



OPEN ACCESS

EDITED BY

Varun Jaiswal,
Gachon University, Republic of Korea

REVIEWED BY

Azhar Azhar Khan,
Shoolini University, India
Sandeep Kushwaha,
National Institute of Animal Biotechnology
(NIAB), India

*CORRESPONDENCE

Sunil Kumar Rai,
✉ s.raimua.edu
Syed Babar Jamal,
✉ babar.jamal@numspak.edu.pk

RECEIVED 26 December 2023

ACCEPTED 08 April 2024

PUBLISHED 17 May 2024

CITATION

Qureshi H, Basheer A, Faheem M, Arshad MW,
Rai SK and Jamal SB (2024), Designing a multi-
epitope vaccine against *Shigella dysenteriae*
using immuno-informatics approach.
Front. Genet. 15:1361610.
doi: 10.3389/fgene.2024.1361610

COPYRIGHT

© 2024 Qureshi, Basheer, Faheem, Arshad, Rai
and Jamal. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other forums is
permitted, provided the original author(s) and
the copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with these
terms.

Designing a multi-epitope vaccine against *Shigella dysenteriae* using immuno-informatics approach

Hurria Qureshi¹, Amina Basheer¹, Muhammad Faheem²,
Muhammad Waqar Arshad³, Sunil Kumar Rai^{4*} and
Syed Babar Jamal^{1*}

¹Department of Biological Sciences, National University of Medical Sciences, Rawalpindi, Punjab, Pakistan, ²Department of Biomedical Sciences, School of Medicine and Health Sciences, University of North Dakota, Grand Forks, ND, United States, ³Department of Urology, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States, ⁴Medical University of the Americas Navis, Charlestown, Saint Kitts and Nevis, West Indies

Shigella dysenteriae has been recognized as the second most prevalent pathogen associated with diarrhea that contains blood, contributing to 12.9% of reported cases, and it is additionally responsible for approximately 200,000 deaths each year. Currently, there is no *S. dysenteriae* licensed vaccine. Multidrug resistance in all *Shigella* spp. is a growing concern. Current vaccines, such as O-polysaccharide (OPS) conjugates, are in clinical trials but are ineffective in children but protective in adults. Thus, innovative treatments and vaccines are needed to combat antibiotic resistance. In this study, we used immuno-informatics to design a new multi-epitope vaccine and identified *S. dysenteriae* strain SD197's membrane protein targets using *in-silico* methods. The target protein was prioritized using membrane protein topology analysis to find membrane proteins. B and T-cell epitopes were predicted for vaccine formulation. The epitopes were shortlisted based on an IC50 value <50, antigenicity, allergenicity, and a toxicity analysis. In the final vaccine construct, a total of 8 B-cell epitopes, 12 MHC Class I epitopes, and 7 MHC Class II epitopes were identified for the Lipopolysaccharide export system permease protein LptF. Additionally, 17 MHC Class I epitopes and 14 MHC Class II epitopes were predicted for the Lipoprotein-releasing ABC transporter permease subunit LolE. These epitopes were selected and linked via KK, AAY, and GGGs linkers, respectively. To enhance the immunogenic response, RGD (arginine-glycine-aspartate) adjuvant was incorporated into the final vaccine construct. The refined vaccine structure exhibits a Ramachandran score of 91.5% and demonstrates stable interaction with TLR4. Normal Mode Analysis (NMA) reveals low eigenvalues (3.925996e-07), indicating steady and flexible molecular mobility of docked complexes. Codon optimization was carried out in an effective microbial expression system of the *Escherichia coli* K12 strain using the recombinant plasmid pET-28a (+). Finally, the entire *in-silico* analysis suggests that the suggested vaccine may induce a significant immune response against *S. dysenteriae*, making it a promising option for additional experimental trials.

KEYWORDS

multi-epitope, *Shigella dysenteriae*, *in silico* vaccine, molecular docking, membrane proteins, shigellosis, multi-drug resistance, comparative analysis

1 Introduction

Shigella dysenteriae is recognized as a major contributor to bacillary dysentery or shigellosis, transmitted to humans through the consumption of contaminated water or food. Additionally, person-to-person transmission occurs through the fecal-oral route (Nasser et al., 2022). *S. dysenteriae* is among the four serological groups present in *Shigella* bacteria. It stands as the sole bacterial strain categorized under the clinical group of *Dysenteriae Shigella* species. This designation stems from its unique capacity to generate Shiga toxins (Stxs) (Pakbin et al., 2022). *Shigella* poses a substantial disease burden worldwide, ranking among those bacteria that causes of fatal diarrheal cases globally. The most severely affected demographic consists of children under 5 years old in developing countries (Kumar and Omeershffudin, 2020).

It has been reported that shigellosis results in over 200,000 fatalities globally. As of now, there are no widely accessible vaccines for shigellosis. Nonetheless, several candidates are undergoing evaluation in clinical, yielding essential data and insights. The emergence of antimicrobial-resistant strains (AMR) has heightened concerns regarding *Shigella* (Raso et al., 2023). Approximately half of *Shigella* strains in numerous regions worldwide are now resistant to multiple drugs (Farsiani and Sasan, 2020). The administration of antibiotics for common bacterial infections significantly contributes to diminishing disease prevalence and mortality rates. Nonetheless, inappropriate usage or excessive reliance on antibiotics for treating diarrhea exacerbates the problem of antibiotic resistance (Ranjbar and Farahani, 2019). Bacteria employ various antibiotic-resistance mechanisms, including the utilization of efflux pumps to expel drugs, alterations in membrane permeability, and the expression of enzymes capable of modifying and deactivating molecules (Raso et al., 2023).

According to World Health Organization (WHO) various antibiotics have been used such as ceftriaxone, azithromycin, ciprofloxacin, and pivmecillinam to treat shigellosis. However, there is an efficient increase in the resistance to quinolones and cephalosporin compounds further complicates the situation (Ahamed and Giria, 2021). As for vaccine treatment, different vaccine candidates for *Shigella*, including orally administered live attenuated or inactivated vaccines, as well as subunit vaccines administered via injection, have undergone human trials to assess their safety and immunogenicity. The most advanced vaccine against *Shigella* is O-polysaccharide (OPS) conjugates, while successful in adults and older children, did not provide protection against shigellosis in children under 3 years old, who are the most vulnerable population (Desalegn et al., 2024).

The Ipa proteins, particularly IpaB, IpaC, and IpaD, are key components of the *Shigella* type 3 secretion system (T3SS), this complex is vital in order to internalized into epithelial cell, a beneficial aspect of whole-cell vaccine strategies involves including Ipa proteins alongside other potentially protective proteins. A compelling alternative approach to subunit vaccines is the invasin complex, or Invaplex technology, developed by WRAIR (Turbyfill et al., 2022). The initial version of Invaplex, known as native Invaplex or InvaplexNAT, comprised a complex containing Ipa antigens (IpaB, IpaC, and IpaD) and LPS extracted from *Shigella* spp. Three intranasally administered doses were found to be well tolerated and immunogenic in phase

I studies conducted in the US. However, in a controlled human infection model study, this formulation failed to provide protection. The extent of added protection achieved by including Ipa proteins in the vaccine are presently uncertain. However, available evidence from animal and human studies indicates that the immune response to IpaB and IpaC is likely to augment protection (MacLennan et al., 2022). The combination of these factors underscores the urgent global public health necessity for developing a vaccine against *Shigella*.

