Research Article

Bioinformatics and Biology Insights
Volume 19: 1–20

DOI: 10.1177/11779322251328302
journals.sagepub.com/home/bbi
Sage

Development of a Novel mRNA Vaccine Against *Shigella* Pathotypes Causing Widespread Shigellosis Endemic: An In-Silico Immunoinformatic Approach

Abdur Razzak[1,*], Otun Saha[1,*], Khandokar Fahmida Sultana[1], Mohammad Ruhul Amin[1], Abdullah bin Zahid[1], Afroza Sultana[1], Uditi Paul Bristi[1], Sultana Rajia[1], Nikkon Sarker[1], Md Mizanur Rahaman[2], Newaz Mohammed Bahadur[3] and Foysal Hossen[1]

Abstract

Shigellosis remains a major global health concern, particularly in regions with poor sanitation and limited access to clean water. This study used immunoinformatics and reverse vaccinology to design a potential mRNA vaccine targeting *Shigella* pathotypes out of 4071 proteins from *Shigella sonnei* str. Ss046, 4 key antigenic candidates were identified: putative outer membrane protein (Q3YZL0), PapC-like porin protein (Q3YZM5), putative fimbrial-like protein (Q3Z3I2), and lipopolysaccharide (LPS)-assembly protein LptD (Q3Z5V5), ensuring broad pathotype coverage. A multitope vaccine was designed incorporating cytotoxic T lymphocyte, helper T lymphocyte, and B-cell epitopes, linked with suitable linkers and adjuvants to enhance immunogenicity. Computational analyses predicted vaccine's favorable antigenicity, solubility, and stability, while molecular docking and dynamic simulations demonstrated strong binding affinity and stability with Toll-like receptor 4 (TLR-4), indicating potential for robust immune activation. Immune simulations predicted strong humoral and cellular immune responses, characterized by significant cytokine production and long-term immune memory. Structural evaluations of the complex, including radius of gyration, root mean square deviation, root mean square fluctuation, and solvent accessibility, confirmed the vaccine's structural integrity, and stability under physiological conditions. This research contributes to the ongoing effort to alleviate the global burden of *Shigella* infections, providing a foundation for future wet laboratory investigations aimed at vaccine development.

Keywords

Shigella species, mRNA, reverse vaccinology, multitope vaccine, molecular docking, MDS

Received: 29 June 2024; accepted: 3 March 2025

Introduction

Bacillary dysentery, caused by the gram-negative bacterium *Shigella*, is a potentially fatal type of diarrhea that may induce severe dehydration.[1] *Shigella dysenteriae*, *S. flexneri*, *S. boydii*, and *S. sonnei* are the 4 most common pathotypes, and they all have unique serotypes and antigenic profiles responsible for disease causation.[2] Infections caused by *S. dysenteriae* and *S. boydii* are uncommon in most parts of the globe. Sub-Saharan Africa and South Asia are common regions for isolating *S. dysenteriae*, whereas the Indian subcontinent is the most common region for isolating *S. boydii*.[3] *S. sonnei* accounts for up to 80% of cases of shigellosis in developed countries,[4] but it has recently

[1]Department of Microbiology, Noakhali Science and Technology University, Noakhali, Bangladesh
[2]Department of Microbiology, University of Dhaka, Dhaka, Bangladesh
[3]Department of Chemistry, Noakhali Science and Technology University, Noakhali, Bangladesh

*Equal Contribution.

Corresponding authors:
Otun Saha, Department of Microbiology, Noakhali Science and Technology University, University road, Noakhali-3814, Bangladesh.
Email: otun.saha@nstu.edu.bd

Foysal Hossen, Department of Microbiology, Noakhali Science and Technology University, Noakhali-3814, Bangladesh.
Email: foysal@nstu.edu.bd

surpassed *S. flexneri* as the most common infectious species in low-income countries with poor access to water, sanitation, and health care.[5,6] The reason behind the historical presence of *S. sonnei* in high-income nations and its recent emergence in low-income countries remains unclear.[7] One possible explanation is natural passive immunization through *Plesiomonas shigelloides*, a bacterium that shares an identical lipopolysaccharide (LPS) O-side chain, which previously conferred protection against *S. sonnei*.[7] However, as economic conditions improve, this protective effect may diminish, contributing to the rising incidence of shigellosis.[8] *S. flexneri* instances have decreased while *S. sonnei* cases have increased, mostly in quickly growing nations in Asia, Latin America, and the Middle East.[7,9-13] Because of its remarkable adaptability and arsenal of antibiotic-resistance genes,[14] *S. sonnei* has also emerged as a major public health danger in recent years. Worldwide, there are an estimated 164.7 million cases of *Shigella* infection each year, with an estimated 1.1 million fatalities, most of them occurring in children less than 5 years old.[15,16] Approximately 1.5 million new cases of shigellosis occur annually in industrialized nations, with the United States accounting for about a quarter of these cases. *S. sonnei* is the predominant strain in the United States, responsible for 77% of all reported infections.[1] Noteworthy vulnerable populations include children, the elderly, and individuals with compromised immune systems.[17] Travelers visiting impoverished nations, members of the armed forces, refugees, and people in institutional settings are also at risk for contracting shigellosis.[18-20] This widespread infection rate is a major contributor to both the global public health concern and the global economic disruption that follows.[7,9,10,15] Antibiotic resistance and a dearth of effective vaccinations have severely constrained existing methods of treating *S.* infection.[21] In response, considerable research efforts have focused on vaccine development, exploring various strategies such as killed whole-cell vaccines, live-attenuated vaccines, subunit vaccines, conjugate vaccines, outer membrane vesicle (OMV) vaccines, and more recently, mRNA vaccines.[22] Despite the advancement of several candidates to clinical trials, including Phase 3, no vaccine has been licensed to date.[23] Challenges in achieving broad, long-lasting protection across diverse Shigella serotypes have hampered progress.[23,24] ZF0901 (Beijing Zhifei Lvzhu Biopharmaceuticals) is a bivalent glycoconjugate vaccine targeting *S. sonnei* and *S. flexneri* 2a, currently in Phase 3 trials.[25] S4V2 (Valneva *SE* and LimmaTech) is a tetravalent bioconjugate vaccine targeting *S. flexneri* 2a, 3a, 6, and *S. sonnei*, undergoing Phase 2 trials in LMICs and a Phase 2b controlled human infection model (CHIM) study in the United States, with Fast Track designation from the U.S. FDA (NCT06615375). Invaplex AR-Detox (WRAIR), a subunit vaccine combining *Shigella* LPS with Ipa proteins, has shown good safety and immunogenicity in Phase 1 trials.[26] SF2a-TT15 (Institut Pasteur), a synthetic glycoconjugate vaccine targeting *S. flexneri* 2a, has demonstrated strong antibody responses in Phase 1 trials.[27]

Given the pressing need for an effective and broadly protective *Shigella* vaccine, we have employed a computational approach to design a multitope mRNA vaccine targeting the most prevalent *Shigella* pathotypes. Using their ability to attach to major histocompatibility complex (MHC) molecules on antigen-presenting cells and T-cell receptors on T-cells, small peptides (epitopes) may activate unique immune responses.[28] Many novel methods such as peptide-based and DNA vaccines showed promising opportunities for developing scalable and rapid vaccines. Advances in immunoinformatics have facilitated the development of peptide-based vaccines, which offer advantages such as minimal toxicity, ease of production, and broad antigenic coverage.[29] Other bacterial and viral diseases, including *Escherichia coli*,[30,31] *Mycobacterium tuberculosis*,[32] *Streptococcus pneumoniae*,[33] hepatitis C virus (HCV),[34] severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2),[35] human immunodeficiency virus (HIV),[36] and Influenza virus subtypes,[37] have all benefited from epitope-based vaccinations. However, despite the promise of peptide-based vaccines, their immunogenicity index was observed to be low.[38] In addition, the use of pDNA within DNA vaccines raised concerns regarding the risk of insertional mutagenesis.[39] In contrast, mRNA vaccines have emerged as a more promising alternative, demonstrating higher efficacy compared to DNA vaccines. This is attributed to the potential of mRNA vaccines to mitigate safety concerns while enhancing efficacy. Notably, mRNA therapeutics are considered noninfectious entities due to their lack of genomic integration and replication, with rare exceptions of recombination between single-stranded RNA molecules.[40] Since *S. sonnei* str. Ss046 is the Uniprot reference strain for *S. sonnei*,[41] we were motivated to adopt a reverse vaccinology technique to find prospective vaccine candidates from its proteome. The exoproteome and secretome were predicted using a suite of bioinformatics algorithms and filters,[42] including those for predicting virulence factors, human homologs, molecular weight, transmembrane helices, antigenicity, conservancy, and population coverage.[43-49] We used immunoinformatics methods to choose 4 proteins that fulfilled all the requirements and to estimate their T-cell and B-cell epitopes.[48,50] By fusing the chosen epitopes with the proper linkers and adjuvants, we were able to create a multitope vaccine. The designed vaccine's physicochemical properties, solubility, secondary, and tertiary structures, conformational B-cell epitopes, disulfide bonds, docking with toll-like receptor 4 (TLR-4), normal mode analysis, and immune simulation were all analyzed.[51-59] So, the overall aim of this study is to provide a safer, more efficient, and broader coverage solution against various *Shigella* strains by in silico approaches, addressing the limitations of current treatment methods.

Materials and Methods

Retrieval of protein sequence

We obtained the amino acid sequences of *S. sonnei* str. Ss046's complete proteome from UniProt, using Taxon ID

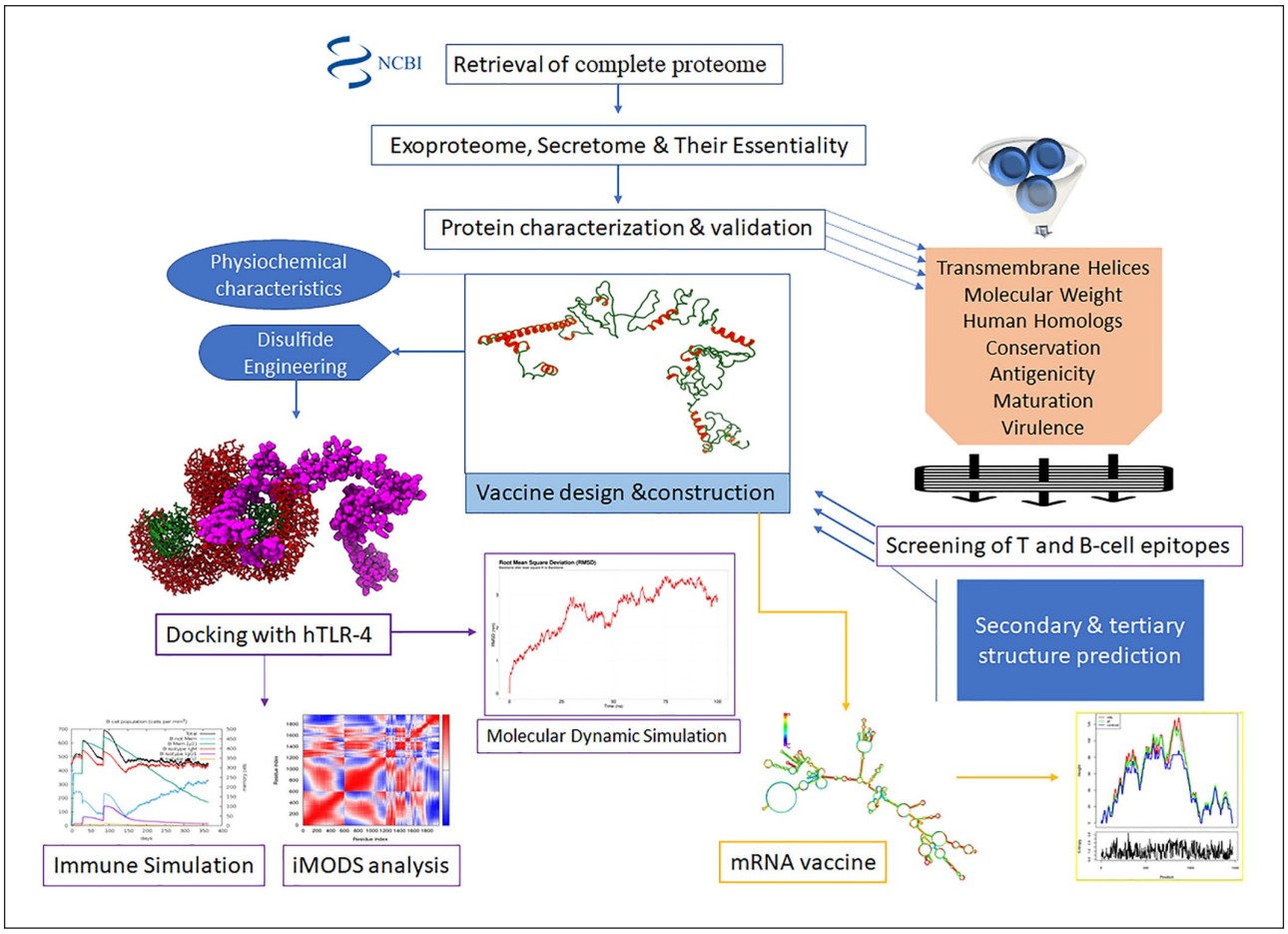

Figure 1. A methodological flow diagram of in silico vaccine design against common *Shigella* pathotypes. Light blue shapes represent ss046's sequence retrieval, analysis, epitope prediction, structural evaluation and yellow boxes showing mRNA vaccine construct, whereas magenta shapes represent simulation tests, respectively.

