
Toxicity (Toxinpred)	Non- Toxin

6-6-Secondary Structure:

Here the SOPMA tool has been used to predict a-helix, B-sheets and coils percentages and the structure consists of 6.80% a-helix, 29.82%extended strands and 63.38% random coils. Results shown in the table below. Table (30):

Table 30 SOMPA results of vaccine's second structure

Feature	AA number	Percentage
a-helix	31	6.80%
extended strands	136	29.82%
random coils	289	63.38%

6-7- Tertiary structure, refinement and validation :

The phyre2 server was used in the homology for both candidates using intensive modeling mode.
Ramachandran plots was created using PROCHEC server. We used the GalaxRefine Server to identify and correct potential errors in the predicted 3D structure .

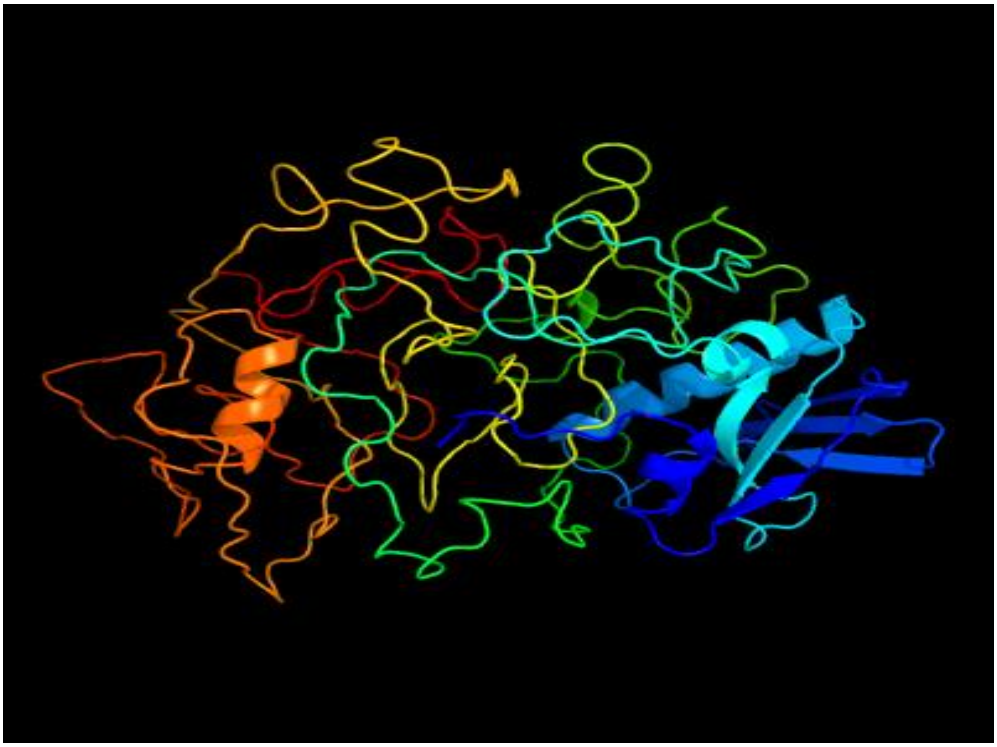

Figure 3 3D model of the vaccine (a) .

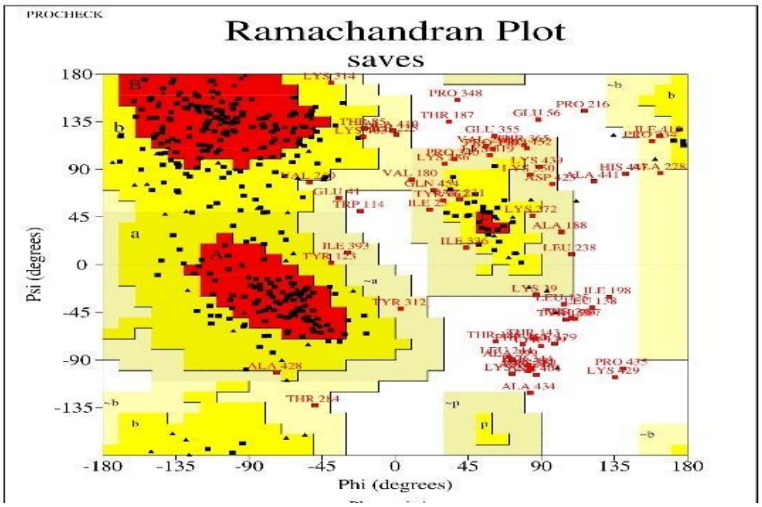

Figure 4 Ramachandran plot for the vaccine modeled.(b).

(a) 3D model of the vaccine protein modeled by the Phyre2 server. 28% of residues modeled at >90%confidence with the confidence of the model. (b) Ramachandran plot for the vaccine modeled protein shows: 54.0% of residues in most favored regions, 30.0% residues in additional allowed regions, 8.5% residues in generously allowed regions and 7.4% residues in disallowed regions, which indicates the high low of the model.

The analysis results indicated that the structure requires refinement and fixing to enhance its quality. Utilizing GalaxyRefine for this purpose can address all the issues, as this platform can generate the best 5 versions of the previous structurel(model -1), Subsequent validation of the refined structure demonstrated its satisfactory quality and stability. The structure was then subjected to docking analysis for further evaluation.

Model 1: has been chosen among all others and a new Ramachandran plot has been generated to the refined model showing the following figure

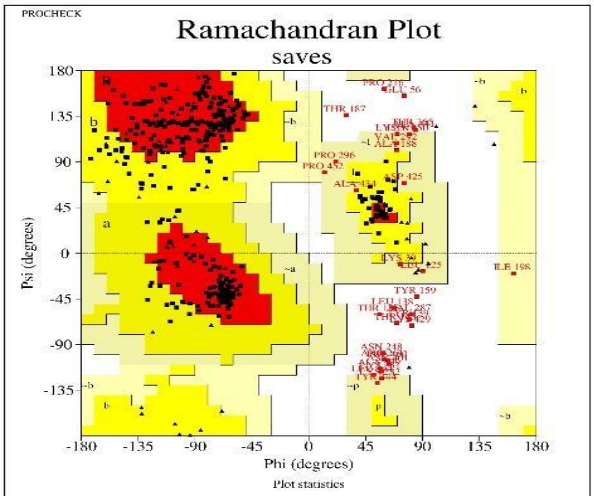

Figure 5 Ramachandran plot for the refined vaccine.(a)

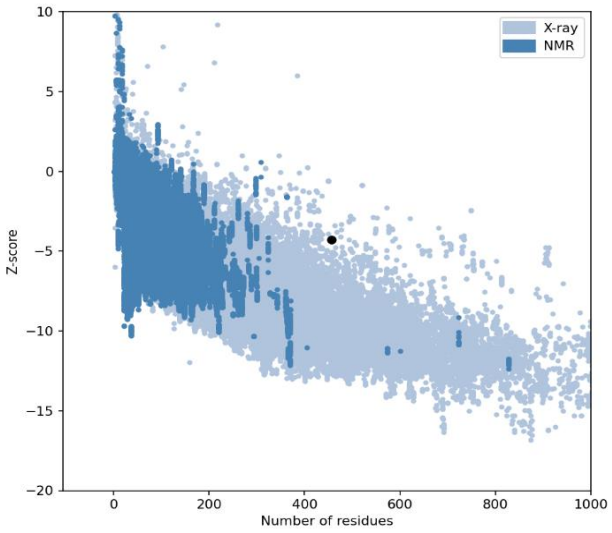

Figure 6 Z-score for the refined vaccine. (b)

a) Ramachandran plots for the refined vaccine shows:71.6 % of residues in most favored regions, 20.4 % residues in additional allowed regions, 3.0 % residues in generously allowed regions and 5.0% residues in disallowed regions.

b) The z-score of the vaccine candidate after refinement= - 4.3 is slightly in range of native protein conformation. It is depicted in a large black spot. z-Score plot consists of z-scores of all experimentally protein chains in PDB defined by NMR spectroscopy (dark blue) and X-ray crystallography (light blue).