This study aims to identify possible vaccine targets against *S. dysenteriae* by using immuno-informatics methodology to design multiepitope vaccine. The discovery of two stable antigenic membrane proteins from the complete reference genome of *S. dysenteriae* strain SD197 holds significance, especially considering its genomic characteristics. *S. dysenteriae* strain SD197, with its fully sequenced genome consisting of a single circular chromosome spanning 4,369,232 nucleotides and encompassing 4,664 genes, provides a thorough genetic profile of this pathogen (Kumar and Omeershffudin, 2020). Lipopolysaccharide export system permease protein LptF and Lipoprotein-releasing ABC transporter permease subunit LolE are surface localizing proteins, showed significant promise as vaccine candidates. The role of LptF in bacterial cell is that LptF interacts with a transmembrane domain (TMD). Additionally, LptF might function as lateral gates. These gates could potentially aid in the passage of lipopolysaccharides (LPS) across the inner membrane and deposits them onto the Lpt periplasmic bridge, which is a structure involved in further transporting LPS to the outer membrane of the cell (Lundstedt et al., 2020). The role of the LolE subunit within this transporter complex is believed to involve recognizing and transporting lipoproteins from the inner membrane to the periplasm. It participates in the energy-dependent translocation of lipoproteins across the inner membrane, collaborating with other components of the Lol system (Okuda and Tokuda, 2009). Additionally, known for their virulence properties, also present themselves as potential peptide-based vaccine candidates against *S. dysenteriae*. In this study, we employed an immune-informatics strategy to analyze the complete proteome of *S. dysenteriae* and identify potential vaccine candidates. This approach was chosen due to its capacity to focus on candidates derived from extracellular and outer membrane proteins, which exhibited efficacy in prior animal studies. Moreover, we mapped T and B cell epitopes for these proteins, utilizing the top-ranking epitopes to formulate a multiepitope vaccine. The predicted B-cell epitopes, MHC class I and MHC class II were merged together by using linkers. A comprehensive characterization of various properties of the vaccine candidates such as antigenicity, allergenicity, and physicochemical properties were conducted. Also, Normal mode analysis was carried out on a vaccine candidate to assess its dynamic behavior and flexibility under diverse conditions. The current research aims to develop a highly effective multi-epitope peptide vaccine against *S. dysenteriae* with the goal of effectively controlling its infection.

2 Materials and methods

2.1 Genomic and core genome analysis

This study conducted a comprehensive pan-genome analysis, utilizing the complete genome sequences of all 27 strains of *S. dysenteriae*. The reference genome and proteome sequences for these

strains were obtained from the National Center for Biotechnology Information (NCBI) and UniProt, respectively (Jamal et al., 2017). The core genome of *S. dysenteriae* was determined using the Efficient Database framework. For comparative genome analyses using the BLAST Score Ratios web tool, the algorithm employed the Basic Local Alignment Search Tool (BLASTp) for proteins, utilizing the standard scoring matrix BLOSUM62 and an E-value cutoff of 1×10^{-5} . Strain SD197 served as the reference for comparative analysis with other *S. dysenteriae* strains. The core genome of *S. dysenteriae* underwent comparison with the human proteome utilizing BLASTp under default conditions (e-value = 0.0001, bit score ≥ 100 , BLOSUM62 scoring matrix, and identity $\geq 25\%$). Only proteins lacking homology with the human proteome database were selected (Qureshi et al., 2021).

2.2 Identification of therapeutic vaccine targets and membrane protein analysis

Proteins lacking similarity with human proteins were filtered based on their essentiality using BLASTp in the Database of Essential Genes (DEG) (Riasat et al., 2021). To identify essential proteins, we utilized the standard scoring matrix BLOSUM62, an e-value of 0.001, and an identity threshold of at least 25%. Strain SD197 of *S. dysenteriae* was selected for the identification of potential therapeutic vaccine targets. Protein sequences from the SD197 strain were acquired from UniProt, and a subtractive analysis was conducted using the online database CELLO to exclude 34 membrane proteins (Scott et al., 2005; Qureshi et al., 2021).

2.3 Prioritization of vaccine targets

The selected 34 membrane proteins were narrow down on the basis of evaluation of its hydrophobicity, physicochemical properties, antigenicity score, allergenicity and membrane topology score, which was carried out using ExPASy ProtParam (<https://web.expasy.org/protparam/>). Additionally, a membrane topology analysis was performed using TMHMM-2.0 (<https://services.healthtech.dtu.dk/services/TMHMM-2.0/>) to further assess the candidates. Among the potential targets, the prioritized proteins were found to be the Lipopolysaccharide export system permease protein LptF and the Lipoprotein-releasing ABC transporter permease subunit LolE to be antigenic (Rouzbahani et al., 2022).

2.4 B cell epitope prediction

The generation of antibodies through B-cell epitopes plays a crucial role in the development of humoral immunity. To enhance the potential of a multi-epitope vaccine, B-cell epitopes for both proteins were predicted using IEDB (<http://tools.iedb.org/bcell/>) with parameters set by previous studies (Rouzbahani et al., 2022).

2.5 T cell epitope prediction

MHC molecules play essential roles in the adaptive immune response and have a significant impact on various host

immunological reactions. To predict T cell epitopes for both proteins, the IDEB Tepi Tool database (<http://tools.iedb.org/tepitool/>) was utilized. This database follows a six-step process that involves selecting classes and alleles, and it provides results based on percentile rank and IC50 values.

For MHC class I epitopes, all peptides with a predicted IC50 value of ≤ 500 nM and a percentile rank of ≤ 1.0 were chosen as predicted epitopes. On the other hand, for MHC class II epitopes, peptides with a percentile rank of ≤ 10.0 and a predicted IC50 value of $\leq 1,000$ nM were selected as predicted epitopes, following the criteria proposed by Paul in 2016 (Paul et al., 2016). The predicted epitopes obtained from this analysis ranged from approximately 9–11 amino acids in length.

2.6 Antigenicity, allergenicity and toxicity analysis of peptides

After predicting the peptides for B and T cells, an additional analysis was conducted to assess their antigenicity using Vaxijen v2.0 (<http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html>). Only epitopes with significant antigenic potential, as indicated by an antigenicity score ranging from 0.4000 to 2.000 were given priority (Saha et al., 2017). To further refine the selection, the predicted epitopes were evaluated for allergenic and non-allergenic properties using AllerTOP v2.0 (<https://www.ddg-pharmfac.net/AllerTOP/>). Additionally, toxicity level of the epitopes was estimated using ToxinPred, an online tool on the parameter established by previous studies. As a result, only epitopes that were prioritized as non-toxic and non-allergenic were chosen for thorough evaluation (Basheer et al., 2023).

2.7 Population coverage analysis

An important factor in developing an effective multi peptide vaccine is the presence of particular HLA alleles between various populations and ethnic groups worldwide. Different HLA alleles express themselves differently in different ethnic groups around the world (Jalal et al., 2022). To determine the population coverage of the prioritized epitopes, the IEDB population coverage tool was utilized (<http://tools.iedb.org/population/>).

2.8 Construction of vaccine

For the effective and highly immunogenic design of multi-epitopes vaccines, essential components are required such as epitopes, adjuvant, and linkers. In this construct, the integrin-binding motif RGD (arginine-glycine-aspartate) was chosen as the adjuvant due to its remarkable potency among various molecules (Sarkar et al., 2022). This adjuvant was attached at the N-terminal of the construct, merged to the first residue by the EAAAK linker, which has been shown to effectively induce the immune system. To ensure proper arrangement, B cell epitopes were joined together using the KK linker, while MHC class I epitopes were connected using the GGS linker, and MHC class II epitopes were linked with the AAY linker. The linkers EAAAK, AAY, and GGS

were employed to join epitopes together (Khamjan et al., 2023). Moreover, EAAAK was utilized for rigidity, reducing interference in the adjuvant-receptor interaction, while GGGS contribute flexibility (Sanches et al., 2021). To enhance the effectiveness of the Multiple epitope vaccine construct, GPGPG and KK linkers were utilized to connect epitopes (Mahmoodi et al., 2023). At the C-terminal, a thrombin cleavage site and a 6X-His tag were introduced. ProtParam was used to evaluate the physicochemical characteristics of the vaccine construct. Additionally, vaccine construct was validated by antigenicity, allergenicity, and toxicity to ensure its safety and immunogenicity.