300269 and entry ID UP000002529, where this strain serves as the reference proteome.[41] Given its significant impact as a predominant pathogen, *S. sonnei* was selected as our primary focus.[11,16] Notably, *S. sonnei* is believed to have evolved more recently compared to other Shigella species and serotypes, and it frequently emerges in low-income countries transitioning from developing to developed status, thus altering the global epidemiological landscape.[60] Its tendency to appear in low-income countries that are emerging from the ranks of the developed has also altered the global epidemiological landscape.[9] Figure 1 depicts the whole study design.

Prediction of exoproteome, secretome, and their essentiality

Using the entire proteome of *S. sonnei* str. Ss046, we employed PSORTb v3.0.2 online server[42] (https://www.psort.org/psortb/) and Geptop 2.0[61] (http://geptop.molgenrug.nl/) to predict the subcellular localization of 4071 proteins. A standard essentiality score threshold of 0.24 was applied. To identify potential vaccination candidates, we focused on proteins predicted to localize in the outer membrane or extracellular space, encompassing both essential and nonessential proteins within these compartments.

Prediction of virulent proteins and transmembrane helices

VirulenPred[43] (http://crdd.osdd.net/raghava/virulenpred/) was used to evaluate the exoproteome and secretome for virulence. We were careful to choose just potentially harmful proteins. In addition, we only chose proteins that included at least one predicted transmembrane helix using TMHMM[46] (http://www.cbs.dtu.dk/services/TMHMM/).

Determination of human homologs and molecular weight estimation

To see whether any of the virulence proteins have human homologs, we ran a BLASTp[44] search (https://blast.ncbi.nlm.nih.gov/Blast.cgi) against the human proteome. Any protein similar to a human protein (35%) was ruled out. Using the Expasy tool (https://web.expasy.org/compute_pi/),[45] we assessed the molecular weight of the remaining proteins and chose only those with molecular weights of less than 110 kDa.

Protein antigenicity, maturation, and conservation analysis with various Shigella *pathotypes*

We assessed the antigenicity of proteins using Vaxijen V 2.049[47] (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html), categorizing those with antigenicity values higher than 0.4. Eleven *Shigella* strains, covering a range of pathologies, were analyzed for their degree of conservation of the proteins specified in the previous rounds. One main goal of this research was to ensure that the developed vaccine will be effective against all pathogenic *Shigella* by doing BLASTp analysis on each protein against a representative strain of each pathotype. Any protein that had a sharing percentage of less than 80% was either removed or suggested for elimination. We used SignalP 5.0[62] Server (http://www.cbs.dtu.dk/services/SignalP/) to estimate the site of the signal peptide, based on the filtered proteins. The full-grown peptides were collected for future study.

Screening for B- and T-cell epitopes

Epitopes for both cytotoxic T lymphocytes (CTLs) and helper T lymphocytes (HTLs) were predicted using the immune epitope database (IEDB) website[63] (http://www.iedb.org/). We predicted epitopes recognized by CTLs that were shared by the HLA allele reference set that covered more than 97% of the population. Antigenicity, allergenicity, toxicity, conservancy, immunogenicity, and population coverage were determined by profiling the top 100 peptides. We employed the IEDB server's conservancy tool (http://tools.iedb.org/conservancy), immunogenicity tool (http://tools.iedb.org/immunogenicity), and population coverage tool (http://tools.iedb.org/population). Antigenicity, toxicity, and allergenicity predictions were made using the VaxiJen (http://www.ddgpharmfac.net/vaxijen/VaxiJen/VaxiJen.html), ToxinPred2, and AllerTOP v.2.0 (https://www.ddg-pharmfac.net/AllerTOP/index.html) software packages. Based on the lowest percentile rank, maximum antigenicity, immunogenicity, and binding affinity, we picked the top 10 peptides that fulfilled all criteria. HTL epitope predictions were made using an HLA allele reference collection with above 99% population coverage. Antigenicity, allergenicity, toxicity, conservancy, induction of interferon-gamma (IFN-γ), interleukin-4 (IL-4), and interleukin-10 (IL-10) as well as population coverage criteria were used to identify the top 100 peptides for profiling. Online resources such as IFNepitope (https://webs.iiitd.edu.in/raghava/ifnepitope/predict.php), IL4pred (https://webs.iiitd.edu.in/raghava/il4pred/index.php), and IL-10Pred (https://webs.iiitd.edu.in/raghava/il10pred/predict) were used for cytokine prediction. The peptides with the greatest antigenicity binding affinity and the lowest percentile rank were chosen as the top 10. We also used docking research using Autodock Vina[64] (https://vina.scripps.edu/) to evaluate the binding affinity of the potential CTLs and HTLs with their corresponding HLA alleles. To do this, the PEP FOLD 3 webserver[65] was used to predict the 3D structure of each peptide. Concurrently, the 3-dimensional structures of HLA-A02:01 (PDB ID 4u6y) and HLA-DRB104:01 (PDB ID 5lax) were obtained from the protein data bank to serve as receptors for MHC-I and MHC-II epitopes, respectively. Linear B-cell epitopes were predicted using the IEDB server's (http://www.iedb.org/) Bepipred Linear Epitope Prediction 2.0, Emini Surface Accessibility Prediction, and Kolaskar and Tongaonkar Antigenicity tools. Antigenicity ratings greater than 0.4 were used to identify peptides with a length between 8 and 20 amino acids. We tested the selected peptides for toxicity and allergenicity using the same servers as previously for single epitope estimation.

Homology modeling and epitope mapping

The 3D structures of the 4 target proteins were modeled using SWISS-MODEL. The highest-quality templates were selected based on sequence identity (>80%), and GMQE (>.50). Structural validation was performed using PROCHECK (Ramachandran plot analysis). After validation, epitope regions were mapped onto the surface of the modeled proteins. PyMOL and Discovery Studio were used to visualize the predicted epitopes, allowing assessment of their spatial distribution, surface accessibility, and potential immunogenic properties. Selected epitopes (CTL, HTL, and B-cell) were color-coded to visualize their locations on each protein.

Design and construction of multi-epitope vaccine candidate

We used GGGS, GPGPG, and KK amino acid linkers to join the top 24 CTL, HTL, and B-cell epitope candidates from the epitope prediction stage to create a possible multitope vaccination. To complete the constructed vaccine, we included the PADRE sequence and defensin adjuvant in addition to the epitopes. Using the same servers as before for single epitope estimate, we analyzed the vaccine candidate's antigenicity score,[47] allergenicity,[66] and toxicity probability.[67] We analyzed the physicochemical characteristics of the proposed vaccination using the ProtParam program, which can be found on the Expasy server[45] (https://web.expasy.org/protparam/).

Secondary and tertiary structure of the SS-Ss046

We used the SOLpro server[51] (http://scratch.proteomics.ics.uci.edu/) and the PSIPRED 4.0 webserver[68] (http://bioinf.cs.ucl.ac.uk/psipred/) to make predictions about the likelihood of overexpression in *E coli* and the protein secondary structure of the multitype construct, respectively. To evaluate the prospective vaccine's interaction with a TLR, we predicted its 3D structure. To do this, we accessed the 3Dpro[69] (http://www.sbg.bio.ic.ac.uk/3dpro/) and Phyle2 webserver[70] (http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index). After predicting the tertiary structure, we enhanced its stability and

energy using the GalaxyRefine server[71] (http://galaxy.seoklab.org/cgi-bin/submit.cgi?type=REFINE), followed by validation through Ramachandran plot analysis[71] and ProSA[72] (https://prosa.services.came.sbg.ac.at/prosa.php).

Prediction of conformational B-cell epitopes

After performing 3D structure prediction and refinement, we made predictions about the conformational B-cell epitopes that would be used in the multitope design. For this prediction, we used ElliPro Server[54] (http://tools.iedb.org/ellipro/), a trustworthy program for identifying B-cell epitopes against a given antigen. A minimum score threshold of 0.5 and maximum distance of 6 angstrom was applied to ensure epitope reliability. ElliPro approximates protein shape using ellipsoids, with higher PI values indicating greater solvent accessibility.

Disulfide engineering of the designed vaccine

We included disulfide links to strengthen the 3D structure of the planned construct before beginning docking research for the created potential vaccine. Disulfide linkages are thought to increase protein stability by altering the protein's geometric shape. For this, our team turned to Disulfide by Design 2.0[55] (http://cptweb.cpt.wayne.edu/dbd2/).

Docking of designed vaccine with hTLR-4

For our epitope-based synthetic vaccine design, we selected hTLR-4 (PDB ID: 4G8A) as the target receptor. Docking analysis between the ligand and receptor was performed using the ClusPro 2.0 server[57] (https://cluspro.org/). By running billions of conformations, aggregating the 1000 lowest energy structures created, and filtering out steric conflicts, this service can accurately predict the optimum docking models. The ligand and receptor PDB files were uploaded to the ClusPro service, and docking was conducted using the default settings.

Dihedral coordinate-based normal-mode analyses

To learn more about how the built epitope vaccine moves with respect to the bound hTLR-4 protein target, we used the iMODS server[58] (http://www.imods.chaconlab.org/). This server's speed and effectiveness are definite pluses. A high eigenvalue, for example, indicates a more severe deformation as predicted by this service.

Immune simulation of the designed vaccine

Using a computational method, we predicted the stimulated immune response to the proposed vaccination using the C-ImmSim server[73] (http://www.iac.rm.cnr.it/filippo/C-IMMSIM/). For this study, we injected the planned vaccine 3 times, each time waiting 4 weeks between doses, following the prime-booster-boost strategy. This strategy was used to induce a persistent immunological response.

Molecular dynamic simulation

A 100 ns Molecular dynamics (MD) simulation was carried out using the GROningen Machine for Chemical Simulations aka GROMACS.[56] The CHARMM36m force field was used for the simulation. Using the TIP3P water model, a water box was constructed with its borders 1 nm from the protein surface. The appropriate ions were used to neutralize the systems. The simulation was performed using periodic boundary conditions and a temporal integration step of 2 fs after energy minimization, isothermal-isochoric (NVT), and Isobaric (NPT) equilibration of the system. For this trajectory analysis, a 100 ps snapshot interval was used. Root mean square deviation (RMSD), root mean square fluctuation (RMSF), radius of gyration (Rg), and solvent accessible surface area (SASA) were calculated after the simulation was finished using the rms, rmsf, gyrate, and sasa modules included into the GROMACS program. As for the plots, they were all made using the ggplot2 tool in RStudio. All MD simulations were performed on high-performance simulation equipment at the Bioinformatics Division, National Institute of Biotechnology, using the Ubuntu 20.04.4 LTS operating system.

Codon-optimization and analysis of the vaccine mRNA

The Java Codon Adaptation Tool (JCat) server (http://www.prodoric.de/JCat) was used to measure the multi-epitope vaccine expression level in *E coli* (strain K12). To ensure maximal expression, Jcat computed the query sequence's GC content and Codon Adaptation Index (CAI).

The secondary structure of mRNA was then predicted using the RNAfold web server (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi). This service forecasts the minimal free energy (MFE) and partition function created thermodynamically by the query mRNA structures. To analyze mRNA folding and vaccine secondary structure, the optimized DNA sequence was transformed to a possible RNA sequence using DNA<->RNA->Protein (http://biomodel.uah.es/en/lab/cybertory/analysis/trans.htm).

Result

Proteome analysis for selection of vaccine candidates

Among *S. sonnei* str. Ss046's total of 4071 proteins, 140 proteins were predicted to be located in the outer membrane or extracellular space from which 81 were projected to be pathogenic. We used further criteria to pick the top 4 proteins that were not similar to proteins in humans, had high antigenicity scores and conserved among 11 *Shigella* strains representing different pathotypes (Supplemental S1 Figure). Q3YZL0 was a putative outer membrane protein, Q3YZM5 was a PapC-like porin protein, Q3Z3I2 was a

Table 1. Characteristics of top-ranked linear MHC-1 epitopes on protein surface, predicted through ElliPro in IEDB-analysis resource from 4 selected vaccine candidates.