6-8-Defining discontinuous B-cell epitope (conformationl):

The refined 3D vaccine model was subjected to B-cell epitopes prediction using Ellipro Server in order to predict conformational B-cell epitopes. Discontinuous B-cell epitopes were predicted with scores ranging from 0.755to 0.514 Amino acid residues, the number of residues, sequence location and their scores have been listed in the table below

Table 31 Discontinuos b.cells epitope prediction results

No	Residues	Number of residues	Score
1	_:N19, _:T24, _:I25, _:N26, _:D27, _:K28, _:I29, _:L30, _:S31, _:Y32, _:T33, _:E34, _:S35, _:M36, _:A37, _:G38, _:K39, _:R40, _:E41, _:M42, _:V43, _:I44, _:I45, _:T46, _:F47, _:K48, _:S49, _:G50, _:A51, _:T52, _:F53, _:Q54, _:V55, _:E56, _:V57, _:P58, _:G59, _:S60, _:Q61, _:H62, _:K74, _:L77, _:R78, _:I79, _:T80, _:Y81, _:L82, _:T83, _:E84, _:T85, _:K86, _:I87, _:D88, _:N94, _:N95, _:K96, _:T97, _:P98, _:N99, _:E107, _:A108, _:A109, _:A110, _:K111, _:K112	65	0.755
2	_:G239, _:V240, _:K241, _:R242, _:L244, _:S245, _:V246, _:K247, _:N248, _:V328, _:D329, _:M330, _:K331, _:K332, _:G333, _:L334, _:N335, _:I337, _:G342, _:G343, _:Y344, _:D346, _:K347, _:P348, _:K349, _:K350, _:A351, _:M352, _:N353, _:G354, _:E355, _:I356, _:V357, _:L358, _:R359, _:T365, _:I366, _:K367, _:K368, _:Y369, _:R370, _:D371, _:N372, _:I373, _:A374, _:K375, _:E376, _:Y377, _:E378, _:N379, _:K380, _:F381, _:K382, _:N383, _:Q384, _:K385, _:K386, _:N387, _:G388, _:K389, _:P390, _:G392, _:I393, _:A395, _:S396, _:G397, _:N398, _:A399, _:C400, _:N401, _:I402, _:K403, _:K404, _:L405, _:G406, _:L407, _:K440, _:A441, _:G442, _:G443, _:D445, _:T446, _:H447, _:I448, _:H449, _:I451, _:P452, _:P453, _:Q454, _:Q455, _:T456	91	0.671
3	_:R115, _:K120, _:A122, _:Y123, _:S124, _:L125, _:Y126, _:D127, _:G128, _:A129, _:T130, _:L131, _:L133, _:A134, _:S137, _:L138, _:Y139, _:D140, _:G141, _:A142, _:T143, _:N145, _:A147, _:A148, _:Y149, _:N150, _:T151, _:D152, _:E153,	97	0.657

	_:F154, _:G155, _:S156, _:S157, _:E158, _:Y159, _:A160, _:A161, _:Y162, _:E164, _:N165, _:G166, _:I167, _:H168, _:K169, _:R170, _:T171, _:Y172, _:G173, _:P174, _:G175, _:P176, _:G177, _:A178, _:I179, _:V180, _:G181, _:G182, _:I183, _:V192, _:G193, _:P194, _:G195, _:P196, _:G197, _:I198, _:V199, _:G200, _:G201, _:I202, _:A203, _:T204, _:G205, _:T206, _:A207, _:V208, _:G209, _:G215, _:P216, _:G217, _:I218, _:P219, _:S264, _:V265, _:K266, _:N267, _:A268, _:V270, _:I271, _:K272, _:G273, _:P274, _:G275, _:P276, _:G277, _:A278, _:L279, _:G280		
4	_:G257, _:G258, _:V259, _:K260, _:R261	5	0.514

6-9 Immune response Simulation :

C-ImmSim is composed by three modules or compartments which corresponds to three different and separate (bone marrow, thymus and lymph node) anatomical regions of a mammalian immune system.The immune simulations utilize the following input parameters: volume (10), HLA types (A0101, A0301, B3501, B4403,DRB1_0701, DRB1_0405), random seed (12345), number of steps (100), and a single injection (set to 1). All other parameters are assumed to be at their default values. The detailed representation of the immune response simulation is illustrated in Figure.

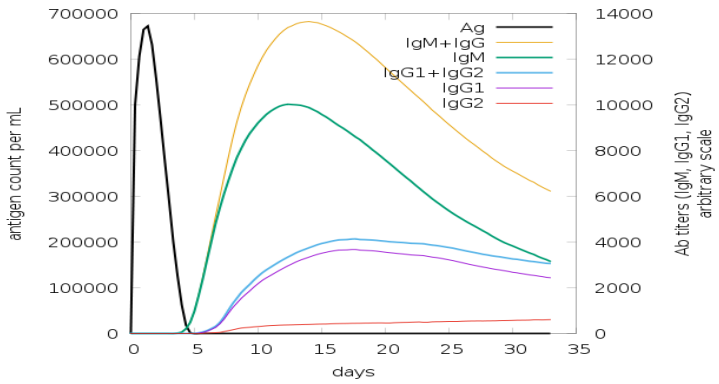

Figure 8

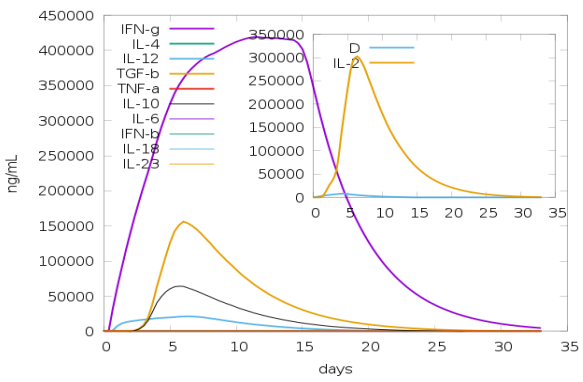

Figure 7

Figure 7: Concentration of cytokines and interleukins. Inset plot shows danger signal together with leukocyte growth.

Figure 8: The immunoglobulins and the immune complexes.

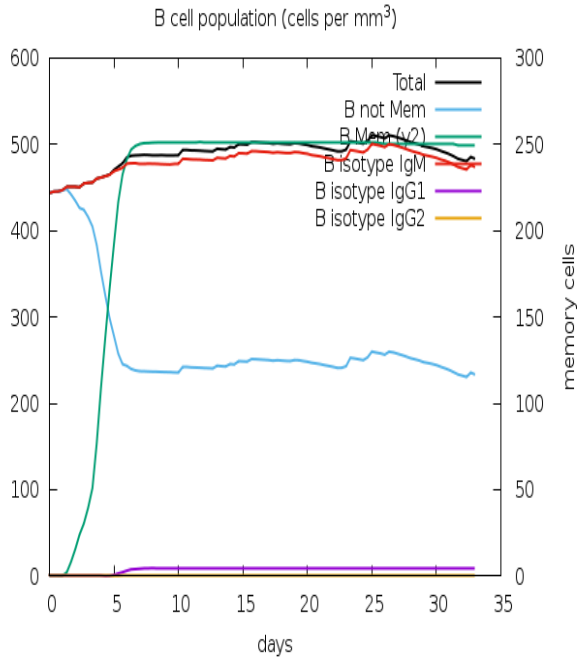

Figure 9 : shows B cell population. B lymphocytes: total count, memory cells, and subdivided in isotypes IgM, IgG1 and IgG2 for vaccine candidate.

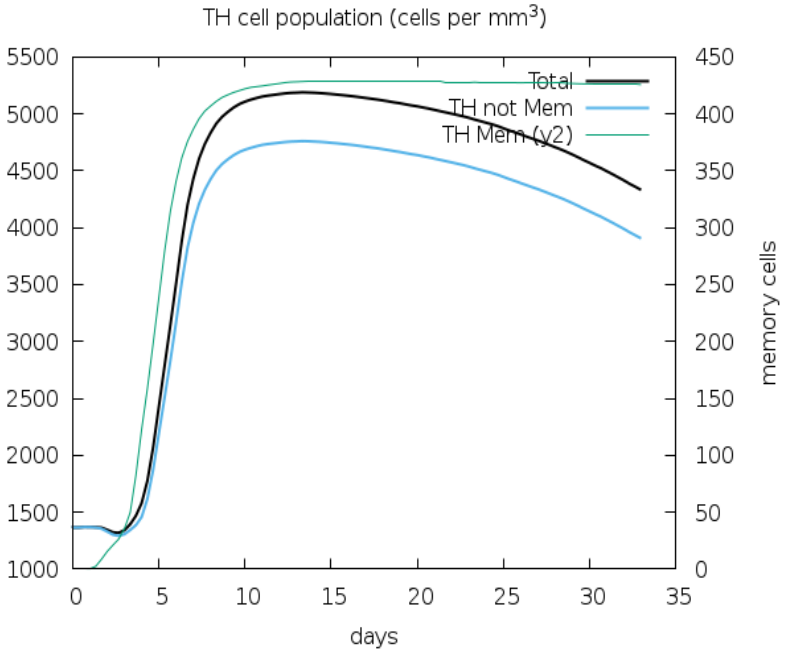

Figure 10 The cell population. CD4 T-helper lymphocytes count (total and memory counts) for vaccine candidate

The response showed high levels of IgG and IgM, followed by an increase in IgG1 along with a decrease in the antigen. Additionally, there was a strong interleukin and cytokine response, and a significant increase in IFN-gamma concentration. The injection also led to an increase

in B-cell and T-helper populations. These results indicate a successful immune response and clearance following subsequent encounters, confirming the immunogenicity of the vaccine candidate.