2.9 Secondary structure prediction and structure modelling of vaccine construct

The secondary structure of the vaccine construct was predicted using PSIPRED. To obtain a reliable 3D structure, the construct was subjected to trRosetta (<https://yanglab.nankai.edu.cn/trRosetta/>), which provides five potential structures in PDB format. These structures were visualized using CHIMERA. The best structure among the five models was selected based on the Ramachandran plot, which evaluates the stereochemistry of the protein backbone. Additionally, the structure was further validated using the ERRAT score, which assesses the overall quality and accuracy of the model (Rouzbahani et al., 2022).

2.10 Docking interaction of vaccine construct with human toll like receptor 4

The binding sites of vaccine construct were predicted by castp (<http://sts.bioe.uic.edu/castp/index.html?2011>) and binding residues were of TLR were predicted by C-port haddock (<https://alcazar.science.uu.nl/services/CPORT/>) prior to docking of Toll like receptor (PDB id: 2Z65) with vaccine construct. Inflammatory responses caused by gram-negative bacteria primarily involve TLR-4. The TLR4 complex is capable of detecting exceptionally low levels of LPS during bacterial infections. This detection initiates the production of diverse pro-inflammatory substances, such as interleukins (e.g., IL-1, IL-6) and tumor necrosis factor (TNF) which assist in the body's defense against pathogens. Therefore, TLR-4 was selected as the receptor for the potential epitope-based constructed vaccine (Mazgaen and Gurung, 2020; Zamyatina and Heine, 2020). Some studies documented that Docking is performed by HADDOCK and UCSF CHIMERA was used to visualize the outcomes. PRODIGY (<https://wenmr.science.uu.nl/prodigy/>) showed the Gibbs free energy and dissociation constant score of docked complexes (Basheer et al., 2023).

2.11 Normal mode analysis of docked complex

Normal mode analysis (NMA) is accomplished on iMODS. By integrating the coordinates of the docked complex, iMODs employs NMA to compute molecular motion and structural flexibility. iMODS shows the results deformability plot; eigenvalue; covariance matrix; elastic network (Rouzbahani et al., 2022).

TABLE 1 Properties of Lipopolysaccharide export system permease protein LptF and Lipoprotein releasing ABC transporter permease subunit LolE membrane proteins.

Protein	Amino acid	Instability index	Estimated half-life	Aliphatic index	Grand average of hydropathicity (GRAVY)	Theoretical pI	Total helix outside the membrane	Antigenicity score
Lipopolysaccharide export system permease protein LptF	399	26.27 (stable protein)	30 h (mammalian reticulocytes, <i>in vitro</i>) >20 h (yeast, <i>in vivo</i>) >10 h (<i>Escherichia coli</i> , <i>in vivo</i>)	119.89	0.465	9.09	1,159	0.6861
Lipoprotein-releasing ABC transporter permease subunit LolE	360	29.57 (stable protein)	30 h (mammalian reticulocytes, <i>in vitro</i>) >20 h (yeast, <i>in vivo</i>) >10 h (<i>Escherichia coli</i> , <i>in vivo</i>)	117.53	0.386	9.26	1,015	0.4101

TABLE 2 Predicted B-cell epitopes of both protein.

Protein	B Cell epitopes	Antigenic score
Lipopolysaccharide export system permease protein LptF	AKANPGMAALAQQGQFQQAT	0.442
	IGHQVAALDPNDTD	0.5790
	QMDMRTLWNTDTDRA	0.6926
	VNPRQGRVLS	0.9061
	NGGKGKLDPT	2.024
	DTVPVRRLRASFSSR	0.58
Lipoprotein-releasing ABC transporter permease subunit LolE	RGRRRGGM	1.0031
	LSGQLDH	0.8104

2.12 In-silico cloning

The *Escherichia coli* K12 strain is widely utilized as the prokaryotic expression vector for *in silico* cloning purposes. To optimize the codon usage for this vector, the Java Codon Adaptation Tool (JCAT) (<http://www.jcat.de/>) is employed. JCAT calculates the Codon Adaptation Index (CAI) score and GC content of the improved sequence, ensuring a better fit for expression in the chosen prokaryotic host. After codon optimization, the improved sequence is then inserted into the recombinant plasmid pET-28a (+) using SnapGene software. SnapGene facilitates visualization and manipulation of DNA sequences, allowing for seamless and efficient virtual cloning procedures (Jalal et al., 2022).

3 Results

3.1 Target protein identification

Filtering out 34 membrane proteins from CELLO database, an extensive analysis was conducted to evaluate their hydrophobicity, physicochemical properties, antigenicity and membrane topology scores. Notably, two proteins, Lipopolysaccharide export system permease protein LptF and Lipoprotein-releasing ABC transporter permease subunit LolE, exhibited remarkable stability. Moreover, both proteins displayed a positive GRAVY (Grand Average of Hydropathy) index, indicating their hydrophobic nature (Table 1).

3.2 B cell epitope prediction

The prediction of B-cell epitopes was carried out using Immune Epitope Database (IEDB) (<http://tools.iedb.org/bcell/>). The predicted B-cell epitopes from both proteins were then subjected to analysis and shortlisting based on their antigenicity, allergenicity, and toxicity. These epitopes ranged in length from 8 to 19 amino acids (mers) (Table 2). There was a total of 6 epitopes for the Lipopolysaccharide export system permease protein LptF and 2 epitopes for the Lipoprotein-releasing ABC transporter permease subunit LolE protein.

3.3 T cell epitope prediction

In this study, the T cell epitopes of both proteins were predicted using IDEB tepi tool database (<http://tools.iedb.org/tepitool/>). The prioritized epitopes are about 9-11-mers in length (Supplementary Table S3A). There were 12 MHC Class I epitopes and 7 MHC Class II epitopes identified for the Lipopolysaccharide export system permease protein LptF. Additionally, 17 MHC Class I epitopes and 14 MHC Class II epitopes were found for the Lipoprotein-releasing ABC transporter permease subunit LolE (Supplementary Table S3B).

3.4 Population coverage analysis

Combinatorial methods were employed to calculate the population coverage of a specific set of T-cell epitopes, which included epitopes of MHC class I and II T cells, along with their corresponding HLA alleles that bind to them (Sarkar et al., 2022). Approximately, 85.72% of the global population is expected to exhibit a response to the chosen epitopes (Figure 1).

3.5 Construction of vaccine

In the process of developing the vaccine construct, the B-cell and T-cell epitopes predicted earlier underwent further analysis. To mitigate the risk of autoimmune errors, any over-lapping sequences among the leading peptides were carefully removed (Yano et al., 2003). The shortlisted peptides for B-cell epitopes comprised 8 amino acids, while the MHC class I epitopes consisted of 29 amino acids, and MHC class II epitopes were 21 amino acids. Subsequently, the final vaccine construct was assembled by merging all these peptides together and incorporating adjuvants and linkers. At the N-terminal, the adjuvant RGD (arginine-glycine-aspartate), along with GGGs, KK, and AAY linkers, were utilized to connect the peptides. Furthermore, thrombin cleavage and 6-his tags were added at the C-terminal end. As a result, the vaccine construct consisted of 372 amino acid residues in total (Figure 2).

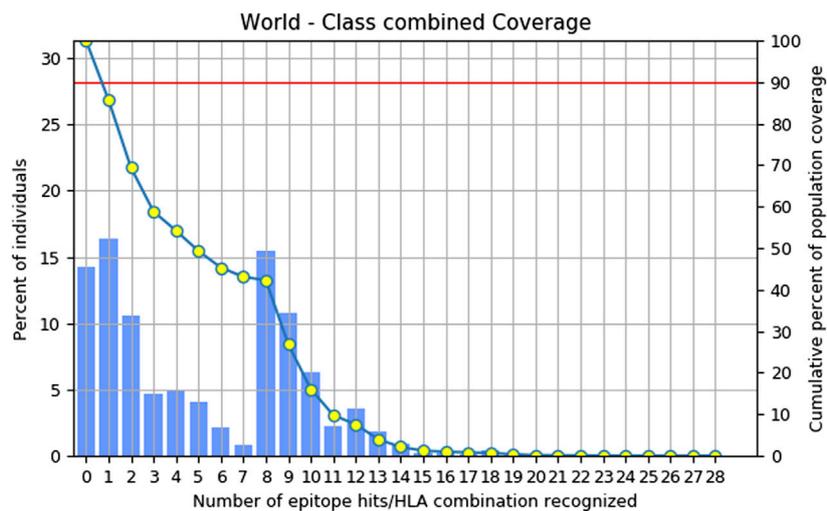


FIGURE 1 Combinatorial graph of population coverage of prioritized T. cell epitopes.