Protein	Epitope	PR	AS	Allergen	Toxicity	Conserv.	Im
Q3YZL0	SVMAGPSVRV	0.06	1.3031	NO	NON	100.00%	0.27134
	SEHQSTLSA	0.41	1.0267	NO	NON	100.00%	0.25668
	SVMAGPSVR	0.09	0.9044	NO	NON	100.00%	0.13014
	FSVMAGPSVR	0.52	0.8800	NO	NON	100.00%	0.09782
	STFSGDYIRV	0.09	0.5648	NO	NON	100.00%	0.05524
	RQLTRYSDTR	0.3	0.5201	NO	NON	100.00%	0.05479
	AMAGVAYSRV	0.34	0.4473	NO	NON	100.00%	0.05051
	VMAGPSVRV	0.09	1.5536	NO	NON	100.00%	−0.0654
	HQSTLSAGY	0.04	0.6384	NO	NON	100.00%	−0.11057
Q3YZM5	SGTDSSQVGY	0.01	1.8295	NO	NON	100.00%	0.13113
	NSFRVSYSK	0.09	0.9562	NO	NON	100.00%	0.10118
	YTATYNQNFR	0.2	1.0213	NO	NON	100.00%	−0.00424
	MYTATYNQNF	0.06	0.6107	NO	NON	100.00%	−0.02146
	LPQAYLEYTY	0.08	0.5769	NO	NON	100.00%	−0.04352
	HYFNIGSIR	0.04	0.4210	NO	NON	100.00%	−0.12523
	TATYNQNFR	0.25	1.0774	NO	NON	100.00%	−0.34168
Q3Z3I2	AVTNFHINYY	0.11	1.2989	NO	NON	100.00%	0.26925
	AVTNFHINY	0.03	1.2660	NO	NON	100.00%	0.25908
	SAEQTVTFK	0.3	1.1121	NO	NON	100.00%	0.25908
	AEQTVTFKV	0.14	1.1015	NO	NON	100.00%	0.11961
	ATNVALQMY	0.06	0.9955	NO	NON	100.00%	0.11961
	TVTFKVDYI	0.33	0.9437	NO	NON	100.00%	0.11961
	TADGVQPTAF	0.35	0.7108	NO	NON	100.00%	0.09376
	VTKNATFTF	0.1	0.4611	NO	NON	100.00%	0.09376
Q3Z5V5	FYLPYYWNI	0.03	2.8380	NO	NON	100.00%	0.37078
	SSRRWLFYW	0.09	1.4386	NO	NON	100.00%	0.37078
	QTLEPRAQY	0.02	1.2333	NO	NON	100.00%	0.26559
	IQATLPKYY	0.02	0.6978	NO	NON	100.00%	0.16132
	HPNDDSSRRW	0.02	0.6744	NO	NON	100.00%	0.10041
	FSEQNTSSY	0.01	0.4384	NO	NON	100.00%	0.01067
	SSRRWLFYW	0.02	1.4386	NO	NON	100.00%	−0.15364
	TTTNYFEFY	0.06	0.8305	NO	NON	100.00%	−0.15364
	IQATLPKYY	0.02	0.6978	NO	NON	100.00%	−0.20109
	RIYGQAVHF	0.01	0.6652	NO	NON	100.00%	−0.2901

Here, PR = Percentile Rank, AS = Antigenicity Score, Conserv. = conservancy, Im = Immunogenicity.

putative fimbrial-like protein, and Q3Z5V5 was the LPS-assembly protein LptD. Supplemental Tables S1 and S2 detail the protein properties and conservancy, respectively.

Prediction of T-cell and B-cell epitopes from the selected vaccine candidates

We used IEBD to generate peptides from the selected vaccine candidates followed by a comprehensive evaluation of their immunological properties. The top 100 generated MHC-1 peptides from IEBD per each protein candidate were estimated for their antigenicity score, allergenicity, toxicity probabilities, conservancy, immunogenicity, binding affinity, and population coverage. Out of the 400 candidates screened, peptides that passed the selection process were shown to be antigenic, with top scores ranging from 1.3031 to 0.4210. Notably, all peptides were predicted to be probable nonallergens and nontoxic, indicating their safety profile for potential vaccine development. Furthermore, these peptides exhibited the lowest percentile rank, ranging from 0.01 to 0.41, suggesting their uniqueness and potential efficacy in eliciting an immune response. Among the selected peptides, some demonstrated exceptionally high immunogenicity scores. The highest immunogenicity score recorded was 0.37078, with binding affinities ranging from −9.7 to −8.6 kcal/mol. In addition, all peptides exhibited complete conservancy, with a 100% conservancy score. This implies that the selected epitopes are highly conserved across different variants or strains of the corresponding proteins. These top-ranked epitopes (MHC-1 peptides) of each protein are listed in Table 1.

The binding affinity of the selected CTL candidates was evaluated through a docking study, with the resulting docked complexes illustrated in Supplemental Figure S2 and the corresponding binding scores presented in Supplemental Table S3. Subsequently, the top 100 generated MHC-2 peptides from IEBD per each protein candidate were thoroughly evaluated for their antigenicity score,

allergenicity, toxicity probabilities, binding affinity, and population coverage as well as the ability to induce IFN-γ, IL-4, and IL-10. From a pool of 400 promising candidates, we identified peptides that surpassed our stringent selection process. These peptides demonstrated high antigenicity scores ranging from 1.3408 to 0.4656 and were predicted to be non-allergenic and non-toxic, reinforcing their suitability for vaccine development. We found that the top-ranked peptides had the lowest percentile ranks, ranging from 0.15 to 5.1. In addition, they displayed an exceptional binding affinity ranging from −9.1 to −7 (kcal/mol), signifying strong interactions with the target MHC-2 molecules. Moreover, we assessed the ability of the epitopes to induce specific cytokine responses. Cytokines, such as IL-4, IFN-γ, and IL-10, play crucial roles in modulating the immune response. Epitopes capable of eliciting positive cytokine responses were considered favorable for their potential to induce robust and targeted immune reactions. For a detailed overview, we have listed the top-ranked epitopes (MHC-2 peptides) of each protein in Table 2.

The binding affinity of the selected HTL candidates was also analyzed, with the docked complexes displayed in Supplemental Figure S3 and their binding scores listed in Supplemental Table S3.

Docking validation for epitopes and MHC molecules

Two control peptides are known to bind with high affinity to the HLA-A*02:01 and DRB-1*04:01 alleles, respectively, and were used to verify the docking simulations. The peptide TSKGLFRAAVPSGAS (alpha-enolase peptide 26-40)[74] and NLVPMVATV (derived from cytomegalovirus protein pp65)[75] served as a positive control in our docking investigation for MHC-I and MHC-II binding, respectively. The docking scores for the MHC-I control peptide were recorded at −8.9 kcal/mol, while the MHC-II control peptide yielded a score of −7.5 kcal/mol.

In comparison, the selected vaccine candidate epitopes exhibited binding energy scores ranging from −8.6 to −9.7 kcal/mol for MHC-I interactions and from −7.0 to −9.1 kcal/mol for MHC-II interactions (Supplemental Table S3), indicating stronger or comparable binding affinities relative to the control peptides. By doing numerous independent dockings with various random seeds, we confirmed that the predicted binding affinities are not skewed by the random source used in the prediction process.

Population coverage of the predicted epitopes

We calculated the population coverage of the predicted epitopes using the IEDB tool. The combined CTL epitopes covered 98.55% of the world population, while the combined HTL epitopes were predicted to cover 99.88% of the global population. The multitope vaccine composed of all the selected epitopes covered 100% of the world population (Figure 2). Further regional-specific population coverage details are elaborated in the Supplemental Table S4, while the Global HLA Allele Reactivity Profiles for Selected Peptides including Coverage and Compatibility Analysis are detailed in the Supplemental Table S5.

Screening for B-cell epitopes

Bepipred Linear Epitope Prediction 2.0, Emini Surface Accessibility Prediction, and Kolaskar and Tongaonkar Antigenicity prediction tools were used for B-cell epitope identification. We identified 13 B-cell epitopes for Q3YZL0_outer membrane protein, 88 for Q3YZM5_PapC-like porin protein, 14 for Q3Z3I2_fimbrial-like protein, and 74 for Q3Z5V5_LPTD. Among the generated peptides for each protein, epitopes with a length between 8:20 amino acids were analyzed for their antigenicity, and peptides with antigenicity score >0.4 were tested for their allergenicity and toxicity. A list of predicted top B-cell epitopes of each protein and their characteristics are given in Table 3.

Homology modeling and epitope mapping

High-confidence 3D models were successfully generated for all target proteins, with validation metrics confirming their structural reliability. The Ramachandran plot analysis revealed that 100% of residues in the putative outer membrane protein (Q3YZL0) model, 99.5% in the PapC-like porin protein (Q3YZM5) model, 98.2% in the putative fimbrial-like protein (Q3Z3I2) model, and 99.7% in the LPS-assembly protein LptD (Q3Z5V5) model were in the favored regions. GMQE scores for the models ranged from 0.58 to 0.88, further supporting their structural accuracy and confidence for downstream analysis.

Molecular visualization of the mapped epitopes revealed their accessibility and immunological relevance. In the putative outer membrane protein (Q3YZL0), epitopes were found to be surface-exposed and readily accessible for immune recognition (Figure 3A). The PapC-like porin protein (Q3YZM5) exhibited epitope regions strategically positioned near porin channel openings, enhancing their potential for immune interaction (Figure 3B). The putative fimbrial-like protein (Q3Z3I2) displayed epitopes distributed across functionally significant domains (Figure 3C), while the LPS-assembly protein LptD (Q3Z5V5) demonstrated epitopes localized to extracellular and solvent-accessible regions (Figure 3D).

Multitope vaccine construction

We chose 8 epitopes per table of 1, 2, and 3 (2 epitopes from each protein candidate) and the sum of which is 24 candidates of CTL, HTL, and B-cell epitopes joined using GGGS, GPGPG, and KK as linkers. β-defensin and PADRE peptide were also incorporated using EAAK linker to finalize a potential vaccine sequence of 490 amino acids in length (Figure 4) and its sequence was as follows:

Table 2. Characteristics of top-ranked linear MHC -2 epitopes on protein surface, predicted through ElliPro in IEDB-analysis resource from 4 selected vaccine candidates.

Protein	Peptide	PR	AS	Tox.	Allergen	IL-4	INF-Y	IL-10
Q3YZL0	RWFSVMAGPSVRVNE	4	1.1085	NON	NON	−	+	+
	NRWFSVMAGPSVRVN	0.15	0.9785	NON	NON	−	+	+
	VRNRWFSVMAGPSVR	1.2	0.7170	NON	NON	+	+	−
	RVSTFSGDYIRVTDN	4.4	0.5795	NON	NON	+	+	+
	DSVRNRWFSVMAGPS	3.6	0.5148	NON	NON	+	+	−
	SRVSTFSGDYIRVTD	4.6	0.5055	NON	NON	+	+	−
	LVTSFSYAGDKNRQL	3.9	0.4735	NON	NON	+	−	−
	NRWFSVMAGPSVRVN	2.6	0.9785	NON	NON	+	−	−
	VRNRWFSVMAGPSVR	0.38	0.7170	NON	NON	+	−	−
	SFSYAGDKNRQLTRY	2.3	0.5926	NON	NON	+	−	−
Q3YZM5	SGQKAMAVLRLQDGS	0.41	0.7443	NON	NON	−	+	+
	YAWRALPSLKAKLAL	0.41	0.9364	Yes	NON	−	+	−
	YAWRALPSLKAKLAL	0.76	0.9364	Yes	NON	+	+	−
	SRYYAWRALPSLKAK	0.41	0.6220	Yes	NON	+	+	−
	DGYYSHDGSLAQVDL	0.71	0.6084	NON	NON	−	−	−
	GYYSHDGSLAQVDLS	0.92	0.4656	NON	NON	−	−	−
	TSVDGYYSHDGSLAQ	1.40	0.6539	NON	NON	−	−	−
	EEQTNYNIMLSHYFN	1.60	0.8998	NON	NON	+	−	−
	REEQTNYNIMLSHYF	1.60	1.1238	NON	NON	+	−	−
Q3Z3I2	PTAFANQATTDAATN	4.6	0.5931	NON	NON	−	+	−
	SGNVNAVTNFHINYY	3.6	1.2477	NON	NON	+	+	−
	SGVKTDVPIALEGCD	3	1.0587	Yes	NON	+	+	−
	NQATTDAATNVALQM	5	1.0395	NON	NON	+	+	−
	QATTDAATNVALQMY	2.7	0.9997	NON	NON	+	+	−
	TSGNVNAVTNFHINY	4	1.3048	NON	NON	−	−	−
	TTDAATNVALQMYLP	2.8	1.1662	NON	NON	+	−	−
	PNVSATKLQTNGAVS	3.9	1.1218	NON	NON	+	−	−
	TDAATNVALQMYLPD	5.1	0.9094	NON	NON	+	−	−
	LQMYLPDGSTSVTPG	2.5	0.8915	NON	NON	+	−	+
Q3Z5V5	AKYTTTNYFEFYLPY	0.18	1.1398	NON	NON	+	−	−
	NAKYTTTNYFEFYLP	0.25	1.2574	NON	NON	+	−	−
	PIFYSPYLQLPVGDK	0.25	0.6786	NON	NON	+	−	+
	KYTTTNYFEFYLPYY	0.29	1.3408	NON	NON	+	−	−
	TTNYFEFYLPYYWNI		1.8695	NON	NON	+	−	+
	EFRYLSQAGAGLMEL	0.56	0.4546	NON	NON	+	+	−
	KYTTTNYFEFYLPYY		1.3408	NON	NON	+	+	−
	DPSYFNDFDNKYGSS	0.74	0.7153	NON	NON	+	+	−
	PSYFNDFDNKYGSST	0.74	0.8819	NON	NON	+	−	−

Here, NON = Absent, PR = Percentile Rank, AS = Antigenicity Score, Tox = Toxicity Negative (−) = Absent, Positive (+) = Present.