6-10 Molecular Docking :

A review of multiple medical articles on molecular docking of vaccine constructs with TLR-2/TLR-4 suggests an in-depth exploration of the interaction between vaccine components and these specific toll-like receptors[56][57]. TLR2 recognizes bacterial lipopolysaccharide (LPS), It plays a crucial role in regulating the immune response to H. pylori by activating NF-κB and inducing cytokine expression in various immune cells such as epithelial cells, monocytes/macrophages, dendritic cells, neutrophils, and B cells. The immune response mediated by TLR2 in response to H. pylori-derived components may significantly influence the outcome of the infection, determining whether the bacteria are eliminated, persist, or lead to pathological reactions. Consequently, TLR2 could be an intriguing therapeutic target for the treatment of H. pylori-related diseases[58].on the other hand, Toll-like receptor 4 (TLR4) plays a critical role in the immune response to Helicobacter pylori infection. H. pylori is a Gram-negative bacterium that possesses lipopolysaccharide (LPS) on its outer membrane, which is a potent activator of TLR4. When TLR4 recognizes the LPS of H. pylori, it triggers a signaling cascade that leads to the activation of pro-inflammatory responses and the production of cytokines[59].

The 3D structure of the receptor (TLR2) was modeled by the phyre2 server after the amino acid sequence was retrieved from the Uniprot (Uniprot ID:O60603).The modeled structure is shown

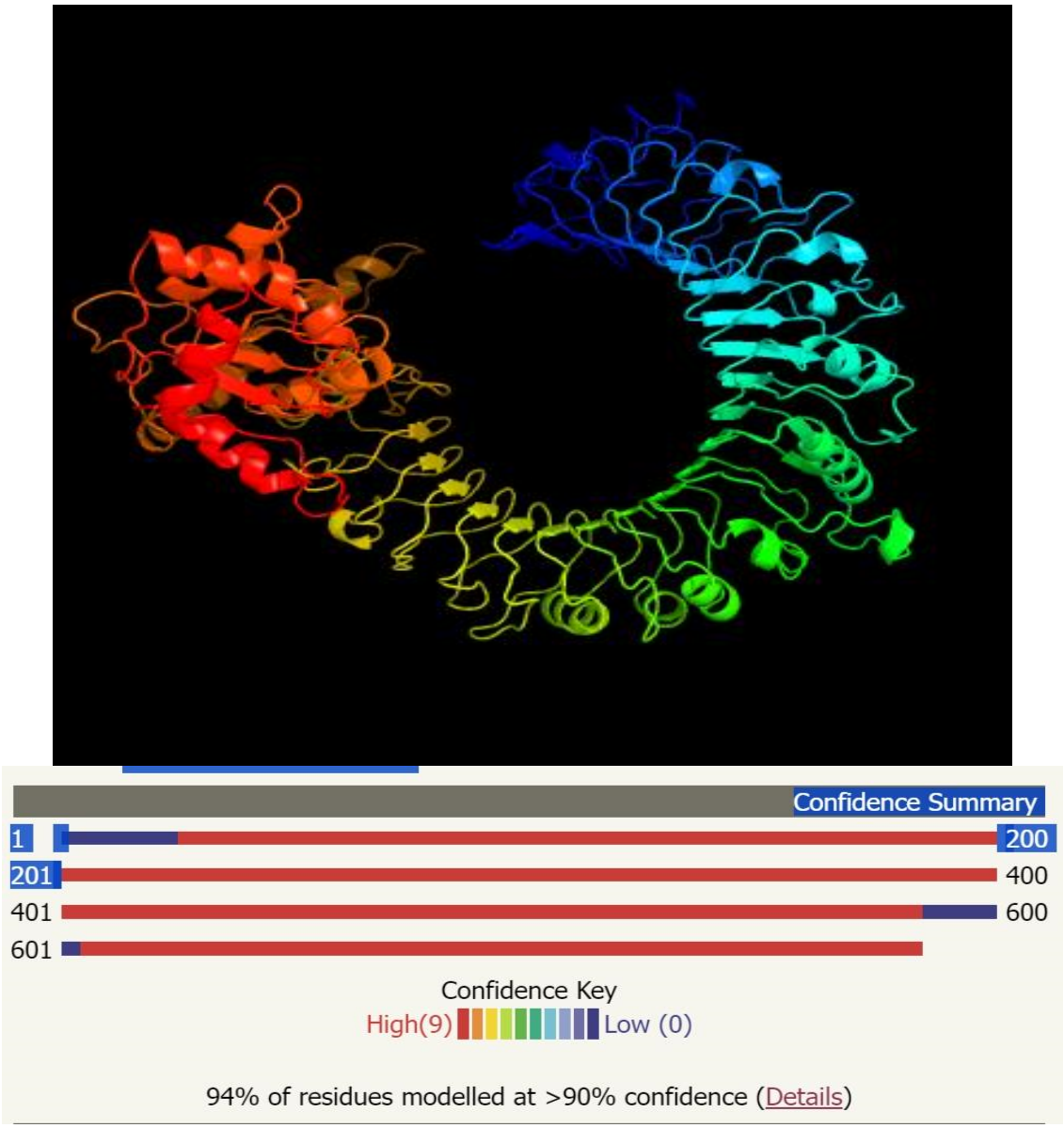

Figure 11 : 3D model of the TLR2 modeled

3D model of the TLR2 modeled by the Phyre2 server.94% of residues modeled at>90% confidence. The confidence of the generated model shows the high confidence of the model.

ClusPro 2.0 was used to initiate docking for TLR2 as a receptor with the ligand (vaccine candidate), which represents the final vaccine model.

After running the docking process, 10 samples were generated at site. The sample with the lowest energy and the highest membership was selected. Subsequently, the selected model (Model 0) with the lowest energy of -1243.8 .

Then Validation by chimera to make the final edited model.

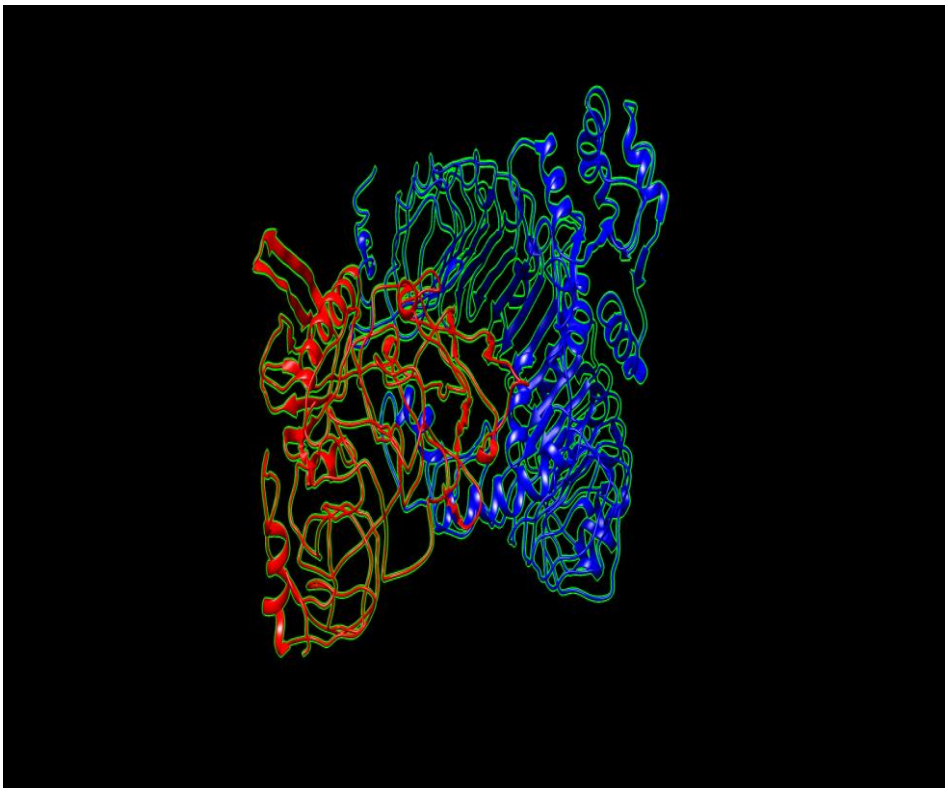

Figure 12 Docked model between TLR2 and ligand

Molecular docking with TLR4 as receptor :

The 3D structure of the receptor (TLR4) was modeled by the phyre2 server after the amino acid sequence was retrieved from the Uniprot (Uniprot ID:O00206).The modeled structure is shown.