RGDEAAAKAKANPGMAALAQQQFQQATKKIGHQAVALDPNDTDKKQMDMRTLWNTDTRAKKVNPRQGRVLSKKDTPVRRRLRASFSRKKRGRRRGGMKKLSGQLDHAAYLVRETLKSQQLAIFIAAYACGLSKAVLVKAAMIAYLLLIFFCQKLVRLGAAYCGLSKAVLVKAAMILAAYVKAAMILAVFTAIVAAYQPKRVRHLIAGILQLAAYQMIRAIMYLMVLIAYYINFTGLVESGANLRGGGSIVAAVNMWGGSSFLGLLMTLGGGSWDTVPVRRRLGGSNLVSLLGLGGSRLSALPSFVGGGSVTNSYVYIKSWGGSSEQRLSALPSFGGSVLVIGVACFLVPRGSHHHHHH

FIGURE 2 The vaccine construct contains a comprehensive sequence of amino acids. At the N terminus of the construct, an RGD adjuvant has been incorporated. Distinctive linkers (EAAAK, AAY, KK, and GGS) were introduced to differentiate specific epitopes within the construct.

3.6 Secondary structure prediction and structure modelling of vaccine construct

PDBsum analysis of vaccine component revealed that it comprises 2 beta sheets, 5 beta hairpin structures, and 15 helices (Figure 3). Meanwhile, trRosetta was utilized to predict the best vaccine construct. The Ramachandran plot demonstrated that approximately 91.5% of residues were in the most favored region, and an additional 7.5% of residues were in the allowed region, indicating good structural quality (Figure 4). The ERRAT score, which was 95.238, further validated the excellent quality of the selected model (Figure 5) (Jalal et al., 2022). Visualization of chosen model was performed using USF CHIMERA (Figure 6).

3.7 Docking interaction of vaccine construct with human toll like receptor 4

Before conducting the docking process between the vaccine construct and human TLR-4. For this purpose, Castp database was employed to identify the binding sites for vaccine construct. While the binding site for TLR-4 was predicted on C-port. The database not only provided information about the binding sites but

also highlighted the active and passive residues of the TLR-4 protein molecule, which are likely to interact with vaccine construct.

After the prediction of active site, the vaccine model was docked with human TLR-4 on online server called HADDOCK. The docking is performed in order to explore the binding affinity, the interactions, and conformation of the docked complex. The docked interactions were visualized on CHIMERA (Figure 7). A negative binding energy provided evidence of a strong and robust affinity between the vaccine and TLR-4 during the interaction process. This negative value indicates a favorable and strong binding affinity between the vaccine and TLR-4. The docked complex was submitted on PDBsum for further analysis, this database revealed that there are 10 binding interactions between the vaccine construct and human TLR-4 (Figure 8). The statistical analysis of HADDOCK docked complex (Table 3). According to PRODIGY predictions, the binding affinity of the docked complex was characterized by a Gibbs free energy (ΔG) of $-9.0 \text{ kcal mol}^{-1}$ and dissociation constant (Kd) values of $2.7 \times 10^{-7} \text{ M}$. These results indicate a highly favorable and tight binding between the vaccine and TLR-4 in the complex. The low ΔG value and the small Kd value suggest that the binding between the vaccine and TLR-4 is strong and stable, reinforcing the potential effectiveness of this interaction in the immune response.

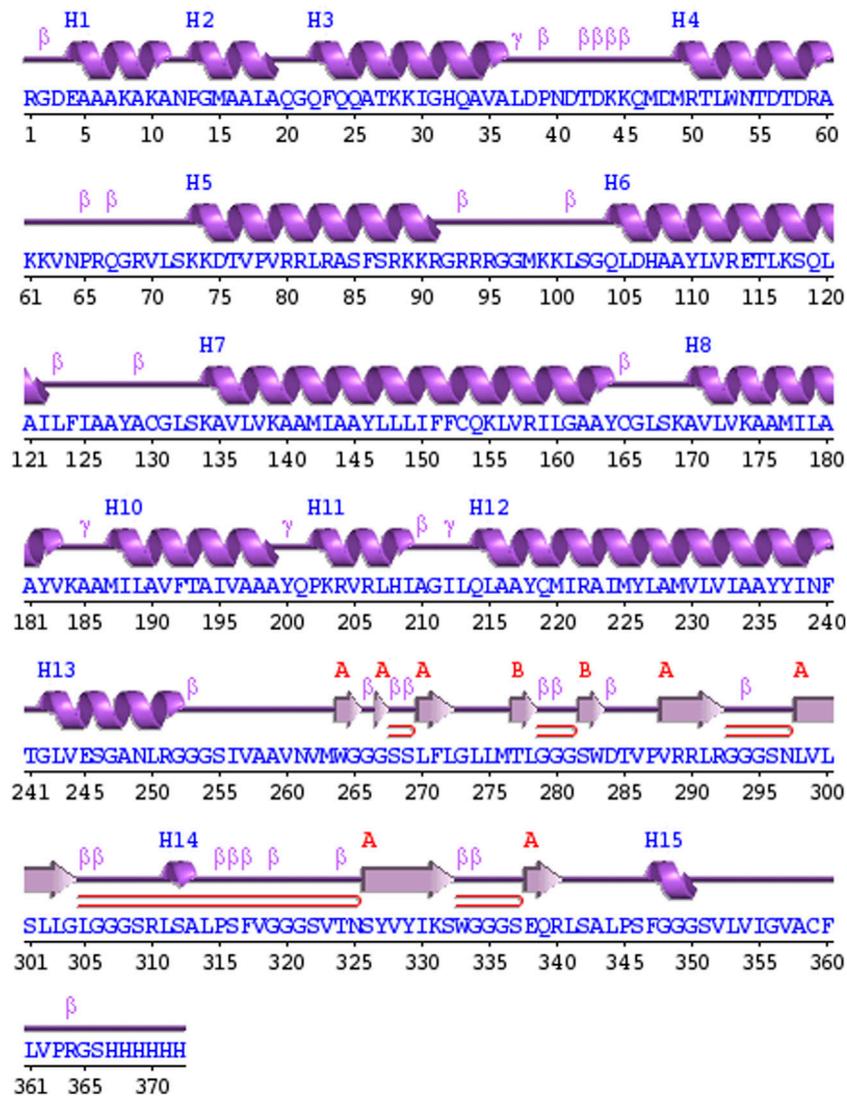


FIGURE 3
 A schematic representation illustrates the secondary structure achieved for the ultimate multi-epitope vaccine. The α -helices are denoted as H, while the β -sheets are indicated by labels A and B. The red loops specifically highlight the β -hairpins.

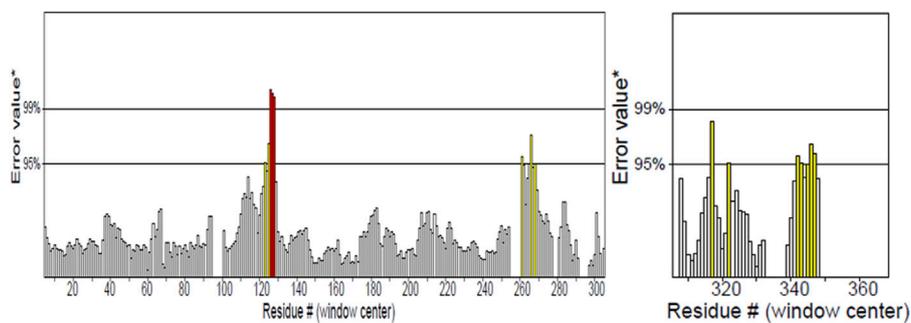


FIGURE 4
 ERRAT showed 95.238 score of vaccine construct, the labeled axis indicating "error" displays two lines that indicate the level of certainty we have in excluding regions where the error values surpass the specified threshold. Residues displaying error values surpassing 99% were identified and marked using the color red. Conversely, areas that exhibit an error value greater than 95% are indicated using the color yellow.