"EAAAKGIINTLQKYYCRVRGGRCAVLS CLPKEEQIGKCSTRGRKCCRRKKEAAA KAKFVAAWTLKAAAEAAAKAKFVAAWT LKAAAGGGSSGTDSSQVGYGGGSNSFRV SYSKGGGSVTKNATFTFGGGSTVTFKVDYIG GGSSVMAGPSVRVGGGSSEHQSTL SAGGGSFYLPYYWNIGGGSSSRRWLFYW GPGPGSGQKAMAVLRLQDGSGPGPGREE QTNYNIMLSHYFGPGPGSGVKTDVP IALEGCDGPGPGQATTDAATNVALQMY GPGPGRWFSVMAGPSVRVNEGPGPGV RNRWFSVMAGPSVRGPGPGKYTTTNY FEFYLPYYGPGPGDPSYFNDFDNKYGSS KKEGNGAAVYTNMKKTGNDKEMYTA TYNQNFKKNVSATKLQTNGAVSGVKKKS GTADGVQPTAFANQATTDAKKSVMAG PSVRVNKKAGDKNRQLTRYSDTRWHKKPSY FNDFDNKYGSSTDGYKKHQKEAP GQGGGS."

This construct was predicted to be nonallergenic, nontoxic, and had an antigenicity score of 1.2568 (VaxiJen).

Physicochemical features, protein solubility assessment, and secondary structure prediction

The designed vaccine construct consists of 490 amino acids, with a molecular weight of 51 523.53 Da. The theoretical isoelectric point (pI) of the vaccine is 9.79, indicating a positively charged nature under physiological

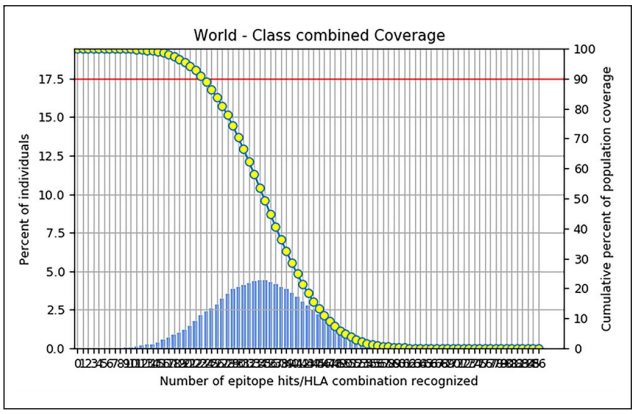

Figure 2. World population coverage of the multitope vaccine composed of all the selected epitopes, showing 100% coverage of the world population. The x-axis indicates the Number of epitope hits/HLA combinations recognized and the y-axis (on the left) shows the Percentage of individuals that can recognize vaccine epitopes. The y-axis on the right shows how many pairs can be recognized by the world population. The blue line displays the total number of pairs, whereas the green line displays the percentage of pairs that are recognized. The ideal situation is depicted by the red line, in which everyone can use at least one vaccination epitope to combat pathogens.

conditions, which is supported by the presence of 64 positively charged residues (Arg + Lys) compared to 31 negatively charged residues (Asp + Glu). Moreover, the protein exhibits moderate stability, as indicated by its estimated half-life of approximately 1 hour in mammalian reticulocytes in vitro and more than 10 hours in Escherichia coli in vivo. The vaccine's structural stability is further supported by a low instability index of 38.77. In addition, the protein demonstrates hydrophilic characteristics, with a negative GRAVY value of −0.629. The physicochemical properties of the predicted vaccine construct are given in the Supplemental Table S6. In addition, the vaccine is predicted to be soluble with a probability score of 0.852386 (SOLpro). The vaccine secondary structure predicted by the PESIPRED server is composed of 14.08% helix, 32.24% strand, and 46.33% coil structure. The structure is given in Supplemental Figure S4.

Tertiary structure prediction, refinement, and validation

Phyle2 and 3Dpro were used to make 3D structure predictions and then Procheck and Prosa were used to double-check those predictions. Validation showed that the model predicted by the 3Dpro server was superior to those projected by phyle2. Figure 5 shows the results of using the GalaxyRefine server, a computer software, to refine the projected 3D structure (3Dpro) of the prospective vaccine and structural validation. We used the Ramachandran plot (PROCHECK) (Figure 5B) and the ProSA web service (Figure 5C) to compare the initial and final structures. Around 92.4% of residues in the optimized structure are in the primary allowed zone, 6.0% in the supplementary permitted region, 1.0% in the generously allowed region, and 0.5% in the unfavorable region. The Z-score went up from −2.54 to −3.41 as well. The main structure included 0.8% of disallowed residues, 1.6% of residues in the generously permitted zone, 1.6% in the extra allowed region, and 84.8% of residues in the most preferred region.

Conformational B-cell epitope prediction

The tertiary structure and folding of the designed vaccine may generate new conformational B-cell epitopes and for this purpose, we used ElliPro server conformational. In the current assessment, the server predicts 11 new epitopes, and their scores were between 0.532 and 0.877 which are given in the Supplemental Table S7. The predicted 3D models of the generated epitopes are shown in Supplemental Figure S5.

Vaccine disulfide engineering

Disulfide by Design 2.0 server suggested that 35 pairs of amino acids are eligible to make disulfide bonds. Mutation was created for pairs having energy values less than 2. Both LEU228-TYR231 and VAL434-ASN437 pairs were engineered to have lower energy scores and chi-3 values of −102.89 and 104.79, respectively.

Table 3. Properties of highly ranked linear B-cell epitopes on protein surface predicted using ElliPro in IEDB-analysis resource for 4 chosen vaccine candidates.

Protein	Peptide	Length	Antigenicity score	Allergenicity	Toxicity
Q3YZL0	AGDKNRQLTRYSDTRWH	17	0.8267	Nonallergen	Toxin
	SVMAGPSVRVN	11	1.5817	Nonallergen	Nontoxin
	QSTLSAGYLHART	13	0.6318	Allergen	Nontoxin
Q3YZM5	EGNGAAVYTNM	11	1.5757	Nonallergen	Nontoxin
	TGNDKEMYTATYNQNF	16	0.9914	Nonallergen	Nontoxin
	PDISGVAHTTAKVTVS	16	0.9926	Nonallergen	Nontoxin
Q3Z3I2	NVSATKLQTNGAVSGVK	17	1.3137	Nonallergen	Nontoxin
	SGTADGVQPTAFANQATTDA	20	1.0285	Nonallergen	Nontoxin
	VSGVKTDVPIALEGCDVT	18	1.1204	Nonallergen	Nontoxin
Q3Z5V5	PSYFNDFDNKYGSSTDGY	18	1.2219	Nonallergen	Nontoxin
	HQKEAPGQ	8	1.2525	Nonallergen	Nontoxin
	KEAPGQPEPV	10	0.9684	Nonallergen	Nontoxin

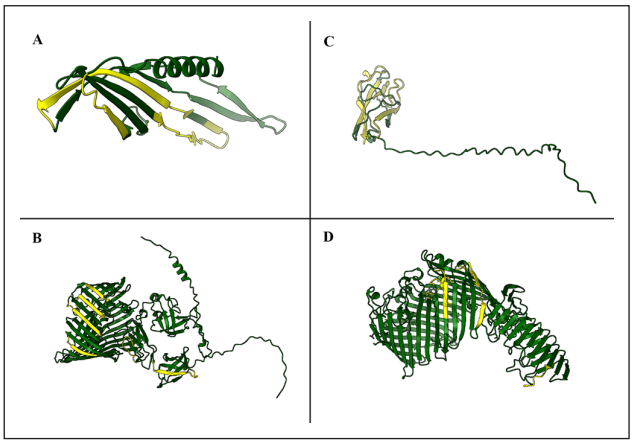

Figure 3. Structural visualization of the modeled target proteins with mapped epitope regions highlighted in yellow: (A) Putative outer membrane protein (Q3YZL0) showing surface-exposed and accessible epitope regions. (B) PapC-like porin protein (Q3YZM5) with epitope regions located near porin channel openings. (C) Putative fimbrial-like protein (Q3Z3I2) displaying epitopes distributed across functional domains. (D) LPS-assembly protein LptD (Q3Z5V5) with epitopes positioned in extracellular and solvent-accessible regions.

Molecular docking of the vaccine with TLR-4

Molecular docking analysis using ClusPro 2.0 generated 30 docking models, with the top-ranked model (model 0.00) exhibiting the lowest binding energy of −1378.5 kcal/mol, indicating a highly stable and strong interaction between the designed multitope vaccine and human TLR-4 (Figure 6). The PDBsum server provided detailed insights into the vaccine-TLR-4 interface, highlighting significant interactions across multiple receptor chains. Among these, chain B demonstrated the most extensive binding with the vaccine, involving 30 and 32 interface residues across an interaction area of 1505 Å² and 1519 Å², respectively. This interaction was stabilized by the formation of 2 salt bridges, 17 hydrogen bonds, and 193 nonbonded contacts, underscoring a robust molecular association (Figure 7). Chain A exhibited moderate interaction, with 7 and 6 interface residues spanning areas of 274 Å² and 292 Å², contributing 1 salt bridge, 7 hydrogen bonds, and 60 nonbonded contacts. Chain D also displayed notable interactions, involving 9 and 15 interface residues over areas of 621 Å² and 554 Å², with 1 salt bridge, 6 hydrogen bonds, and 76 nonbonded contacts. These findings collectively suggest that the designed vaccine establishes strong and stable interactions with TLR-4, potentially enhancing immune activation and downstream signaling.

Dihedral coordinate-based normal-mode analyses

NMA was used on the iMODS simulation server to evaluate the stability and adaptability of the vaccine-TLR-4 complex. We first compared the vaccine-TLR-4 complex to the TLR-4 protein using deformability (Figure 8A) and B-factor (Figure 8B) analysis and found that the vaccine-TLR-4 complex had much less distortion. In comparison to the eigenvalue of the target TLR-4 protein, which was 2.480454e 05, the eigenvalue of the vaccine-TLR-4 complex was 3.0261702e 08 (Figure 8 C), showing that much less energy is required to deform the complex. In addition, the variance analysis was used to represent the stiffness of the complex (Figure 8D), and it was shown to be inversely proportional to the eigenvalue. The iMOS generated a covariance matrix showing the linked residue pairs with either anti-correlated (blue) or correlated (red) movements, or without any correlation (white). Predicted associated residue-pair movements were greater for the protein complex than for TLR-4, while uncorrelated motions were lower for the complex protein (Figure 8E). In addition, an elastic network analysis was carried out, with each spring standing in for a specific atomic pair. Analysis of the residues close to the carboxy terminus showed the dark gray band around the spring and noncontinuous gray bands around the same immobility normal string (Figure 8F).

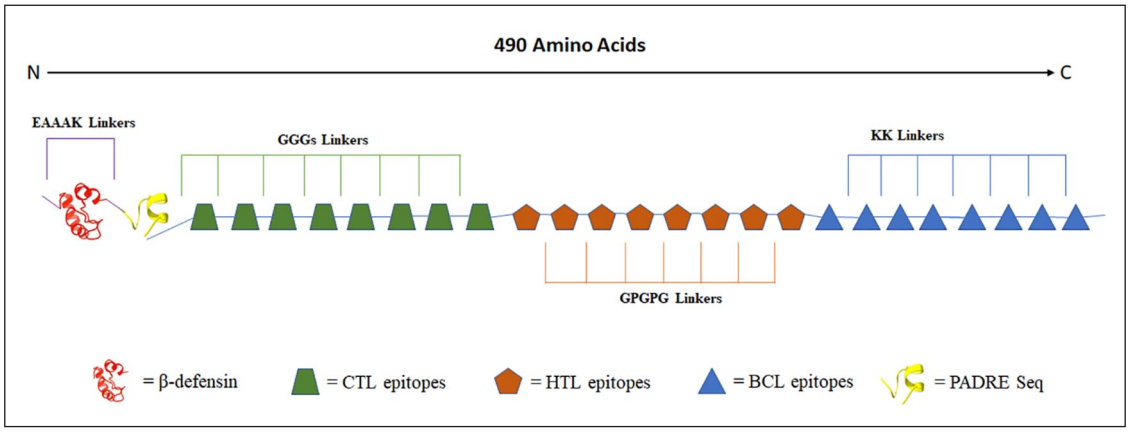

Figure 4. The multitope vaccine is built using different sequences connected by specific linkers. The adjuvant (red) and PADRE (yellow) sequences are joined at the N- and C-terminal respectively. Eight CTL epitopes are connected with GGGS linkers (Green), while 8 HTL epitopes are connected with GPGPG linkers (orange). Finally, 8 BCL epitopes are connected with KK linkers (Blue).