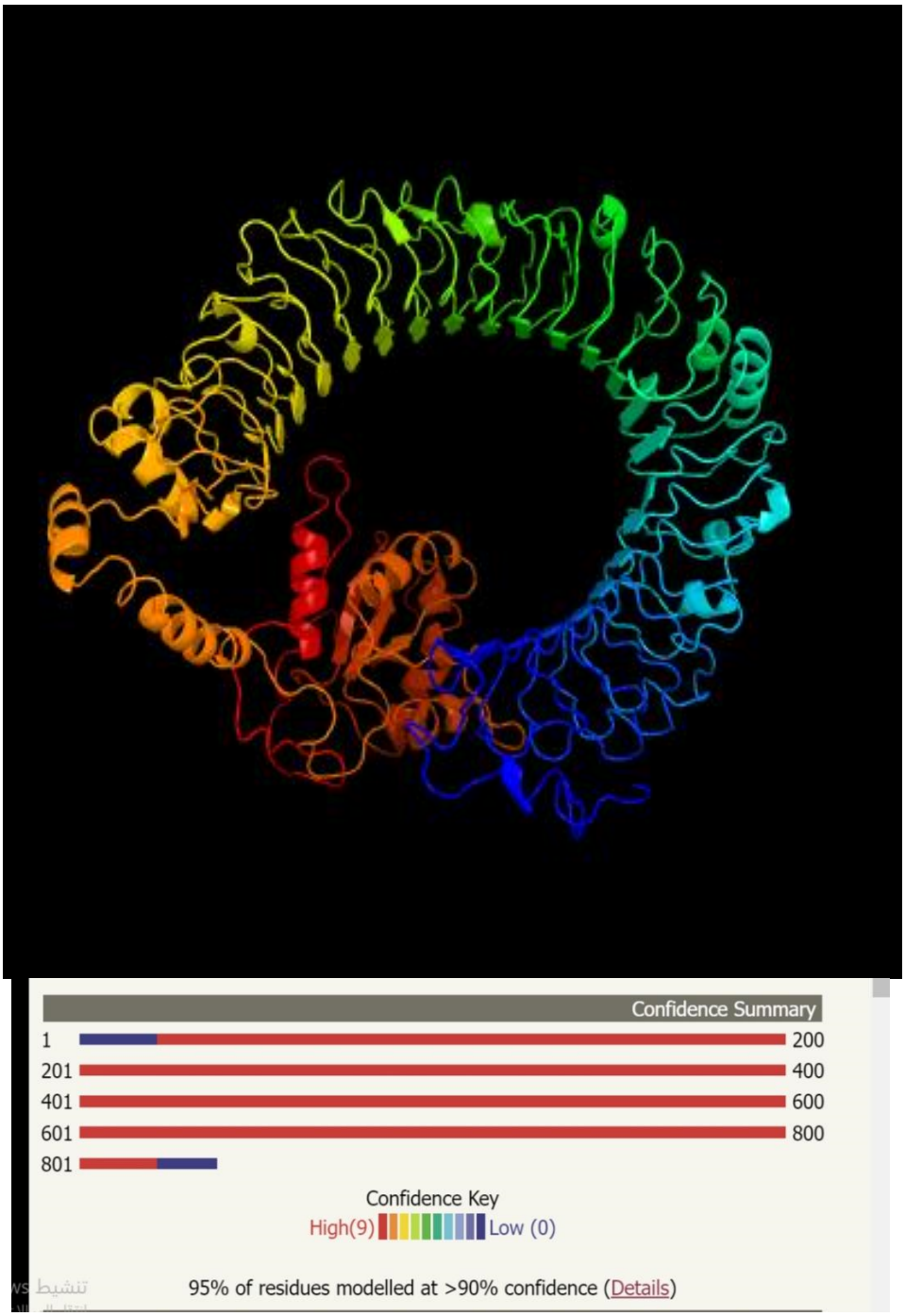

Figure 13 3D model of the TLR4 modeled

3D model of the TLR4 modeled by the Phyre2 server. 95% of residues modeled at >90% confidence. The confidence of the generated model shows the high confidence of the model.

TLR-4 as a receptor enters with the vaccine candidate as ligand to clusPro to give many models. The best model with lowest energy of -1297.1 .

Then Validation by chimera to make the final edited model.

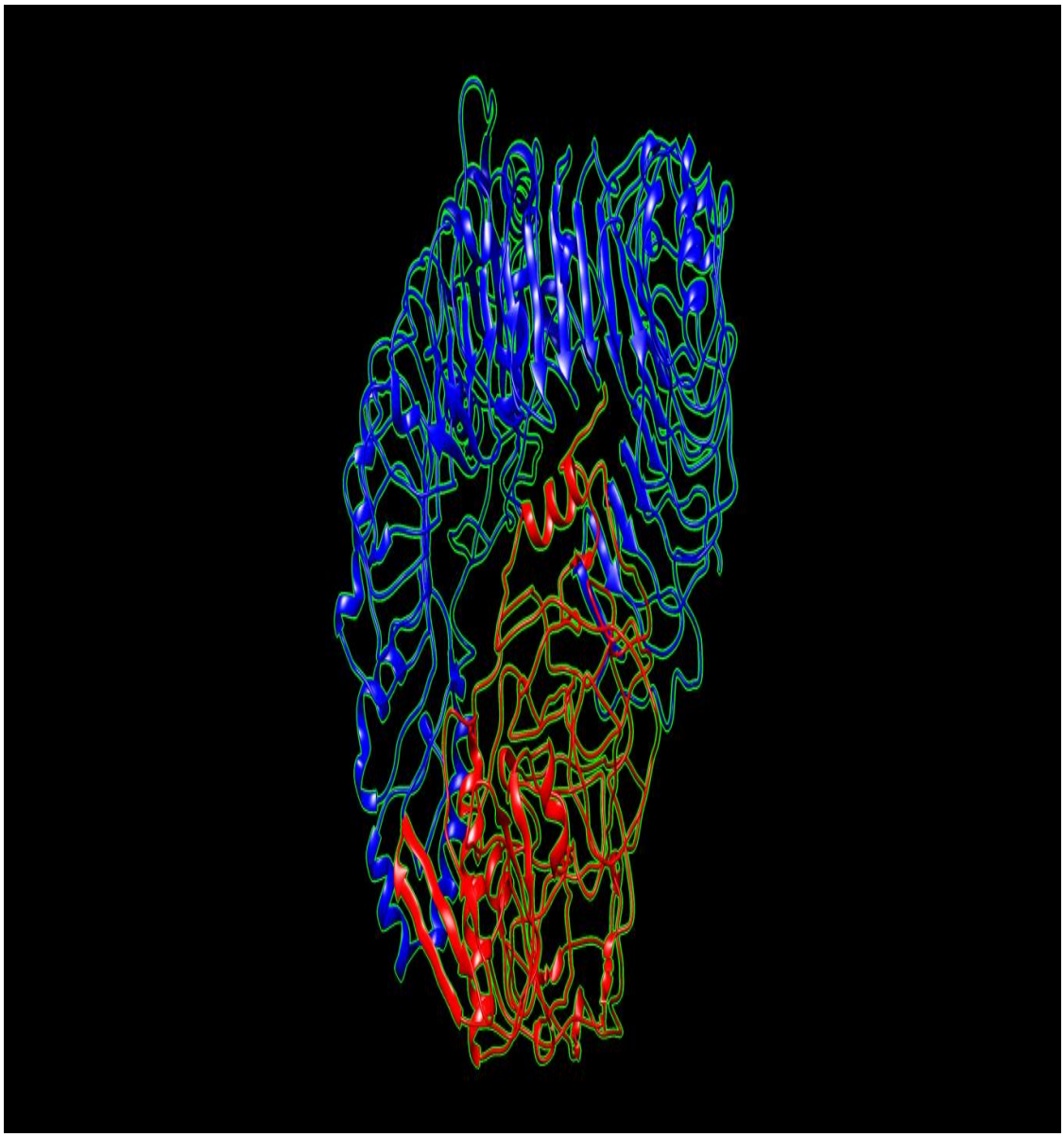

Figure 14 : Docked model between TLR4 and ligand

6-11 Codon adaptation and in silico:

It is vital to inform whether such multi-epitope vaccines can be cloned and expressed in an appropriate expression vector. Therefore, in silico cloning was utilized to examine the cloning and expression efficiency of the vaccine construct in an expression vector. In the vector E. coli (strain K12), the codon sequence was improved utilizing vectorbuilder . It gave important parameters like a GC content of 54.% percent and a Codon Adaptation Index (CAI) of 0.94. The results show that the designed vaccine design may be expressed efficiently in the E. coli K12 (host) expression vector. The multi-epitope vaccine-adapted codon sequence was also inserted into the pET28a (+) vector using the SnapGene tool for restriction cloning.

Improved DNA**[1]: GC=54.34%, CAI=0.94**

CTGTGTGCCCATGGTGCCCCGCAGAGCATTACCGAACTGTGCAGCGAATACCGCAATACCCAGATTT
ATACCATCAACGATAAAATTCTGAGCTATACCGAATCAATGGCGGGCAAACGTGAAATGGTGATTA
TTACCTTTAAAAGCGGCGCAACCTTTCAGGTGGAAGTGCCGGGCAGCCAGCACATTGATAGCCAGA
AAAAAGCGATTGAACGCATGAAAGATACCCTGCGCATTACCTATTTAACCGAAACCAAAATTGATAA
ACTGTGTGTGTGGAATAACAAAACCCCGAATAGCATTGCCGCGATTAGCATGGAAGCAGCAGCCAA
AAAAGTGTGGCGCATCCAGGCGGGTAAGGCGGCCTATAGCCTGTATGATGGCGCAACCCTGAATCT
GGCCGCCTATAGCCTGTATGATGGCGCGACCCTGAACCTGGCCGCGTATAACACCGATGAATTTGG
CTCAAGCGAATATGCGGCATATTACGAAAACGGCATTCATAAACGTACCTATGGCCCGGGCCCGGG
TGCGATTGTGGGTGGCATTGCGACCGGCACCGCGGTGGGTACCGTTGGCCCGGGCCCGGGCATTG
TTGGCGGCATTGCGACGGGCACCGCCGTGGGTACGGTGAGCGGCCCGGGCCCGGGCATTCCGGCG
ATTGTGGGCGGCATTGCGACCGGCACCGCAGTGGGCGGCCCGGGCCCGGGCCTGGGCGTGAAAC
GCACCCTGAGTGTGAAAAACGCGCTGGTGATTGGTCCGGGCCCGGGCGGCGTGAAACGCACCCTG
AGCGTGAAAAACGCCCTGGTGATCAAAGGCCCGGGCCCGGGCGCGCTGGGCGTGAAACGCACCCT
GTCGGTGAAAAACGCCCTGGTTGGCCCGGGTCCGGGCCGCAATGGCTTCTACCTGGGCCTGAATTT
TGCGGAAGGCTCATATAAAAAACATTACTGGATTAAAGGTGGCCAGTGGAACAAACTGGAAGTGG
ATATGAAAAAAGGCCTGAACATTATTGCGCCGCCGGAAGGCGGCTACAAAGACAAACCGAAAAAA
GCGATGAATGGCGAAATCGTGCTGCGCCCGGATCCGAAACGCACCATTAAAAAATATCGTGATAAC
ATTGCGAAAGAATATGAAAACAAATTTAAAAACCAGAAAAAAAATGGTAAACCGTGGGGCATTAAC
GCGAGCGGGAACGCCTGCAACATTAAAAAACTGGGTCTGAACTTTGCGGAAGGCAGCTATATCAA
AGGCCAGGGCAGCAAAAAAGCCGGCGATGTGAGCGCCAAAATGGTGCAGACCGCGCCGGTGACC
ACCAAAAAAGCCGGCGGCATTGATACCCATATTCATTTTATTCCGCCTCAGCAGACCTAA

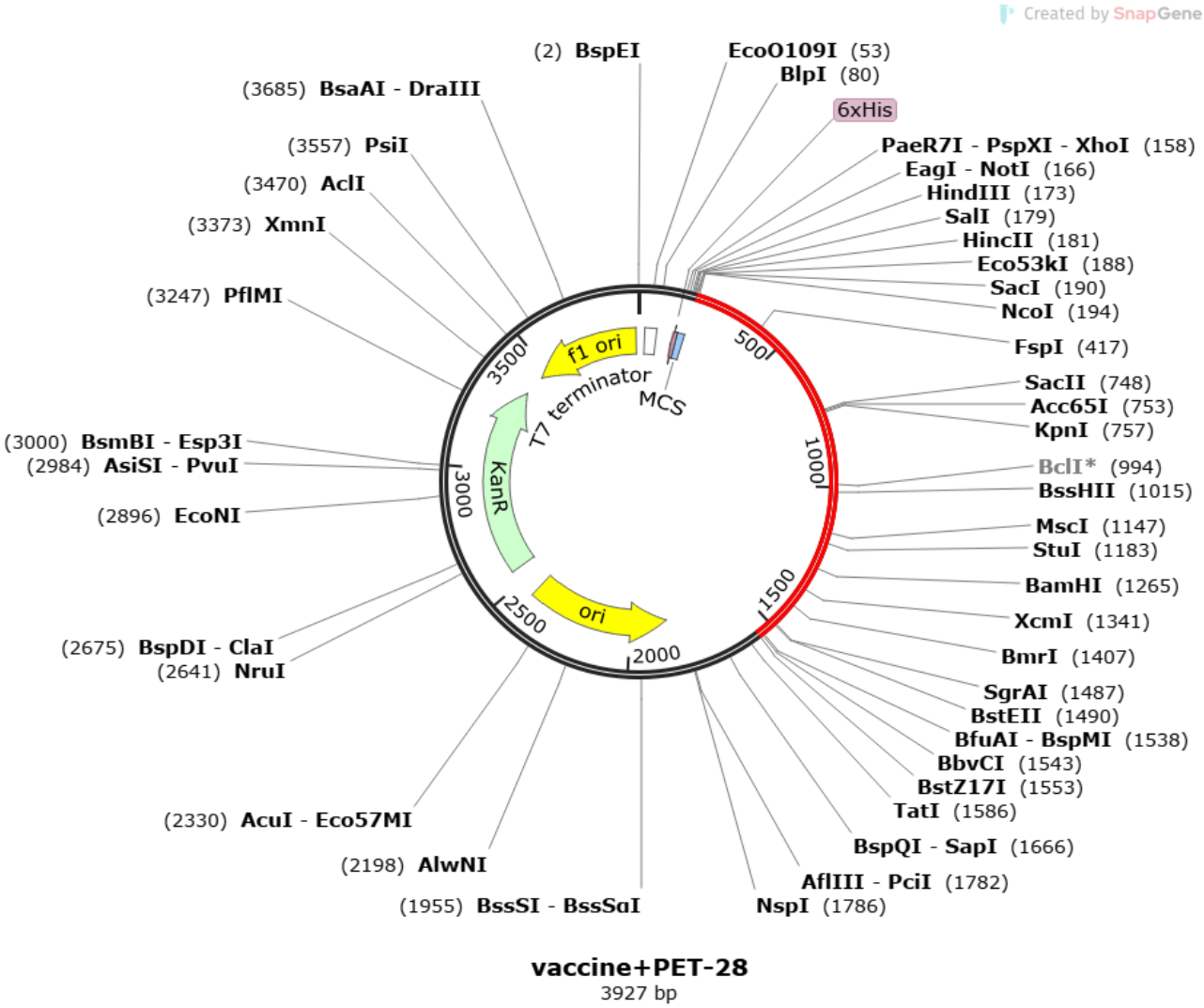

Figure 15 Restriction cloning of final multi-epitope vaccine using PET-28a (+) expression vector using the in silico space

Restriction cloning of final multi-epitope vaccine using PET-28a (+) expression vector using the in silico space. The Black circle indicates the vector , and the red part is the place in which the vaccine is inserted.