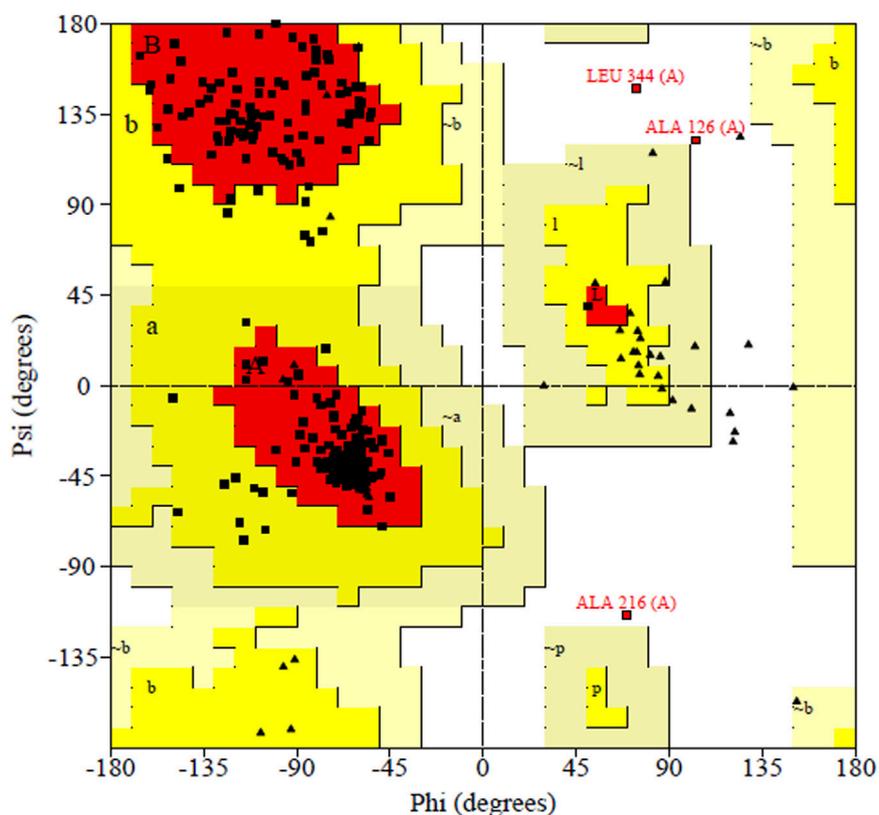


FIGURE 5
Ramachandran plot shown the most favored regions and additional allowed regions the, the increased red color indicates the residues are present in most favored region.

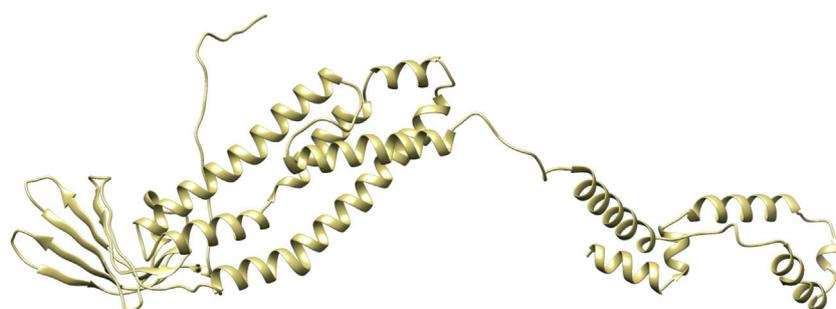


FIGURE 6
Vaccine construct in 3D visualized on CHIMERA.

3.8 Normal mode analysis of docked complex

iMODs employs Normal Mode Analysis (NMA) to assess the molecular mobility and structural flexibility of docked complex. The hypothesis behind NMA of proteins is that the lowest frequency vibrational normal modes (referred to as “soft modes”) describes the most significant movements in a protein and are therefore functionally relevant. The study of biomolecular dynamics has been greatly facilitated by normal

mode analysis, which is a dependable and effective method. Additionally, it is particularly well-suited for investigating slow dynamics and large conformational changes in protein (Skjaerven et al., 2009; Soltan et al., 2021). The deformability of the complex is primarily determined by the individual distortion of each residue’s C α atom (alpha-carbon), which is represented by green colored hinges in the chain (Figure 9A). The higher peak indicates higher deformability (Figure 9B) indicated the eigenvalue 3.925996e-07. It relates to the stiffness of the protein structure; lowest eigenvalue indicated easier

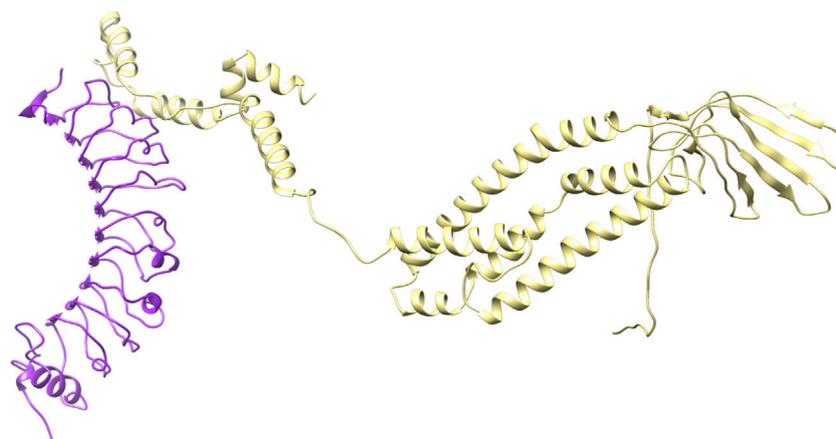


FIGURE 7
3D structure of docked complex.

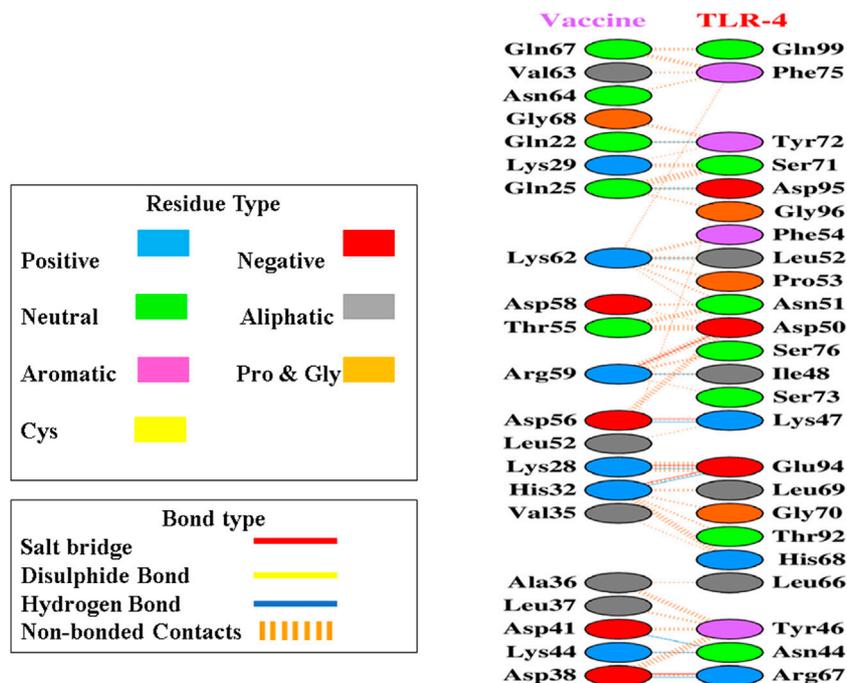


FIGURE 8
Upon docking, the complex exhibited a considerable number of non-bonded contacts between the vaccine and TLR-4 depicted as dashed lines, along with some noteworthy hydrogen bonds and salt bridges, disulphides bonds as well. Key hydrogen bonds are highlighted in blue, while salt bridges are denoted by red lines. Furthermore, identification of the participating amino acid residues categorized them as aliphatic, aromatic, positive, and negative within these interactions.

deformability. The motion between two residues correlates indicated in red color, anti-correlated motions were in blue color, uncorrelated motions were in white color (Figure 9C). In the elastic network model, the connected residues pairs are represented by strings. The dot on the graphs indicates that single spring coinciding to it residues pair, with darker gray indicating a stiffer spring. The stiffness of the springs is represented by the dots (Figure 9D).