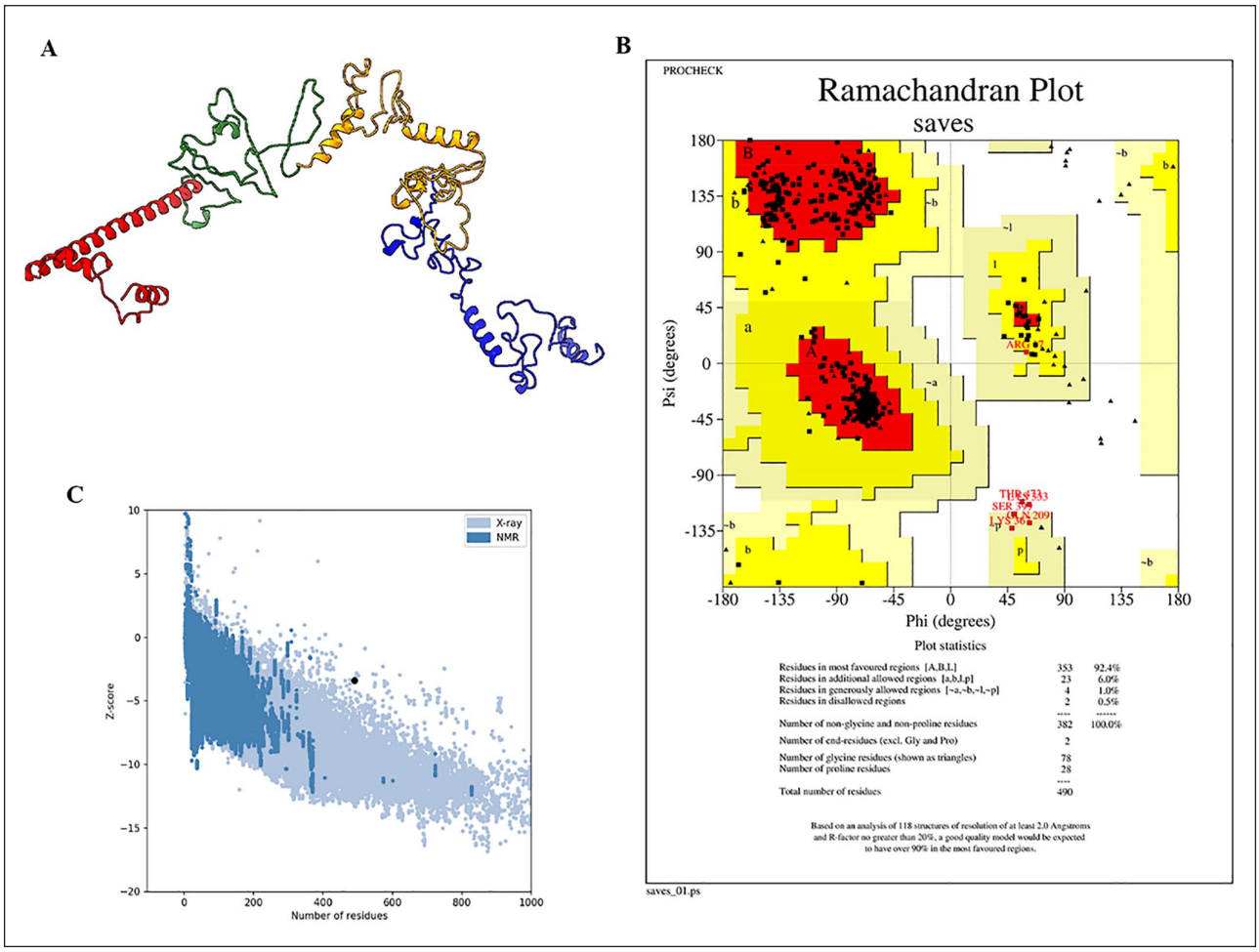

Figure 5. The 3-dimensional structure of vaccine obtained after molecular refinements and validation by Ramachandran plot analysis and ProSA. (A) The constructed vaccine model, with adjuvants highlighted in red, CTL epitopes in green, HTL epitopes in yellow, and B-cell epitopes in blue. (B) The red, yellow, gray, and white color sections represent residues in the most favored, additional permitted, generously allowed, and disallowed regions, respectively. (C) A Z-plot of the revised vaccination model derived from the Pro-SA webserver. The Y-axis depicts the Z-score acquired by NMR or X-ray crystallography for natural proteins, while the X-axis represents the number of residues. The black dot on the Z-plot represents our vaccine's achieved Z-score.

Immune simulation of the designed vaccine

Secondary immunological responses, such as IgM + IgG, IgM, IgG1 + IgG2, IgG1, and IgG2, were shown to be highly stimulated by the potential vaccine, and they increased with further doses of the vaccine (Figure 9A). There was an uptick in the production of cytokines such as interferon-gamma (IFN-γ), transforming growth factor beta (TGF-β), and interleukin-10 (IL-10) after vaccination. A significant elevation in IFN-γ levels was seen after the first vaccination, with concentrations reaching over 400 000 ng/mL. This is consistent with the activation of cellular immune responses. After the second dosage, IFN-γ levels remained stable, demonstrating a sustained induction. Interferon-γ levels were high but slightly rose after the third treatment, reaching over 450 000 ng/mL (Figure 9B). Transforming growth factor-β levels followed a distinct pattern, with a first-dose induction of about 150 000 ng/mL. As can be seen in Figure 7B, Transforming growth factor-β concentrations steadily declined throughout many administrations, from a peak of about 75 000 ng/mL after the second dosage to around 10 000 ng/mL after the third dose. After the first 2 vaccination doses, IL-10 levels peaked at approximately 50 000 ng/mL, suggesting an early immunological response. However, IL-10 concentrations dropped dramatically after the third dosage (Figure 9B), indicating that IL-10 production naturally declines with time. It was also discovered that the vaccination doses prompted a rise in the number of activated B and T cells. Figure 9C shows that the third vaccination dosage produced the largest number of B cells, whereas Figure 9D shows that the second vaccine dose produced the highest number of T cells. This suggests that both T and B cell populations expand in response to vaccination, with separate maxima occurring after different doses of the vaccine.

Molecular dynamic simulation

We performed several investigations using MD simulations based on Cα atoms, to determine the stability of the vaccine-TLR-4 complex. Protein conformational changes during ligand binding were analyzed using the RMSD. After a

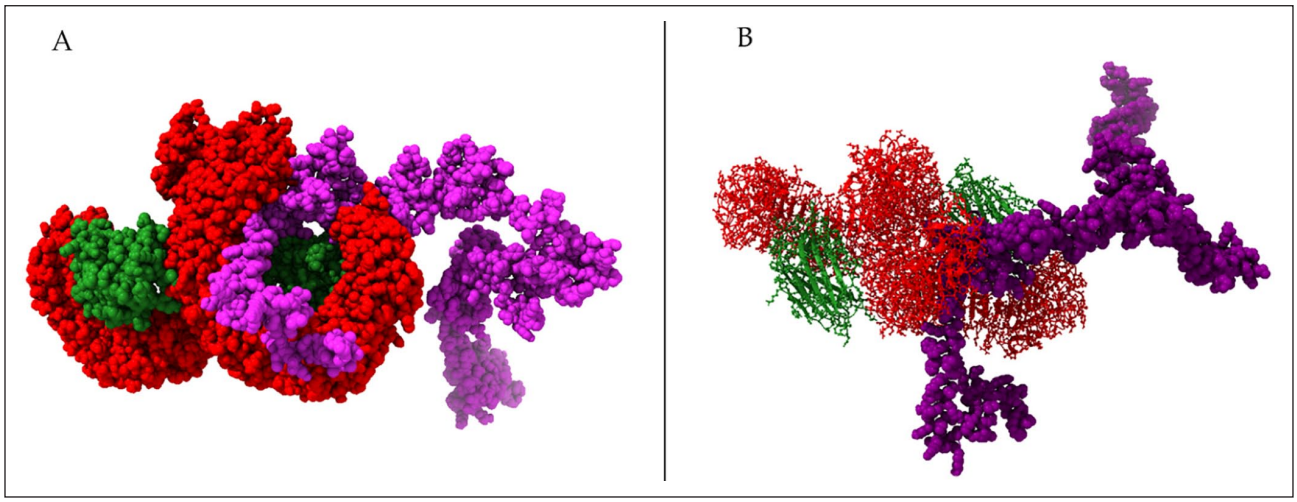

Figure 6. Docked complex of vaccine constructs with human TLR-4 (A, B); vaccine constructs in magenta color and TLR-4 receptor in red and MD2 in green color.

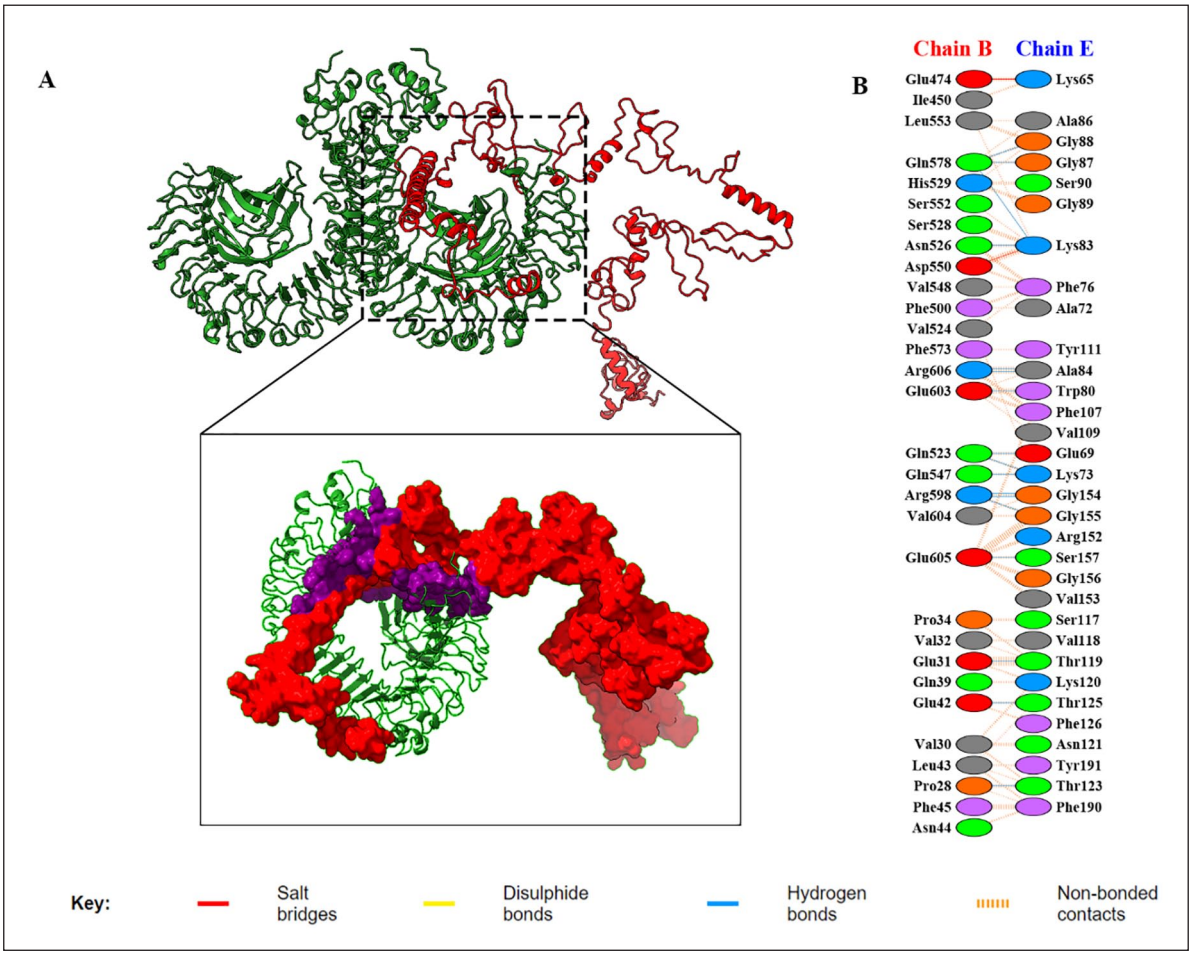

Figure 7. Receptor-vaccine docking analysis. (A) Depicts the docked complex of TLR-4/MD2-vaccine, with the vaccine construct highlighted in red, where the specific interacting residues are highlighted in purple. (B) The interacting residues in the interface of the docked complex, where interacting chains are connected by colored lines, each denoting a different type of interaction as indicated in the key provided.

dramatic spike in the first 10 ns, the complex's RMSD value stabilized at roughly 0.2 nm for the rest of the run. After reaching its peak at 30 ns, it gradually dropped over the next 20 ns. At the end of the molecular dynamic simulations, the complex settled into a rather stable form (Figure 10A). We used the RMSF to assess the protein's local adaptability.