Discussion

Helicobacter pylori is widely recognized as a major contributor to gastrointestinal diseases, being strongly linked to conditions such as gastritis, peptic ulcers, and gastric cancer [60] . The reliance on antibiotics to manage these infections poses significant challenges; not only is it costly, but it also leads to the ongoing emergence of antibiotic-resistant strains[6].
In contrast, vaccination presents a viable alternative for large-scale prevention of infections [61] .

With the extensive data now available on the genomes and proteomes of various pathogens, the development of innovative reverse vaccines has become increasingly achievable. Traditional methods for vaccine development are proving to be less efficient and more expensive in terms of both financial resources and time. Consequently, employing immunoinformatics-based approaches for vaccine design is viewed as a more reliable, safe, and cost-effective strategy that provides enhanced specificity and efficacy in combating infections[10][8].
Despite these advancements, a significant challenge remains: there is currently no effective vaccine against Helicobacter pylori infection. To address this issue, researchers have utilized proteins such as VacA, HpaA, Omp 6, OipA, UreB, and IceA2 to develop a new vaccine[62]. Experimental results indicate that this vaccine can elicit protective immune responses from both humoral and cellular components of the immune system, providing effective protection against Helicobacter pylori infection.
The study utilizes immunoinformatics and structural analyses to pinpoint potential T cell and B cell epitopes for the purpose of devising a novel vaccine through a logical approach. In designing a peptide vaccine against the chosen protein, T cell and B cell epitopes are identified. In the final vaccine design, the presence of IFN-gamma-inducing epitopes is verified, and the chosen epitopes are confirmed to be IFN-gamma-inducing. For the development of the vaccine construct, the targeted T and B cell epitopes are fused using the KK, AAY, and GPGPG linkers [63]. Acknowledging the limited number of epitopes involved, subunit vaccines are recognized for having poor immunogenicity, but this can be improved with the addition of an adjuvant. To create the vaccine, Heat-labile enterotoxin B chain from E.COLI (LTB) [64] is incorporated at the N-terminus using the EAAAK linker.
Antigenic and non-allergenic property predictions indicate the vaccine's efficacy and safety , The vaccine protein is determined to be antigenic with a score of 0.7752. The final vaccine is found to be 456 amino acids long, with a non-allergenic construct that

has the ability to elicit an antigenic response. Physicochemical parameters of the chimeric protein are investigated. Importantly, the designed vaccine is revealed to be highly soluble, facilitating purification. The vaccine's secondary and tertiary structure is generated by SOMPA and the Phyre2 server. Random coils dominate in a higher percentage of secondarily designed vaccine constructs (63.38%), with 6.80%, and 29.82% of alpha helix and extended strand, , respectively.

To detect errors in the 3D model, validation servers are utilized. The Ramachandran plot analysis, conducted using the Procheck server, confirms that most of the vaccine residues are located within the most favorable regions. In recent years, TLR agonists have emerged as effective adjuvants, with several already incorporated into approved vaccines. Molecular docking studies are performed using ClusPro 2.0 to assess the binding affinity between the chimeric vaccine and TLR2/TLR4 [65][66]. The vaccine has been successfully cloned into the pET28a (+) vector for expression in a bacterial medium, facilitating production and purification processes. However, prior to public use, the vaccine must undergo rigorous testing through real-time in vitro experiments. Overall, this multi-peptide subunit vaccine holds significant promise as an alternative therapeutic option to combat Helicobacter pylori infections.

Conclusion

In conclusion, this study presents an innovative multi-epitope vaccine aimed at Helicobacter pylori. The extensive immunoinformatics analyses, along with the evaluation of protein structure and detailed physicochemical and molecular docking studies, underscore the vaccine's potential to elicit significant humoral and cellular immune responses. These promising results strongly advocate for further experimental validation to confirm the vaccine's efficacy.

Table 32: Study tools :

NO.	TOOL	Description
1	ANTIGENpro	ANTIGENpro is a sequence-based, alignment-free and pathogen-independent predictor of protein antigenicity.
2	Vaxijen	Vaxijen is a server which is designed for predicting protective antigens, tumor antigens, and subunit vaccines, and was employed to assess the immunogenic potential of each epitope sequence.
3	Psortb	It is a server, which is used for bacterial protein subcellular localization prediction.

4	AllerTop	Is a server to eliminate allergic proteins. AllerTop employs an alignment-free technique based on key physicochemical features of proteins, demonstrating a server sensitivity of approximately 94% in predicting allergens.
5	ToxinPred	To evaluate the toxicity of selected epitopes, utilizing the support vector machine (SVM) approach on the server, with all parameters set to their default values. SVM, a widely adopted machine-learning technique for toxicity prediction, effectively distinguishes between toxic and nontoxic epitopes.
6,7,8,9	IEDB	A website which provides a collection of tools for the prediction and analysis of immune epitopes. It serves as a companion site to the Immune Epitope Database (IEDB), a manually curated database of experimentally characterized immune epitopes. It contains MHC I and II epitopes prediction, Class I Immunogenicity prediction and population coverage of T-cell epitope prediction
10	ABCpred	used to predict B cell epitopes which are ranked according to their score obtained by trained recurrent neural networks
11	ProtParam	It scrutinize the physical and chemical characteristics of the multi-epitope vaccines. This contained an assessment of amino acid composition, estimated half-life, instability index, extinction coefficient, theoretical pI, atomic composition, molecular weight, and the grand average of hydropathicity (GRAVY) to facilitate subsequent experimental investigations.
12	SOMPA	SOMPA tool predicts the secondary structure Secondary attributes, encompassing α-helix, β-strand, and random coils
13	SOLPro	To forecast protein solubility, this tool determines the likelihood of a protein being soluble upon overexpression in E. coli through a two-stage SVM architecture. Each classifier in the initial layer receives distinct sets of attributes to characterize the sequence. The final SVM classifier synthesizes the data, predicting the protein's solubility and the associated probability
14	Protein-Sol	Protein-sol is a simple and free, web based suite of theoretical calculations and predictive algorithms for understanding protein solubility and stability. The scaled solubility value (QuerySol) is the predicted solubility. The population average for the experimental dataset (PopAvrSol) is 0.45, and therefore any scaled solubility value greater than 0.45 is predicted to have a higher solubility than the average soluble E.coli protein from the experimental solubility dataset, and any protein with a lower scaled solubility value is predicted to be less soluble.

15	Phyre2	This tool develops the tertiary structure of the subunit vaccine protein for its functional characterization.
16	PROCHECK	This tool provides a Ramachandran plot. This plot, detailing the percentage of residues in favored, allowed, and outlier regions, was employed to evaluate the modeled tertiary structure's quality.
17	GalaxyRefine	GalaxyRefine used to enhance the model to its perfect specialisation.
18	ProSA-web	a widely-used tool, assesses 3D models for potential errors using the z-score, indicating overall model quality by measuring the deviation of the structure's total energy from an energy distribution derived from random conformations. Z-scores outside the range characteristic for native proteins indicate potential structural errors
19	Ellipro Server	B.cell epitopes prediction
20	C-ImmSim	is composed by three modules or compartments which corresponds to three different and separate (The bone marrow, the thymus and the lymph node) anatomical regions of a mammalian immune system.
21	Clus-Pro 2.0	ClusPro employs a docking approach that utilizes the pairwise RMSD histogram of all docked conformations to rapidly categorize them. It supports rigid body docking techniques like DOT and ZDOCK, both based on the FFT correlation approach.
22	Chimera	which is a program for the interactive visualization and analysis of molecular structures and related data, including density maps, trajectories, and sequence alignments
23	vectorbuilder	The multi-epitope vaccine sequence underwent reverse translation and subsequent optimization for Escherichia coli codon usage, enhancing the expression efficiency when cloned into the specified expression system. The vectorbuilder server facilitated this process
24	SnapGene	The engineered construct was cloned into the pET28a (+) expression vector using SnapGene