3.9 Codon optimization and *in silico* cloning

In silico cloning was used to determine how our multi-epitope vaccine construct will express on *E. coli* hosts. The vaccine construct was optimized for expression in *E. coli* (K12 strain) using the Java Codon Adaptation (JCat) tool. To mitigate potential errors, we applied two filtering parameters. Firstly, we excluded rho-independent transcription terminators

TABLE 3 The statistical analysis of docked complex.

HADDOCK score	-122.6+/4.0
Cluster size	112
RMSD from the overall lowest-energy structure	28.5+/-0.1
Van der Waals energy (kcal mol ⁻¹)	-71.5+/-12.4
Electrostatic energy (kcal mol ⁻¹)	-349.7+/-57.4
Desolvation energy (kcal mol ⁻¹)	-10.2+/-4.4
Restraints violation energy (kcal mol ⁻¹)	290.2+/-76.4
Buried surface area (Å ²)	2,171.6+/-25.5
Z-score	-1.5

to maintain the integrity of transcriptional processes and prevent premature termination. These terminators are capable of terminating transcription (Łobocka et al., 2004). Secondly, we avoided prokaryotic ribosome binding sites to prevent interference with translation initiation and ensure accurate protein synthesis (Shine and Dalgarno, 1974). These errors can lead to incomplete or inefficient expression of the vaccine construct. The optimized gene had 1,100 nucleotides. The CAI value of optimized vaccine construct was 0.92 (range must be of 0.8–1.0) this value is deemed acceptable as it surpasses the cutoff threshold of 0.8. Furthermore, the G.C. content of the optimized construct was assessed. It was determined that an optimal G.C. content percentage of approximately 30%–70% is preferable for efficient expression in the *E. coli*. The average G.C. percentage of the optimized construct was calculated to be 53.118% (Dey et al.,

2022). At end N and C termini of optimized construct, two restricted enzymes (*EcoRI* and *BamHI*) were inserted to facilitate the cloning procedure. We used the SnapGene software to clone the vaccine construct into the pET-28a (+) vector. The size of complete clone was 6.486 kb (Figure 10).

4 Discussion

Shigella spp. are Gram-negative, rod-shaped bacteria that lack motility and spores. They are responsible for diarrhea, which can advance to bloody mucoid diarrhea, commonly referred to as bacillary dysentery or shigellosis (Kahsay and Muthupandian, 2016). While antibiotic therapy can mitigate the disease, but the rising prevalence of antibiotic resistance in *Shigella* isolates underscores the need for enhanced preventive measures, such as improved hygiene, access to clean water, and the development of pathogen-specific immunity (Walker et al., 2021). Recent studies have shown that *Shigella* have taken the lead of over 200,000 fatalities, which is more common in low-income countries globally (Nuzhat et al., 2022). Even after a century of vaccine development, there remains no licensed vaccine to date for safeguarding against shigellosis, which is the important cause of fatal bacterial diarrhea in children worldwide (MacLennan et al., 2022).

The *in silico* vaccine has gained significant acceptance and due to its duration of development and cost efficiency, it is commonly utilized as the primary stage in vaccine design (Tahir et al., 2019). The *in silico* approach can easily predict whether vaccine construct can induce a strong immune response and overcome major challenges in vaccine development, such as antigenic shifts, genetic dissimilarities. This method offers a

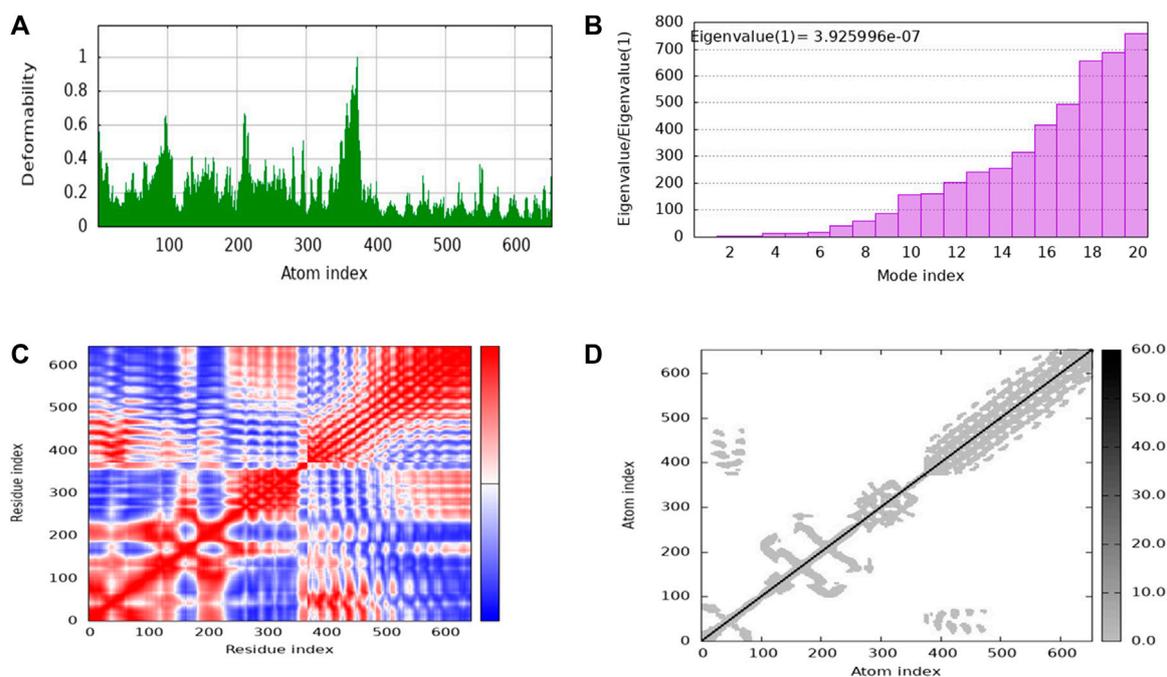


FIGURE 9

The stability evaluation of the protein-protein complex between the multi-epitope vaccine and TLR-4 utilized molecular dynamics simulation. This analysis encompassed the exploration of (A) Deformability, (B) Eigenvector, (C) Covariance of residue index and (D) elastic network characteristics.

After meeting all the required criteria, the peptides were linked together using adjuvants and specific linkers. The adjuvants used in this process were RGD (arginine–glycine–aspartate). The data provided clear evidence that peptides with the RGD motif at the N-terminal side of the -KK- linker significantly induced the production of peptide anti-bodies on the C-terminal side of the -KK- linker (Zamyatina and Heine, 2020), along with GGGs, KK, and AAY linkers, which played a crucial role in facilitating the effective joining of the peptides to form the final multivalent epitope-based vaccine. Through the comprehensive analyses, we designed a vaccine construct that consists of 372 amino acids. The projected vaccine exhibited desirable characteristics, being stable, soluble, highly antigenic, and non-allergenic. These features make it a promising candidate for potential therapeutic applications in combating the target pathogen. Following the design of the vaccine construct, a 3D model was created and its quality was verified by generating a Ramachandran plot. The majority of residues (91.5%) were found within the favored or allowed regions of the plot, indicating that the model exhibited good quality and stability. Furthermore, the ERRAT score for the model was calculated to be 95.238, further confirming the reliability and accuracy of the constructed vaccine model.