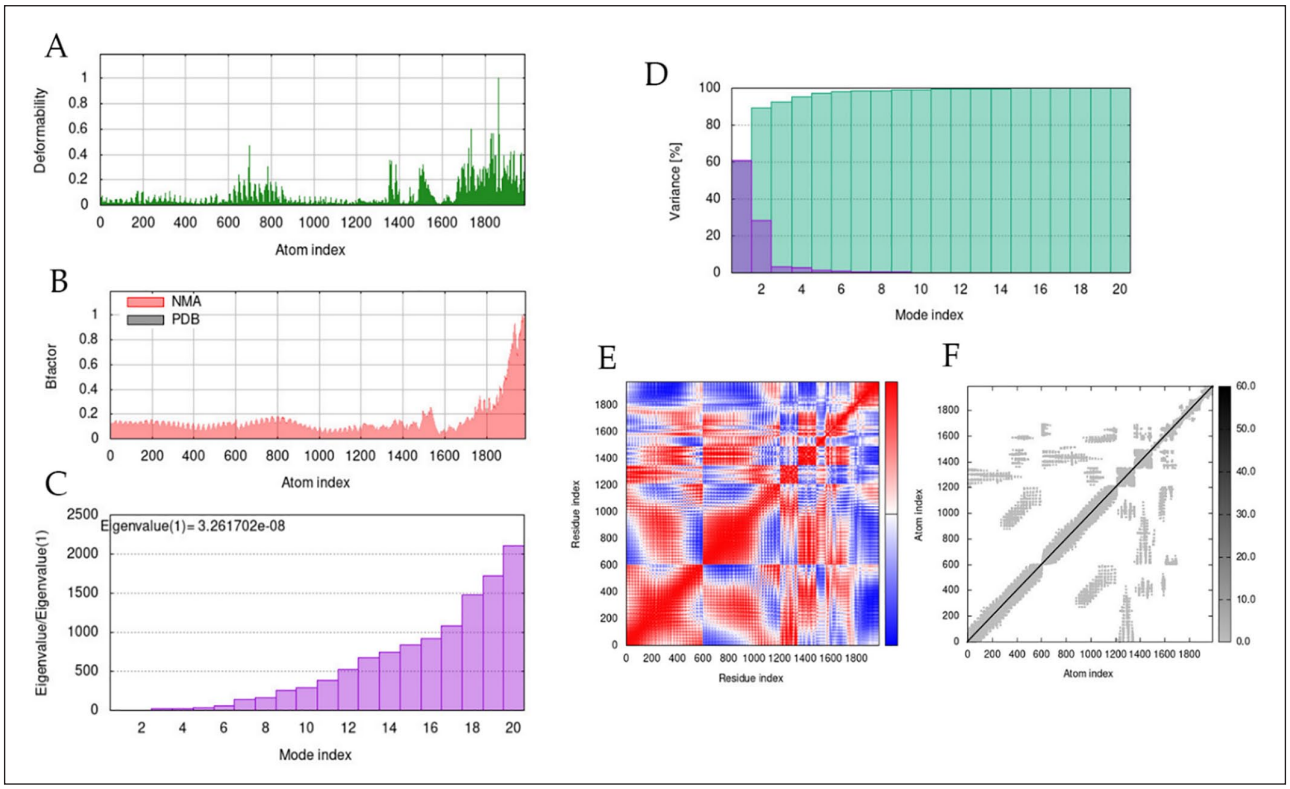

Figure 8. Dihedral coordinate-based normal-mode analyses of the vaccine-hTLR-4 complex; ligand-receptor interaction was assessed throughout comparative (A) Deformabilities, (B) b-factor, (C) eigenvalues, (D) variance, (E) covariance of residue indices, and (F) elastic network analysis.

Increases in RMSF values suggest more positional flexibility in certain amino acids. Particularly flexible were residues at positions 10, 250, and 450 (Figure 10B), indicating a possible role in interactions with the receptor. Using the Rg, we could evaluate how compact the protein was. Stable protein folding is represented by a constant Rg value, whereas changes in this value imply unfolding. Our investigation showed that after 40 ns, the Rg rapidly decreased, then increased, and then decreased again, resulting in a compact condition (Figure 10C). Solvent-accessible surface area research was used to evaluate the protein's hydrophobic core stability. Increases in SASA reflects a higher propensity for protein instability due to solvent accessibility. The SASA values dropped steadily during the simulation and reached rock bottom at the conclusion. Figure 10D shows that this arrangement decreases the likelihood of solvent disruption and implies that the protein is more stable.

Codon optimization and mRNA prediction of the vaccine construct

We employed a sophisticated codon optimization strategy for the *E coli* K12 strain, using the JCAT tool for precision engineering of the vaccine's coding sequence. This optimization yielded a codon sequence with a length of 1470 base pairs (bp). Significantly, the optimized sequence achieved a CAI of 0.98, surpassing the benchmark value of 0.8 and indicating a high level of expression efficiency in the target host system. Furthermore, the GC content of the optimized sequence was measured at 51.22%, aligning closely with the native GC content of the *E coli* K12 strain (50.73%), an essential factor for ensuring transcriptional and translational fidelity.

Advancing the molecular architecture of the vaccine, we leveraged the RNAfold server to predict the mRNA secondary structure of our vaccine construct (Figure 11). This critical step, informed by our codon-optimized sequence, allowed for a rigorous assessment of the mRNA's structural integrity. The server's analysis predicted a highly stable mRNA structure characterized by substantial base pairing and a marked absence of unstable regions, underpinning the vaccine's potential efficacy. Notably, the minimum free energy (MFE) of the optimal secondary structure was calculated to be −475.90 kcal/mol, with the secondary centroid structure exhibiting a free energy of −369.28 kcal/mol. Furthermore, the ensemble free energy was determined to be −503.05 kcal/mol, underscoring the vaccine mRNA's robust stabilization across a spectrum of potential structural conformations.

Discussion

Millions of children under the age of 5 are disproportionately affected by shigellosis, a severe diarrhea illness caused by the gram-negative bacteria *Shigella*.[16] Long-term effects include things like stunted development[76] (in children), mental retardation[77] (in adults), and reactive arthritis[5] (in children). Despite its widespread impact, the

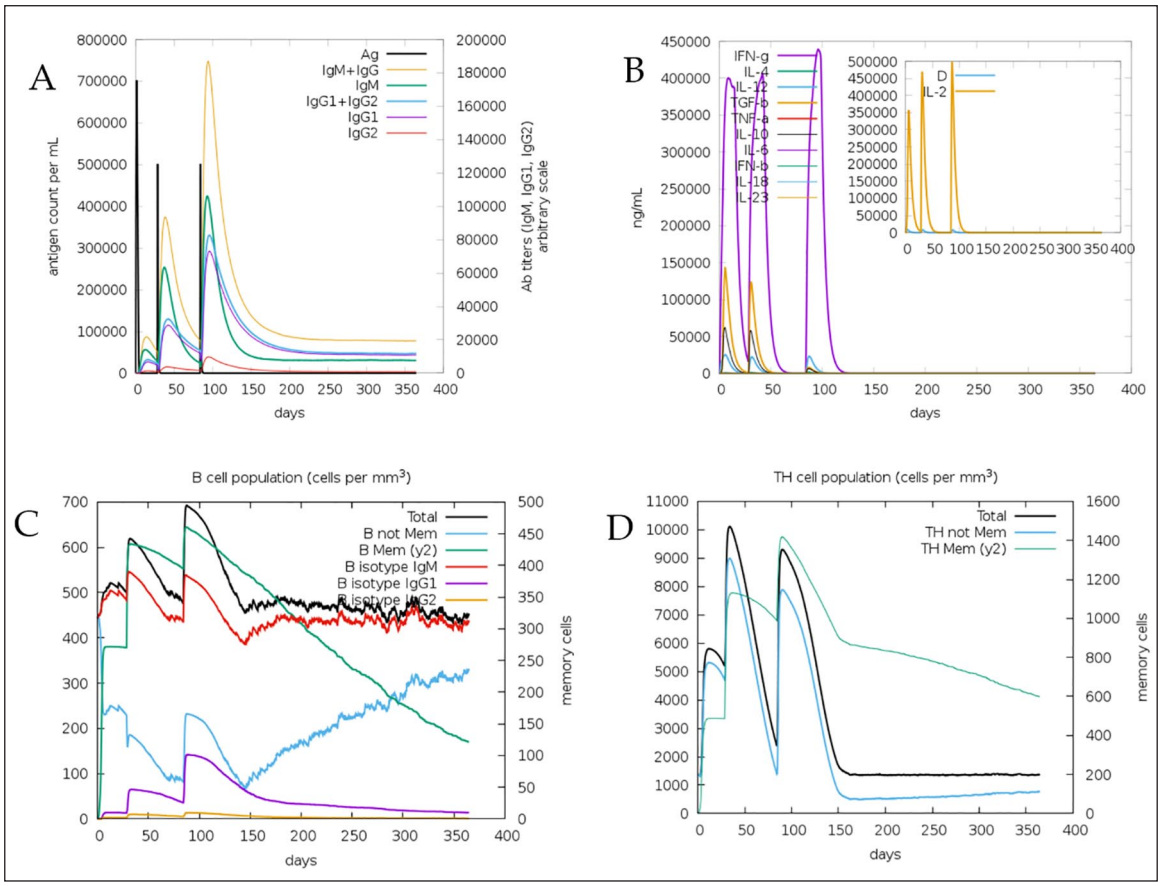

Figure 9. Anticipated immune response following the injection of the designed vaccine. Panel (A) and (B) display the number of antibodies and cytokines, respectively, in response to the vaccine. Panel (C) and (D) depict the population of B and T cells, respectively.

development of effective vaccines for shigellosis has been elusive, and current treatment options are limited.[78] While some promising vaccine candidates are in trials,[27,79] the absence of a licensed vaccine persists. Therefore, the purpose of this research was to develop an efficient epitope-based vaccination against prevalent *Shigella* pathotypes using a reverse vaccinology strategy to overcome these obstacles.

The computational approach of reverse vaccinology was used to find possible vaccine candidates.[80] In this research, vaccine candidates were derived from the proteome of *S. sonnei* str. Ss046, the type strain of *S. sonnei*. Putative outer membrane protein (Q3YZL0), PapC-like porin protein (Q3YZM5), Putative fimbrial-like protein (Q3Z3I2), and LPS-assembly protein LptD (Q3Z5V5) were chosen because they met all of the necessary criteria, such as virulence, lack of homology to human proteins, antigenicity, and conservation among different *Shigella* strains. The putative outer membrane protein of *S. sonnei* has yet to be fully characterized, although it is thought to be involved in important bacterial processes including adhesion and invasion.[81] Porin-mediated food absorption, transporter function, and the development of extracellular encapsulating structures are probably all aided by the PapC-like porin protein.[82] Possible fimbrial-like proteins contribute to adhesion, biofilm formation, and bacterial colonization through fimbriae-mediated adhesion.[83] Lipopolysaccharide production and transport rely on the LPS-assembly protein LptD.[84]

Our in-silico approach integrated multiple computational methods to predict T-cell and B-cell epitopes, ensuring their immunogenicity, stability, and broad population coverage. The successful generation of high-confidence 3D models for all target proteins highlights their structural reliability and suitability for epitope mapping. Validation metrics, including Ramachandran plot analysis and GMQE scores, confirmed the model's accuracy. Mapping of CTL, HTL, and B-cell epitopes onto these structures revealed their surface-exposed and solvent-accessible locations, critical features for effective immune recognition and antigenicity.[85] The distribution of epitopes across extracellular loops and functional domains, as observed in the putative outer membrane protein (Q3YZL0) and LPS-assembly protein LptD (Q3Z5V5), emphasizes their immunological relevance. Similarly, the positioning of epitopes near porin channel openings in PapC-like porin protein (Q3YZM5) and across functional domains in the putative fimbrial-like protein (Q3Z3I2) underscores their potential for eliciting robust immune responses.