References

1. Parikh, N. S., and Ahlawat, R. (2024). Helicobacter Pylori. In *StatPearls*. StatPearls Publishing.

2. Hooi, J. K. Y., Lai, W. Y., Ng, W. K., Suen, M. M. Y., Underwood, F. E., Tanyingoh, D., Malfertheiner, P., Graham, D. Y., Wong, V. W. S., Wu, J. C. Y., Chan, F. K. L., Sung, J. J. Y., Kaplan, G. G., and Ng, S. C. (2017). Global Prevalence of Helicobacter pylori Infection: Systematic Review and Meta-Analysis. *Gastroenterology*, *153*(2), 420–429. https://doi.org/10.1053/j.gastro.2017.04.022

3. Wroblewski, L. E., Peek, R. M., and Wilson, K. T. (2010). Helicobacter pylori and gastric cancer: factors that modulate disease risk. *Clinical Microbiology Reviews*, *23*(4), 713–739. https://doi.org/10.1128/CMR.00011-10

4. Schistosomes, liver flukes and Helicobacter pylori. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Lyon, 7-14 June 1994. (1994). *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans / World Health Organization, International Agency for Research on Cancer*, *61*, 1–241.

5. Talebi Bezmin Abadi, A., and Yamaoka, Y. (2018). Helicobacter pylori therapy and clinical perspective. *Journal of Global Antimicrobial Resistance*, *14*, 111–117. https://doi.org/10.1016/j.jgar.2018.03.005

6. Boyanova, L., Hadzhiyski, P., Gergova, R., and Markovska, R. (2023). Evolution of Helicobacter pylori Resistance to Antibiotics: A Topic of Increasing Concern. *Antibiotics (Basel, Switzerland)*, *12*(2). https://doi.org/10.3390/antibiotics12020332

7. Yuan, C., Yu, C., Sun, Q., Xiong, M., Zhou, S., Zeng, M., and Song, H. (2023). Research on antibiotic resistance in Helicobacter pylori: a bibliometric analysis of the past decade. *Frontiers in Microbiology*, *14*. https://doi.org/10.3389/fmicb.2023.1208157

8. Yurina, V., and Adianingsih, O. R. (2022). Predicting epitopes for vaccine development using bioinformatics tools. *Therapeutic Advances in Vaccines and Immunotherapy*, *10*, 25151355221100216. https://doi.org/10.1177/25151355221100218

9. Raoufi, E., Hemmati, M., Eftekhari, S., Khaksaran, K., Mahmodi, Z., Farajollahi, M. M., and Mohsenzadegan, M. (2020). Epitope Prediction by Novel Immunoinformatics Approach: A State-of-the-art Review. *International Journal of Peptide Research and Therapeutics*, *26*(2), 1155–1163. https://doi.org/10.1007/s10989-019-09918-z

10. Kanampalliwar, A. M. (2020). Reverse vaccinology and its applications. *Methods in Molecular Biology*, *2131*, 1–16. https://doi.org/10.1007/978-1-0716-0389-5_1

11. Ma, J., Qiu, J., Wang, S., Ji, Q., Xu, D., Wang, H., Wu, Z., and Liu, Q. (2021). A Novel Design of Multi-epitope Vaccine Against Helicobacter pylori by Immunoinformatics Approach. *International Journal of Peptide Research and Therapeutics*, *27*(2), 1027–1042. https://doi.org/10.1007/s10989-020-10148-x

12. *UniProt*. (n.d.). Retrieved July 18, 2024, from https://www.uniprot.org/uniprotkb/Q48245/entry

13.*UniProt*. (n.d.). Retrieved July 19, 2024, from https://www.uniprot.org/uniprotkb/Q48244/entry.

14. *UniProt*. (n.d.). Retrieved July 19, 2024, from https://www.uniprot.org/uniprotkb/B4XMC6/entry

15. *UniProt*. (n.d.). Retrieved July 19, 2024, from https://www.uniprot.org/uniprotkb/O30511/entry

16. *UniProt*. (n.d.). Retrieved July 19, 2024, from https://www.uniprot.org/uniprotkb/Q56S04/entry

17. *UniProt*. (n.d.). Retrieved July 19, 2024, from https://www.uniprot.org/uniprotkb/A0A448UFR2/entry

18. ***UniProt*. (n.d.). Retrieved July 19, 2024, from** https://www.uniprot.org/uniprotkb/U3GLE7/entry

19. *UniProt*. (n.d.). Retrieved July 19, 2024, from https://www.uniprot.org/uniprotkb/A0A7K1N3E4/entry

20. *UniProt*. (n.d.). Retrieved July 19, 2024, from https://www.uniprot.org/uniprotkb/Q9ZNC3/entry

21. *UniProt*. (n.d.). Retrieved July 19, 2024, from https://www.uniprot.org/uniprotkb/L7SY90/entry

22. *Page restricted | ScienceDirect*. (n.d.). Retrieved July 19, 2024, from https://www.sciencedirect.com/science/article/abs/pii/S1359511322002471

23. *Page restricted | ScienceDirect*. (n.d.). Retrieved July 19, 2024, from https://www.sciencedirect.com/science/article/abs/pii/S1567134817300497

24. *Page restricted | ScienceDirect*. (n.d.). Retrieved July 19, 2024, from https://www.sciencedirect.com/science/article/abs/pii/S0041010116301106

25. Voland, P., Hafsi, N., Zeitner, M., Laforsch, S., Wagner, H., and Prinz, C. (2003). Antigenic properties of HpaA and Omp18, two outer membrane proteins of Helicobacter pylori. *Infection and Immunity*, *71*(7), 3837–3843.

26. *Just a moment...* (n.d.). Retrieved July 19, 2024, from https://www.hindawi.com/journals/bmri/2015/571280/

27. *Page restricted | ScienceDirect*. (n.d.). Retrieved July 19, 2024, from https://www.sciencedirect.com/science/article/abs/pii/S0944501323002410

28. O'Toole, P. W., Janzon, L., Doig, P., Huang, J., Kostrzynska, M., and Trust, T. J. (1995). The putative neuraminyllactose-binding hemagglutinin HpaA of Helicobacter pylori CCUG 17874 is a lipoprotein. *Journal of Bacteriology*, *177*(21), 6049–6057.

29. Jones, A. C., Logan, R. P., Foynes, S., Cockayne, A., Wren, B. W., and Penn, C. W. (1997). A flagellar sheath protein of Helicobacter pylori is identical to HpaA, a putative N-acetylneuraminyllactose-binding hemagglutinin, but is not an adhesin for AGS cells. *Journal of Bacteriology*, *179*(17), 5643–5647.

30. Lundström, A. M., Blom, K., Sundaeus, V., and Bölin, I. (2001). HpaA shows variable surface localization but the gene expression is similar in different Helicobacter pylori strains. *Microbial Pathogenesis*, *31*(5), 243–253. https://doi.org/10.1006/mpat.2001.0466

31. Oktem-Okullu, S., Karaman, T., Akcelik-Deveci, S., Timucin, E., Sezerman, O. U., Mansur-Ozen, N., Buyukcolak, Y., and Tiftikci, A. (2023). Effect of the switch status of Helicobacter pylori outer inflammatory protein A on gastric diseases. *AMB Express*, *13*(1), 109. https://doi.org/10.1186/s13568-023-01621-z

32. Motamedi, H., Abiri, R., Salari, F., Jalili, C., and Alvandi, A. (2023). Reduction of UreB and CagA expression level by siRNA construct in Helicobacter pylori strain SS1. *BMC Microbiology*, *23*(1), 401. https://doi.org/10.1186/s12866-023-03143-x

33. Wu, J., Zhu, X., Guo, X., Yang, Z., Cai, Q., Gu, D., Luo, W., Yuan, C., and Xiang, Y. (2021). Helicobacter urease suppresses cytotoxic CD8 + T cell responses through activating Myh9-dependent induction of PD-L1. *International Immunology*. https://doi.org/10.1093/intimm/dxab044

34. *Page restricted | ScienceDirect*. (n.d.). Retrieved July 19, 2024, from https://www.sciencedirect.com/science/article/abs/pii/S0006291X22006325