Once the 3D model of the vaccine was generated, a molecular docking analysis was conducted to investigate the interactions between the vaccine and the Toll-like receptor molecule 4 (TLR-4). Previous studies have shown the strong activation of TLR-4 and the production of cytokines. The vaccine demonstrated an interaction with TLR-4, exhibiting negative energy in the binding process. This negative binding energy signifies a stronger binding affinity between the vaccine and TLR-4 (Adhikari et al., 2018). It has been documented that TLR4, a receptor found on immune cells, detects lipopolysaccharide (LPS) molecules derived from Gram-negative bacteria. LBP, a protein present in the bloodstream, aids in the transportation of LPS to CD14, which can be found in either soluble or membrane-bound forms on the cell surface. Upon binding to LPS, CD14 transfers it to the TLR4/MD-2 receptor complex, initiating a conformational change in TLR4 and leading to the formation of receptor pairs. This activation triggers intracellular signaling pathways, culminating in the synthesis and release of proinflammatory cytokines such as interleukins (e.g., IL-1 and IL-6) and tumor necrosis factor (TNF), which play crucial roles in the immune response against bacterial infections (Mazgaen and Gurung, 2020).

The NMA study was conducted to analyze the molecular mobility and comparative deformability of the system. Notably, the calculated eigenvalue of suggests that the docking complex possesses enhanced flexibility. The NMA study was conducted to assess the molecular mobility and comparative deformability of the docked complex. The calculated eigenvalue, 3.925996e-07 indicates that the complex exhibits improved flexibility.

The study verified that the interactions between the vaccine and its targets were balanced. Additionally, *in silico* cloning was successfully carried out, indicating the potential for efficient expression in *E. coli* for large-scale production of the vaccine. Nevertheless, it is essential to validate these findings through *in-vivo* studies. Collectively, these studies suggest that the vaccine

candidate holds promise as a prevented therapy against *S. dysenteriae* infections. The significant proteins identified in this study provide a basis for developing a vaccine model aimed at eliciting an immune response against *S. dysenteriae*. Nevertheless, validating the stimulation of a robust immune response against *S. dysenteriae* requires *in vivo* expression and evaluation of the multiepitope vaccine.

5 Conclusion

Globally, diarrheal diseases resulting from *Shigella* species are linked to both mortality and long-term dysfunction. This bacterium also tends to get resistant to major antibiotics. Hence, an effective vaccine design is needed cope up with this problem. The *in silico* immuno-informatics approach is demonstrated to develop a multiepitope vaccine which can elicit both innate and adaptive immunity. The vaccine is constructed by predicting the B and T-cells, the antigenicity, allergenicity and toxicity are assessed. The interaction of 3D structure of vaccine construct and TLR-4 showed good binding affinity. Finally, we conducted *in silico* cloning to assess the expression of the final vaccine construct. Our *in silico* approach validates the designed vaccine can stimulate immune response against *S. dysenteriae*. There is a crucial need to carry out *in-vitro* and *in-vivo* studies to confirm its efficacy, immune stimulation.

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

Author contributions

HQ: Conceptualization, methodology, software, validation, formal analysis, investigation, data curation, writing–original draft preparation, writing–review and editing, visualization. AB: validation, data curation, writing–review and editing. MF: Conceptualization, writing–review and editing, visualization, supervision. MWA: methodology, writing–review and editing. SKR resources, writing–review and editing. SBJ: Conceptualization, resources, supervision, project administration, Writing–review and editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

We acknowledge, National University of Medical Sciences for supporting the study.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2024.1361610/full#supplementary-material>