The designed vaccine construct demonstrated favorable physicochemical properties, including solubility, stability, non-allergenicity, and non-toxicity. After 3 vaccinations, the body produced abundant quantities of antibodies (IgM + IgG, IgG1 + IgG2) and cytokines (IFN-γ, TGF-β,

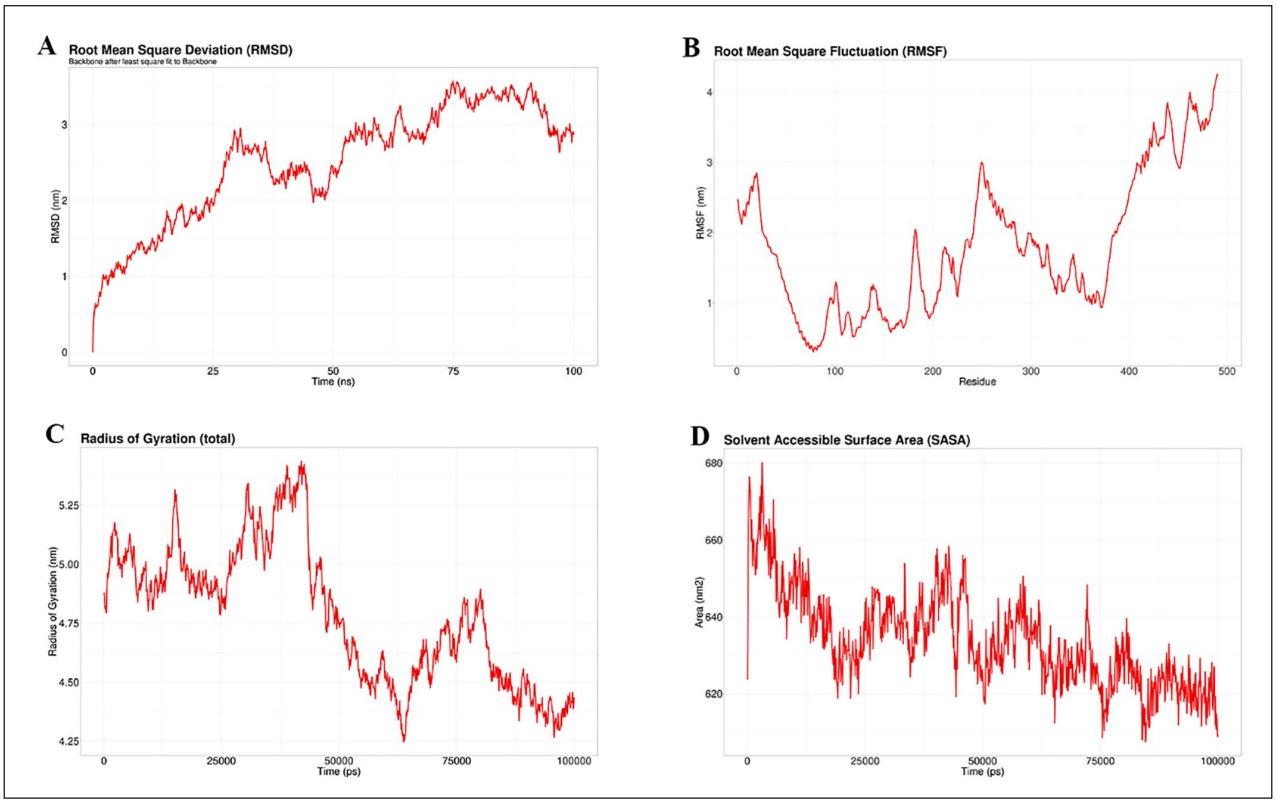

Figure 10. Analysis of conformational changes, flexibility, compactness, and hydrophobic core stability of the vaccine-TLR-4 complex across 100 ns explicit molecular dynamics simulation run. (A) RMSD, (B) RMSF, (C) RG, and (D) SASA. Here, RMSD, root mean square deviation; RMSF, root mean square fluctuation; RG, mean radius of gyration, and SASA meaning solvent-accessible surface area.

IL-10). These predicted responses align with known protective immune mechanisms observed in *Shigella* infections, where humoral immunity, particularly IgG responses, plays a crucial role in pathogen clearance,[86] and IFN-γ-mediated cellular immunity is essential for controlling intracellular bacterial replication.[87] The presence of regulatory cytokines such as TGF-β and IL-10 further suggests a balanced immune response, which is critical for mitigating excessive inflammation while ensuring effective pathogen elimination.[88] Notably, the molecular docking analysis revealed significant binding affinities between epitopes and HLA alleles, further validating the vaccine's potential to elicit both humoral and cellular immunity, which is consistent with previous findings.[89,90] Moreover, TLR-4 an important receptor for innate immunity and adjuvant action, showed high affinity and stability with the vaccination.

Comparing this approach to other in silico-designed peptide vaccines, such as those targeting *Salmonella*, *Escherichia coli*, and *Klebsiella pneumoniae*, our vaccine design methodology demonstrates superior antigen selection by focusing on highly conserved proteins involved in key bacterial processes.[31,91,92] Unlike traditional peptide vaccine designs that rely on empirical screening, our strategy integrates structural modeling, normal mode analysis, and MD simulations to refine the vaccine construct. Structural validation analyses, including the Ramachandran plot and the ProSA web service, were conducted to evaluate the stability and quality of the vaccine construct. The optimized structure showed 92.4% of residues in the primary allowed region, 6.0% in the supplementary permitted region, 1.0% in the generously allowed region, and 0.5% in the forbidden region. The Z-score improved from −2.54 to −3.41, indicating enhanced structural reliability. These findings are consistent with previous studies on similar vaccine constructs, reinforcing the structural integrity and potential effectiveness of the designed vaccine.[93,94] The vaccine's structure remained resilient, demonstrated by its low deformability and high stiffness in normal mode analysis. The vaccine-TLR-4 complex was modeled using MD, and the results shed light on the stability and dynamics of this complex. Changes in the RMSD value suggested structural changes in the protein in response to ligand interaction. Particular amino acids surrounding the 10th, 250th, and 450th residues have been singled out as being potentially important in receptor interactions because of their proximity to extremely flexible areas. The protein's folding and unfolding behavior, as well as the subsequent transition to a compact form, were also revealed by the examination of the Rg. This result indicates that the compound has a stable structural conformation. With decreasing SASA values, the probability of solvent-induced protein instability has been shown to diminish, as measured by the SASA study. Identification of key flexible regions provides potential targets for site-directed mutagenesis studies, in which specific amino acids are modified to improve vaccine efficacy, stability, or binding affinity to the receptor.[95]

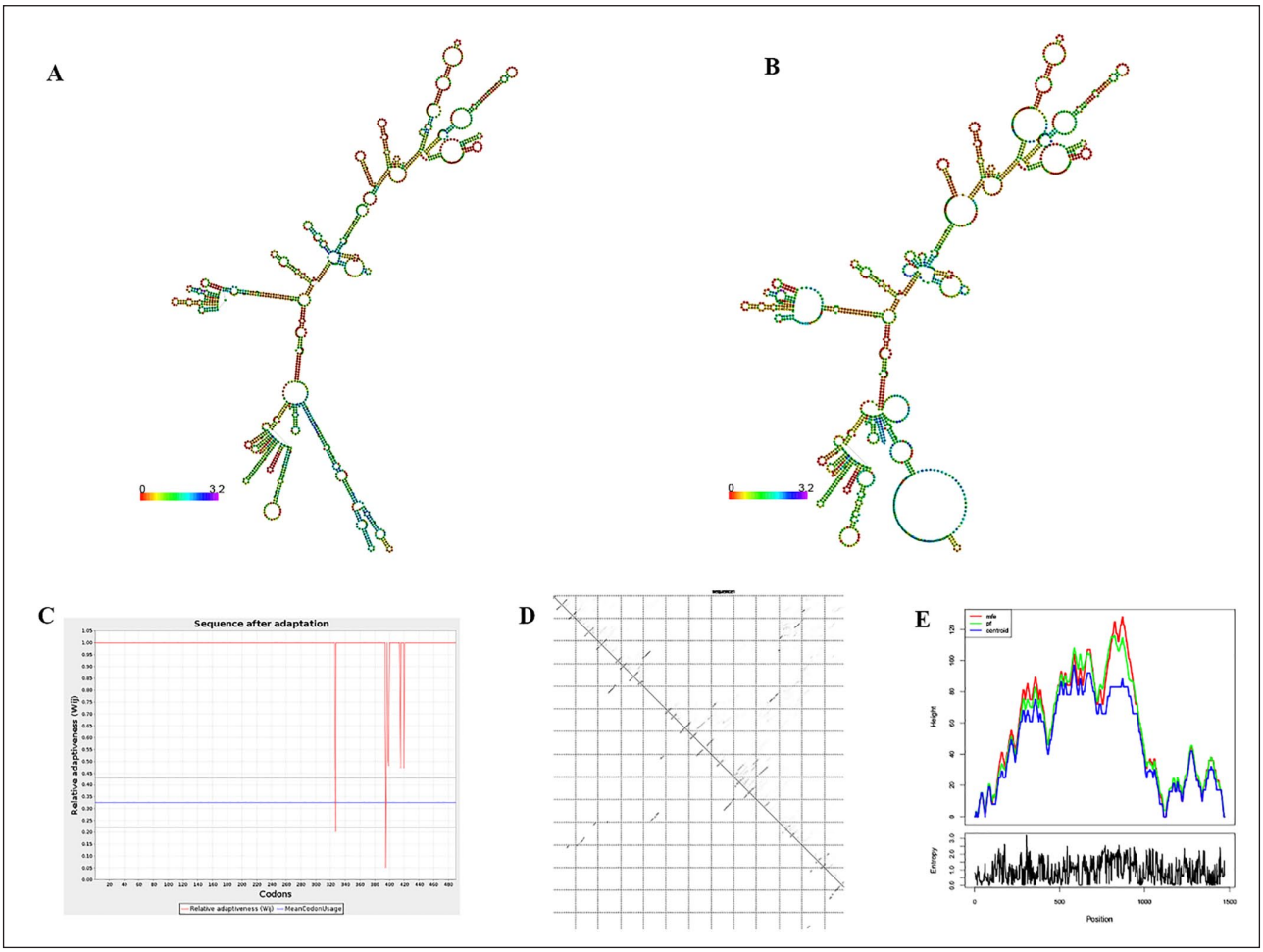

Figure 11. Optimizing codons and predicting mRNA vaccine structure. (A) Optimal secondary structure, (B) Centroid secondary structure of the vaccine mRNA retrieved using RNAfold Webserver. (C) CAI value, (D) a dot plot containing the base pair probabilities, (E) a mountain plot representation of the MFE structure, the thermodynamic ensemble of RNA structures, and the centroid structure along with the positional entropy for each position.

A key innovation of this study is the development of an mRNA-based vaccine, which offers significant advantages over conventional DNA-based MEVs that require plasmid construction and cloning.[96] The mRNA vaccine design leveraged the RNAfold server and related algorithms to optimize secondary structure stability and free energy profiles. The codon optimization process tailored the mRNA sequence for optimal expression in *E. coli* K12, yielding a CAI value of 0.98, indicating efficient translation potential. The predicted secondary structure revealed substantial base pairing and minimal unstable regions, suggesting enhanced stability and translational efficiency.[97]

Overall, the results of this work highlight the promise of reverse vaccinology for developing mRNA *Shigella* vaccines based on epitopes. The proposed vaccine shows potential for providing widespread protection against *Shigella* infection by targeting conserved antigens engaged in essential activities for bacterial survival and pathogenesis. Importantly, it overcomes the safety issues, antigenic diversity, serotype specificity, and inadequate immunogenicity of traditional vaccinations, hence solving the challenge of broad-spectrum protection.[29] Despite the positive findings, it is important to note the study's limitations. Additional research is needed to characterize the proteins' functions, especially the Putative outer membrane protein. The vaccine's efficacy was also evaluated computationally; more experimental validation and clinical studies are needed to verify the vaccine's efficacy and safety in practice. To overcome these obstacles, researchers need to undertake in vitro and in vivo experiments to confirm the vaccine's immunogenicity, protective effectiveness, and safety. The vaccine design process might be improved by investigating new bioinformatics techniques, using omics data, and taking into account a variety of *Shigella* strains.

In summary, our research offers a potential reverse vaccinology mRNA vaccine based on multiple epitopes that can be used to protect against the most frequent *Shigella* pathotypes. The findings provide credence to the vaccine's promise of widespread protection, help researchers find ways to work around its current limits, and forward the research into vaccines for shigellosis and other infectious illnesses. Further study and experimental validation of this vaccine design technique are required before it can be used in clinical practice.

Conclusion

The goal of this research was to create an mRNA vaccine against the most prevalent *Shigella* pathotypes that cause shigellosis. It is found all across the world, but no reliable vaccination has yet been developed to prevent it. We set out on an untried strategy for developing vaccines using computational methods. By analyzing 4 essential proteins found in most of the *Shigella* strains and essential to their survival, we were able to locate immunogenic regions that may trigger an immune response. Together, these pieces form a composite vaccination with added ingredients for increased potency. Vaccine properties and structure were analyzed computationally, as was the vaccine's interaction with TLR-4, a key protein in immune recognition. Remarkably, our results highlighted long-lasting compatibility between the vaccination and hTLR-4. Using state-of-the-art computer models, we were able to predict the vaccine's ability to provoke a strong immunological response, one that includes antibodies, cytokines, and activated B and T cells. All of these positive results point to the vaccine's potential to prevent *Shigella* infections. Based on our findings, this vaccine seems to be a strong contender that should undergo more testing in the real world.

Acknowledgments

The authors acknowledge the Research Cell, Noakhali Science and Technology University for providing the research facilities.

ORCID iDs

Otun Saha https://orcid.org/0000-0001-9159-0437

Newaz Mohammed Bahadur https://orcid.org/0009-0005-7258-4636

Statements and Declarations

Author Contributions

Abdur Razzak: Conceptualization; Methodology; Data curation; Investigation; Formal analysis; Validation; Visualization; Writing—original draft; Software; Writing—review & editing.
Otun Saha: Conceptualization; Methodology; Investigation; Formal analysis; Software; Resources; Supervision; Writing—original draft; Writing—review & editing.
Khandokar Fahmida Sultana: Resources; Writing—review & editing.
Mohammad Ruhul Amin: Supervision.
Abdullah bin Zahid: Data curation; Visualization; Writing—original draft.
Afroza Sultana: Methodology.
Uditi Paul Bristi: Validation; Writing—review & editing.
Sultana Rajia: Validation; Writing—review & editing.
Nikkon Sarker: Formal analysis; Resources; Writing—review & editing.
Md Mizanur Rahaman: Supervision; Writing—review & editing.
Newaz Mohammed Bahadur: Writing—review & editing.
Foysal Hossen: Writing—review & editing.

Funding

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

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Data Availability Statement

Data are contained within the article and Supplementary Material.

Supplemental Material

Supplemental material for this article is available online.