35. *Human Leukocyte Antigen (HLA) System - Human Leukocyte Antigen (HLA) System - MSD Manual Professional Edition*. (n.d.). Retrieved July 19, 2024, from https://www.msdmanuals.com/professional/immunology-allergic-disorders/biology-of-the-immune-system/human-leukocyte-antigen-hla-system

36. Ikhtiar, A. M., Jazairi, B., Khansa, I., and Othman, A. (2018). HLA class I alleles frequencies in the Syrian population. *BMC Research Notes*, *11*(1), 324. https://doi.org/10.1186/s13104-018-3427-1

37. *Just a moment...* (n.d.). Retrieved July 19, 2024, from https://www.beckman.com/resources/cell-types/blood-cells/leukocytes/lymphocytes/t-cells/t-helper-cells

38. Jazairi, B., Khansaa, I., Ikhtiar, A., and Murad, H. (2016). Frequency of HLA-DRB1 and HLA-DQB1 Alleles and Haplotype Association in Syrian Population. *Immunological Investigations*, *45*(2), 172–179. https://doi.org/10.3109/08820139.2015.1131293

39. Saha, S., and Raghava, G. P. S. (2007). Prediction methods for B-cell epitopes. *Methods in Molecular Biology*, *409*, 387–394. https://doi.org/10.1007/978-1-60327-118-9_29

40. *B Cell Epitope Prediction Service - Creative Biolabs*. (n.d.). Retrieved July 19, 2024, from https://www.creative-biolabs.com/b-cell-epitopes-prediction-service.html?msclkid=42615b52e8bc179af1703ed9895e335a&utm_source=bing&utm_medium=cpc&utm_campaign=lyj-com-8.0+-+All+Campiagns-2021.03&utm_term=b+cell+epitope&utm_content=2.5.4.5.2.9+B+Cell+Epitopes+Prediction+Service

41. Doytchinova, I. A., and Flower, D. R. (2007). VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics*, *8*, 4. https://doi.org/10.1186/1471-2105-8-4

42. Dimitrov, I., Flower, D. R., and Doytchinova, I. (2013). AllerTOP--a server for in silico prediction of allergens. *BMC Bioinformatics*, *14 Suppl 6*(Suppl 6), S4. https://doi.org/10.1186/1471-2105-14-S6-S4

43. Dimitrov, I., Flower, D. R., and Doytchinova, I. (2013). AllerTOP--a server for in silico prediction of allergens. *BMC Bioinformatics*, *14 Suppl 6*(Suppl 6), S4. https://doi.org/10.1186/1471-2105-14-S6-S4

44. Li, S., Zhao, W., Xia, L., Kong, L., and Yang, L. (2023). How Long Will It Take to Launch an Effective Helicobacter pylori Vaccine for Humans? *Infection and Drug Resistance*, *16*, 3787–3805. https://doi.org/10.2147/IDR.S412361

45. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., and Bairoch, A. (2005). Protein identification and analysis tools on the expasy server. In J. M. Walker (Ed.), *The proteomics protocols handbook* (pp. 571–607). Humana Press. https://doi.org/10.1385/1-59259-890-0:571

46. Hebditch, M., Carballo-Amador, M. A., Charonis, S., Curtis, R., and Warwicker, J. (2017). Protein-Sol: a web tool for predicting protein solubility from sequence. *Bioinformatics*, *33*(19), 3098–3100. https://doi.org/10.1093/bioinformatics/btx345

47. Magnan, C. N., Randall, A., and Baldi, P. (2009). SOLpro: accurate sequence-based prediction of protein solubility. *Bioinformatics*, *25*(17), 2200–2207. https://doi.org/10.1093/bioinformatics/btp386

48. *Just a moment...* (n.d.). Retrieved July 19, 2024, from https://www.pnas.org/doi/full/10.1073/pnas.0811922106

49. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols*, *10*(6), 845–858. https://doi.org/10.1038/nprot.2015.053

50. *Just a moment...* (n.d.). Retrieved July 19, 2024, from https://onlinelibrary.wiley.com/doi/abs/10.1107/s0021889892009944

51. Wiederstein, M., and Sippl, M. J. (2007). ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Research*, *35*(Web Server issue), W407-10. https://doi.org/10.1093/nar/gkm290

52. Ponomarenko, J., Bui, H.-H., Li, W., Fusseder, N., Bourne, P. E., Sette, A., and Peters, B. (2008). ElliPro: a new structure-based tool for the prediction of antibody epitopes. *BMC Bioinformatics*, *9*, 514. https://doi.org/10.1186/1471-2105-9-514

53. Rapin, N., Lund, O., Bernaschi, M., and Castiglione, F. (2010). Computational immunology meets bioinformatics: the use of prediction tools for molecular binding in the simulation of the immune system. *Plos One*, *5*(4), e9862. https://doi.org/10.1371/journal.pone.0009862

54.*Just a moment...* (n.d.). Retrieved July 19, 2024, from https://www.cell.com/structure/fulltext/S0969-2126(20)30209-4?_returnURL=https%3A%2F%2Flinkinghub.elsevier.com%2Fretrieve%2Fpii%2FS0969212620302094%3Fshowall%3Dtrue

55.*Just a moment...* (n.d.). Retrieved July 19, 2024, from https://onlinelibrary.wiley.com/doi/10.1002/pro.3943

56. Meliț, L. E., Mărginean, C. O., Mărginean, C. D., and Mărginean, M. O. (2019). The Relationship between Toll-like Receptors and Helicobacter pylori-Related Gastropathies: Still a Controversial Topic. *Journal of Immunology Research*, *2019*, 8197048. https://doi.org/10.1155/2019/8197048

57. Smith, S. M. (2014). Role of Toll-like receptors in Helicobacter pylori infection and immunity. *World Journal of Gastrointestinal Pathophysiology*, *5*(3), 133–146. https://doi.org/10.4291/wjgp.v5.i3.133

58.*Page restricted | ScienceDirect*. (n.d.). Retrieved July 19, 2024, from https://www.sciencedirect.com/science/article/abs/pii/S0024320517301546

59. Su, B., Ceponis, P. J. M., Lebel, S., Huynh, H., and Sherman, P. M. (2003). Helicobacter pylori activates Toll-like receptor 4 expression in gastrointestinal epithelial cells. *Infection and Immunity*, *71*(6), 3496–3502. https://doi.org/10.1128/IAI.71.6.3496-3502.2003

60. Malfertheiner, P., Camargo, M. C., El-Omar, E., Liou, J.-M., Peek, R., Schulz, C., Smith, S. I., and Suerbaum, S. (2023). Helicobacter pylori infection. *Nature Reviews. Disease Primers*, *9*(1), 19. https://doi.org/10.1038/s41572-023-00431-8

61. *Page restricted | ScienceDirect*. (n.d.). Retrieved July 27, 2024, from https://www.sciencedirect.com/science/article/pii/S0264410X18300173?via%3Dihub

62. Calado, C. R. C. (2022). Antigenic and conserved peptides from diverse Helicobacter pylori antigens. *Biotechnology Letters*, *44*(3), 535–545. https://doi.org/10.1007/s10529-022-03238-x

63. Ghosh, P., Bhakta, S., Bhattacharya, M., Sharma, A. R., Sharma, G., Lee, S.-S., and Chakraborty, C. (2021). A Novel Multi-Epitopic Peptide Vaccine Candidate Against Helicobacter pylori: In-Silico Identification, Design, Cloning and Validation Through Molecular Dynamics. *International Journal of Peptide Research and Therapeutics*, 1–18. https://doi.org/10.1007/s10989-020-10157-w

64. Rezaee, M. A., Rezaee, A., Moazzeni, S. M., Salmanian, A. H., Yasuda, Y., Tochikubo, K., Pirayeh, S. N., and Arzanlou, M. (2005). Expression of Escherichia coli heat-labile enterotoxin B subunit (LTB) in Saccharomyces cerevisiae. *Journal of Microbiology*, *43*(4), 354–360.

65. Torok, A. M., Bouton, A. H., and Goldberg, J. B. (2005). Helicobacter pylori induces interleukin-8 secretion by Toll-like receptor 2- and Toll-like receptor 5-dependent and -independent pathways. *Infection and Immunity*, *73*(3), 1523–1531. https://doi.org/10.1128/IAI.73.3.1523-1531.2005

66. *Just a moment...* (n.d.). Retrieved July 27, 2024, from https://journals.aai.org/jimmunol/article/173/2/1406/81021/Essential-Role-of-MD-2-in-TLR4-Dependent-Signaling