References

- Adhikari, U. K., Tayebi, M., and Mizanur Rahman, M. (2018). Immunoinformatics approach for epitope-based peptide vaccine design and active site prediction against polyprotein of emerging oropouche virus. *J. Immunol. Res.* 2018, 6718083. doi:10.1155/2018/6718083
- Ahamed, S. T., and Giria, N. (2021). Shigellosis and development of multiple antimicrobial resistance mechanisms of *Shigella* spp. *Biosci. Biotechnol. Res. Asia* 18 (4), 703–718. doi:10.13005/bbra/2953
- Akhtar, N., Joshi, A., Kaushik, V., Kumar, M., and Amin-ul Mannan, M. (2021). In-silico design of a multi-valent epitope-based vaccine against *Candida auris*. *Microb. Pathog.* 155, 104879. doi:10.1016/j.micpath.2021.104879
- Basheer, A., Jamal, S. B., Alzahrani, B., and Faheem, M. (2023). Development of a tetravalent subunit vaccine against dengue virus through a vaccinomics approach. *Front. Immunol.* 14, 1273838. doi:10.3389/fimmu.2023.1273838
- Desalegn, G., Tamilselvi, C. S., Lemme-Dumit, J. M., Heine, S. J., Dunn, D., Ndungo, E., et al. (2024). *Shigella* virulence protein VirG is a broadly protective antigen and vaccine candidate. *npj Vaccines* 9 (1), 2. doi:10.1038/s41541-023-00797-6
- Dey, J., Mahapatra, S. R., Patnaik, S., Lata, S., Kushwaha, G. S., Panda, R. K., et al. (2022). Molecular characterization and designing of a novel multi-epitope vaccine construct against *Pseudomonas aeruginosa*. *Int. J. Peptide Res. Ther.* 28 (2), 49. doi:10.1007/s10989-021-10356-z
- Farsiani, H., and Sasan, M. S. (2020). High rate of resistance to ceftriaxone and azithromycin among *Shigella* spp. isolates at three children's referral hospitals in Northeast Iran. *J. Infect. Chemother.* 26 (9), 955–958. doi:10.1016/j.jiac.2020.04.022
- Jalal, K., Abu-Izneid, T., Khan, K., Abbas, M., Hayat, A., Bawazeer, S., et al. (2022). Identification of vaccine and drug targets in *Shigella dysenteriae* sd197 using reverse vaccinology approach. *Sci. Rep.* 12 (1), 251. doi:10.1038/s41598-021-03988-0
- Jamal, S. B., Hassan, S. S., Tiwari, S., Viana, M. V., Benevides, L. D., Ullah, A., et al. (2017). An integrative in-silico approach for therapeutic target identification in the human pathogen *Corynebacterium diphtheriae*. *PLoS One* 12 (10), e0186401. doi:10.1371/journal.pone.0186401
- Kahsay, A. G., and Muthupandian, S. (2016). A review on Sero diversity and antimicrobial resistance patterns of *Shigella* species in Africa, Asia and South America, 2001–2014. *BMC Res. notes* 9 (1), 422–426. doi:10.1186/s13104-016-2236-7
- Khamjan, N. A., Lohani, M., Khan, M. F., Khan, S., and Algaissi, A. (2023). Immunoinformatics strategy to develop a novel universal multiple epitope-based COVID-19 vaccine. *Vaccines* 11 (6), 1090. doi:10.3390/vaccines11061090
- Kumar, S., and Omeershffudin, U. N. (2020). In silico functional and toxicity prediction of *Shigella dysenteriae* sd197 hypothetical proteins. *Int. J. Med. Toxicol. Leg. Med.* 23 (1and2), 246–250. doi:10.5958/0974-4614.2020.00035.2
- Lobocka, M. B., Rose, D. J., Plunkett, I. I. G., Rusin, M., Samojedny, A., Lehnerr, H., et al. (2004). Genome of bacteriophage P1. *J. Bacteriol.* 186 (21), 7032–7068. doi:10.1128/JB.186.21.7032-7068.2004
- Lundstedt, E., Kahne, D., and Ruiz, N. (2020). Assembly and maintenance of lipids at the bacterial outer membrane. *Chem. Rev.* 121 (9), 5098–5123. doi:10.1021/acs.chemrev.0c00587
- MacLennan, C. A., Grow, S., Ma, L. F., and Steele, A. D. (2022). The *Shigella* vaccines pipeline. *Vaccines* 10 (9), 1376. doi:10.3390/vaccines10091376
- Mahmoodi, S., Amirzakaria, J. Z., and Ghasemian, A. (2023). In silico design and validation of a novel multi-epitope vaccine candidate against structural proteins of Chikungunya virus using comprehensive immunoinformatics analyses. *Plos one* 18 (5), e0285177. doi:10.1371/journal.pone.0285177
- Mazgaen, L., and Gurung, P. (2020). Recent advances in lipopolysaccharide recognition systems. *Int. J. Mol. Sci.* 21 (2), 379. doi:10.3390/ijms21020379
- Nasser, A., Mosadegh, M., Azimi, T., and Shariati, A. (2022). Molecular mechanisms of *Shigella* effector proteins: a common pathogen among diarrheic pediatric population. *Mol. Cell. Pediatr.* 9 (1), 12–21. doi:10.1186/s40348-022-00145-z
- Nuzhat, S., Das, R., Das, S., Islam, S. B., Palit, P., Haque, M. A., et al. (2022). Antimicrobial resistance in shigellosis: a surveillance study among urban and rural children over 20 years in Bangladesh. *PLoS one* 17 (11), e0277574. doi:10.1371/journal.pone.0277574
- Okuda, S., and Tokuda, H. (2009). Model of mouth-to-mouth transfer of bacterial lipoproteins through inner membrane LolC, periplasmic LolA, and outer membrane LolB. *Epub* 106 (14), 5877–5882. doi:10.1073/pnas.0900896106
- Oyarzún, P., and Kobe, B. (2016). Recombinant and epitope-based vaccines on the road to the market and implications for vaccine design and production. *Hum. vaccines Immunother.* 12 (3), 763–767. doi:10.1080/21645515.2015.1094595
- Pakbin, B., Didban, A., Brück, W. M., and Alizadeh, M. (2022). Phylogenetic analysis and antibiotic resistance of *Shigella sonnei* isolates. *FEMS Microbiol. Lett.* 369 (1), fnac042. doi:10.1093/femsle/fnac042
- Paul, S., Sidney, J., Sette, A., and Peters, B. (2016). TepiTool: a pipeline for computational prediction of T cell epitope candidates. *Curr. Protoc. Immunol.* 114 (1), 1–18. doi:10.1002/cpim.12
- Qureshi, N. A., Bakhtiar, S. M., Faheem, M., Shah, M., Bari, A., Mahmood, H. M., et al. (2021). Genome-based drug target identification in human pathogen *Streptococcus galloyticus*. *Front. Genet.* 12, 564056. doi:10.3389/fgene.2021.564056
- Ranjbar, R., and Farahani, A. (2019). *Shigella*: antibiotic-resistance mechanisms and new horizons for treatment. *Infect. drug Resist.* 7, 3137–3167. doi:10.2147/IDR.S219755
- Raso, M. M., Arato, V., Gasperini, G., and Nicoli, F. (2023). Toward a *Shigella* vaccine: opportunities and challenges to fight an antimicrobial-resistant pathogen. *Int. J. Mol. Sci.* 24 (5), 4649. doi:10.3390/ijms24054649
- Riasat, I., Bakhtiar, S. M., Faheem, M., Jaiswal, A. K., Naeem, M., Khan, R., et al. (2021). Application of pan genomics towards the druggability of *Clostridium botulinum*. *Appl. Nanosci.* 12, 3237–3249. doi:10.1007/s13204-021-02005-0
- Rouzbahani, A. K., Kheirandish, F., and Hosseini, S. Z. (2022). Design of a multi-epitope-based peptide vaccine against the S and N proteins of SARS-COV-2 using immunoinformatics approach. *Egypt. J. Med. Hum. Genet.* 23 (1), 16. doi:10.1186/s43042-022-00224-w
- Saha, C. K., Hasan, Md M., Hossain, Md S., Jahan, Md A., and Azad, A. K. (2017). In silico identification and characterization of common epitope-based peptide vaccine for Nipah and Hendra viruses. *Asian Pac. J. Trop. Med.* 10 (6), 529–538. doi:10.1016/j.apjtm.2017.06.016
- Sanches, R. C., Tiwari, S., Ferreira, L. C., Oliveira, F. M., Lopes, M. D., Passos, M. J., et al. (2021). Immunoinformatics design of multi-epitope peptide-based vaccine against *Schistosoma mansoni* using transmembrane proteins as a target. *Front. Immunol.* 12, 621706. doi:10.3389/fimmu.2021.621706
- Sarkar, B., Ullah, Md A., Araf, Y., Islam, N. N., and Zohora, U. S. (2022). Immunoinformatics-guided designing and in silico analysis of epitope-based polyvalent vaccines against multiple strains of human coronavirus (HCoV). *Expert Rev. vaccines* 21 (12), 1851–1871. doi:10.1080/14760584.2021.1874925
- Scott, M. S., Calafell, S. J., Thomas, D. Y., and Hallett, M. T. (2005). Refining protein subcellular localization. *PLoS Comput. biol.* 1 (6), e66. doi:10.1371/journal.pcbi.0010066
- Shine, J., and Dalgarno, L. (1974). The 3'-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. *Proc. Natl. Acad. Sci.* 71 (4), 1342–1346. doi:10.1073/pnas.71.4.1342
- Skjaerven, L., Siv, M., and Reuter, N. (2009). Normal mode analysis for proteins. *J. Mol. Struct. THEOCHEM* 898 (1-3), 42–48. doi:10.1016/j.theochem.2008.09.024
- Soltan, M. A., Elbassiouny, N., Gamal, H., Elkaeed, E. B., Eid, R. A., Eldeen, M. A., et al. (2021). In silico prediction of a multitope vaccine against *Moraxella catarrhalis*: reverse vaccinology and immunoinformatics. *Vaccines* 9 (6), 669. doi:10.3390/vaccines9060669

- Tahir, ul Q., Muhammad, S. S., Ali Ashfaq, U., Bari, A., Anwar, F., and Alqahtani, S. (2019). Epitope-based peptide vaccine design and target site depiction against Middle East Respiratory Syndrome Coronavirus: an immune-informatics study. *J. Transl. Med.* 17, 1–14. doi:10.1186/s12967-019-2116-8
- Turbyfill, K. R., Clarkson, K. A., Oaks, E. V., and Kaminski, R. W. (2022). From concept to clinical product: a brief history of the novel shigella invaplax vaccine's refinement and evolution. *Vaccines* 10 (4), 548. doi:10.3390/vaccines10040548
- Walker, R., Kaminski, R. W., Porter, C., Choy, R. K., White, J. A., Fleckenstein, J. M., et al. (2021). Vaccines for protecting infants from bacterial causes of diarrheal disease. *Microorganisms* 9 (7), 1382. doi:10.3390/microorganisms9071382
- Yano, A., Onozuka, A., Matin, K., Imai, S., Hanada, N., and Nisizawa, T. (2003). RGD motif enhances immunogenicity and adjuvanicity of peptide antigens following intranasal immunization. *Vaccine* 22 (2), 237–243. doi:10.1016/s0264-410x(03)00561-9
- Yu, C. S., Cheng, C. W., Su, W. C., Chang, K. C., Huang, S. W., Hwang, J. K., et al. (2014). CELLO2GO: a web server for protein subCELLular LOcalization prediction with functional gene ontology annotation. *PLoS one* 9 (6), e99368. doi:10.1371/journal.pone.0099368
- Zamyatina, A., and Heine, H. (2020). Lipopolysaccharide recognition in the crossroads of TLR4 and caspase-4/11 mediated inflammatory pathways. *Front. Immunol.* 11, 585146. doi:10.3389/fimmu.2020.585146