References

1. Kotloff KL, Riddle MS, Platts-Mills JA, Pavlinac P, Zaidi AKM. Shigellosis. *Lancet*. 2018;391:801-812. doi:10.1016/S0140-6736(17)33296-8
2. Prabhurajeshwar C, Chandrakanth Kelmani R. Shigellosis: a conformity review of the microbiology, pathogenesis and epidemiology with consequence for prevention and management issues. *J Pure Appl Microbiol*. 2018;12:405-417. doi:10.22207/JPAM.12.1.48
3. Garcia-Williams A, Amanda Garcia-Williams K, Logan N. *Shigellosis*. National Center for Emerging and Zoonotic Infectious; 2023.
4. Gu B, Cao Y, Pan S, et al. Comparison of the prevalence and changing resistance to nalidixic acid and ciprofloxacin of *Shigella* between Europe–America and Asia–Africa from 1998 to 2009. *Int J Antimicrob Agents*. 2012;40:9-17. doi:10.1016/j.ijantimicag.2012.02.005
5. Khalil IA, Troeger C, Blacker BF, et al. Morbidity and mortality due to Shigella and enterotoxigenic *Escherichia coli* diarrhoea: the Global Burden of Disease Study 1990–2016. *Lancet Infect Dis*. 2018;18:1229-1240. doi:10.1016/S1473-3099(18)30475-4
6. Shad AA, Shad WA. *Shigella sonnei*: virulence and antibiotic resistance. *Arch Microbiol*. 2021;203:45-58. doi:10.1007/s00203-020-02034-3
7. Thompson CN, Duy PT, Baker S. The rising dominance of *Shigella sonnei*: an intercontinental shift in the etiology of bacillary dysentery. *Plos Negl Trop Dis*. 2015;9:e0003708. doi:10.1371/journal.pntd.0003708
8. Ram PK, Crump JA, Gupta SK, Miller MA, Mintz ED. Part II. Analysis of data gaps pertaining to *Shigella* infections in low and medium human development index countries, 1984–2005. *Epidemiol Infect*. 2008;136:577-603. doi:10.1017/S0950268807009351
9. Vinh H, Nhu NTK, Nga TVT, et al. A changing picture of shigellosis in southern Vietnam: shifting species dominance, antimicrobial susceptibility and clinical presentation. *BMC Infect Dis*. 2009;9:204. doi:10.1186/1471-2334-9-204
10. Qu F, Bao C, Chen S, et al. Genotypes and antimicrobial profiles of *Shigella sonnei* isolates from diarrheal patients circulating in Beijing between 2002 and 2007. *Diagn Microbiol Infect Dis*. 2012;74:166-170. doi:10.1016/j.diagmicrobio.2012.06.026
11. Ashkenazi S, Levy I, Kazaronovski V, Samra Z. Growing antimicrobial resistance of *Shigella* isolates. *J Antimicrob Chemother*. 2003;51:427-429. doi:10.1093/jac/dkg080
12. Sousa MÂ, Mendes EN, Collares GB, Péret-Filho LA, Penna FJ, Magalhães PP. *Shigella* in Brazilian children with acute diarrhoea: prevalence, antimicrobial resistance and virulence genes. *Mem Inst Oswaldo Cruz*. 2013;108:30-35. doi:10.1590/S0074-02762013000100005

13. Tajbakhsh M, García Migura L, Rahbar M, et al. Antimicrobial-resistant *Shigella* infections from Iran: an overlooked problem? *J Antimicrob Chemother*. 2012;67:1128-1133. doi:10.1093/jac/dks023
14. Baker S, The HC. Recent insights into *Shigella*. *Curr Opin Infect Dis*. 2018;31:449-454. doi:10.1097/QCO.0000000000000475
15. Ashkenazi S, Cohen D. An update on vaccines against *Shigella*. *Ther Adv Vaccines*. 2013;1:113-123. doi:10.1177/2051013613500428
16. World Health Organization. *Shigella (Shigellosis)*. World Health Organization.
17. DuPont HL, Levine MM, Hornick RB, Formal SB. Inoculum size in shigellosis and implications for expected mode of transmission. *J Infect Dis*. 1989;159:1126-1128. doi:10.1093/infdis/159.6.1126.
18. von Seidlein L, Kim DR, Ali M, et al. A multicentre study of *Shigella* diarrhoea in six Asian countries: disease burden, clinical manifestations, and microbiology. *PLoS Med*. 2006;3:e353. doi:10.1371/journal.pmed.0030353
19. Ross AGP, Olds GR, Cripps AW, Farrar JJ, McManus DP. Enteropathogens and chronic illness in returning travelers. *N Engl J Med*. 2013;368:1817-1825. doi:10.1056/NEJMra1207777
20. Riddle MS, Sanders JW, Putnam SD, Tribble DR. Incidence, etiology, and impact of diarrhea among long-term travelers (US military and similar populations): a systematic review. *Am J Trop Med Hyg*. 2006;74:891-900.
21. Phalipon A, Sansonetti PJ. Shigella' s ways of manipulating the host intestinal innate and adaptive immune system: a tool box for survival? *Immunol Cell Biol*. 2007;85:119-129. doi:10.1038/sj.icb7100025.
22. MacLennan CA, Grow S, Ma L, fu Steele AD. The Shigella vaccines pipeline. *Vaccines*. 2022;10:1376. doi:10.3390/vaccines10091376
23. Raso MM, Arato V, Gasperini G, Micoli F. Toward a *Shigella* vaccine: opportunities and challenges to fight an antimicrobial-resistant pathogen. *Int J Mol Sci*. 2023;24:4649. doi:10.3390/ijms24054649.
24. Lu T, Das S, Howlader DR, Picking WD, Picking WL. *Shigella* vaccines: the continuing unmet challenge. *Int J Mol Sci*. 2024;25:4329. doi:10.3390/ijms25084329.
25. Mo Y, Fang W, Li H, et al. Safety and immunogenicity of a *Shigella bivalent* conjugate vaccine (ZF0901) in 3-month- to 5-year-old children in China. *Vaccines*. 2021;10:33. doi:10.3390/vaccines10010033
26. Roozen GVT, Sukwa N, Chirwa M, et al. Safety, tolerability, and immunogenicity of the InvaplexAR-DetoxShigella vaccine co-administered with the dmLT adjuvant in Dutch and Zambian adults: study protocol for a multi-center, randomized, double-blind, placebo-controlled, dose-escalation phase Ia/b clinical trial. *Vaccines*. 2025;13:48. doi:10.3390/vaccines13010048
27. Meron-Sudai S, Asato V, Adler A, et al. A *Shigella flexneri* 2a synthetic glycan-based vaccine induces a long-lasting immune response in adults. *NPJ Vaccines*. 2023;8:35. doi:10.1038/s41541-023-00624-y.
28. Sette A, Rappuoli R. Reverse vaccinology: developing vaccines in the era of genomics. *Immunity*. 2010;33:530-541. doi:10.1016/j.immuni.2010.09.017
29. Delany I, Rappuoli R, Seib KL. Vaccines, reverse vaccinology, and bacterial pathogenesis. *Cold Spring Harb Perspect Med*. 2013;3:a012476. doi:10.1101/cshperspect.a012476
30. Mehla K, Ramana J. Identification of epitope-based peptide vaccine candidates against enterotoxigenic *Escherichia coli*: a comparative genomics and immunoinformatics approach. *Mol Biosyst*. 2016;12:890-901. doi:10.1039/C5MB00745C
31. Soltan MA, Behairy MY, Abdelkader MS, et al. In silico designing of an epitope-based vaccine against common *E. Coli* pathotypes. *Front Med*. 2022;9:829467. doi:10.3389/fmed.2022.829467
32. Gong W, Pan C, Cheng P, Wang J, Zhao G, Wu X. Peptide-based vaccines for tuberculosis. *Front Immunol*. 2022;13:830497. doi:10.3389/fimmu.2022.830497
33. Bahadori Z, Shafaghi M, Madanchi H, Ranjbar MM, Shabani AA, Mousavi SF. In silico designing of a novel epitope-based candidate vaccine against *Streptococcus pneumoniae* with introduction of a new domain of PepO as adjuvant. *J Transl Med*. 2022;20:389. doi:10.1186/s12967-022-03590-6.
34. Behmard E, Abdulabbas HT, Abdalkareem Jasim S, et al. Design of a novel multi-epitope vaccine candidate against hepatitis C virus using structural and nonstructural proteins: an immunoinformatics approach. *PLoS ONE*. 2022;17:e0272582. doi:10.1371/journal.pone.0272582.
35. Tahir ul, Qamar M, Shahid F, Aslam S, et al. Reverse vaccinology assisted designing of multiepitope-based subunit vaccine against SARS-CoV-2. *Infect Dis Poverty*. 2020;9:132. doi:10.1186/s40249-020-00752-w.
36. Karpenko LI, Bazhan SI, Eroshkin AM, Antonets DV, Chikaev AN, Ilyichev AA. Artificial epitope-based immunogens in HIV-vaccine design. In: *Advances in HIV and AIDS Control*. IntechOpen; 2018. doi:10.5772/intechopen.77031
37. Mintaev RR, Glazkova DV, Orlova OV, Bogoslovskaya EV, Shipulin GA. Development of a universal epitope-based influenza vaccine and evaluation of its effectiveness in mice. *Vaccines*. 2022;10(4):534. doi:10.3390/vaccines10040534
38. Li W, Joshi M, Singhania S, Ramsey K, Murthy A. Peptide vaccine: progress and challenges. *Vaccines*. 2014;2:515-536. doi:10.3390/vaccines2030515
39. Suschak JJ, Williams JA, Schmaljohn CS. Advancements in DNA vaccine vectors, non-mechanical delivery methods, and molecular adjuvants to increase immunogenicity. *Hum Vaccin Immunother*. 2017;13:2837-2848. doi:10.1080/21645515.2017.1330236
40. Jäschke A, Helm M. RNA sex. *Chem Biol*. 2003;10:1148-1150. doi:10.1016/j.chembiol.2003.12.003.
41. UniProt. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res*. 2019;47:D506-D515. doi:10.1093/nar/gky1049
42. Yu NY, Wagner JR, Laird MR, et al. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics*. 2010;26:1608-1615. doi:10.1093/bioinformatics/btq249
43. Garg A, Gupta D. VirulentPred: a SVM based prediction method for virulent proteins in bacterial pathogens. *BMC Bioinformatics*. 2008;9:62. doi:10.1186/1471-2105-9-62
44. Altschul S. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. 1997;25:3389-3402. doi:10.1093/nar/25.17.3389
45. Gasteiger E, Hoogland C, Gattiker A, et al. Protein Identification and Analysis Tools on the ExPASy Server. In: The Proteomics Protocols Handbook. Humana Press; 2005:571-607. doi:10.1385/1-59259-890-0
46. Krogh A, Larsson B, von Heijne G, Sonnhammer ELL. Predicting transmembrane protein topology with a hidden

Markov model: application to complete genomes11Edited by F. Cohen. *J Mol Biol*. 2001;305:567-580. doi:10.1006/jmbi.2000.4315
47. Doytchinova IA, Flower DR. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics*. 2007;8:4. doi:10.1186/1471-2105-8-4
48. Vita R, Mahajan S, Overton JA, et al. The Immune Epitope Database (IEDB): 2018 update. *Nucleic Acids Res*. 2019;47:D339-D343. doi:10.1093/nar/gky1006
49. Bui HH, Sidney J, Dinh K, Southwood S, Newman MJ, Sette A. Predicting population coverage of T-cell epitope-based diagnostics and vaccines. *BMC Bioinformatics*. 2006;7:153. doi:10.1186/1471-2105-7-153
50. Jespersen MC, Peters B, Nielsen M, Marcatili P. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Res*. 2017;45:W24-W29. doi:10.1093/nar/gkx346
51. Magnan CN, Randall A, Baldi P. SOLpro: accurate sequence-based prediction of protein solubility. *Bioinformatics*. 2009;25:2200-2207. doi:10.1093/bioinformatics/btp386
52. Jones DT. Protein secondary structure prediction based on position-specific scoring matrices 1 1Edited by G. *Von Heijne. J Mol Biol*. 1999;292:195-202. doi:10.1006/jmbi.1999.3091
53. Cheng J, Randall AZ, Sweredoski MJ, Baldi P. SCRATCH: a protein structure and structural feature prediction server. *Nucleic Acids Res*. 2005;33:W72-W76. doi:10.1093/nar/gki396.
54. Ponomarenko J, Bui HH, Li W, et al. ElliPro: a new structure-based tool for the prediction of antibody epitopes. *BMC Bioinformatics*. 2008;9:514. doi:10.1186/1471-2105-9-514
55. Craig DB, Dombkowski AA. Disulfide by Design 2.0: a web-based tool for disulfide engineering in proteins. *BMC Bioinformatics*. 2013;14:346. doi:10.1186/1471-2105-14-346
56. Abraham MJ, Murtola T, Schulz R, et al. GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*. 2015;1-2:19-25. doi:10.1016/j.softx.2015.06.001
57. Kozakov D, Hall DR, Xia B, et al. The ClusPro web server for protein–protein docking. *Nat Protoc*. 2017;12:255-278. doi:10.1038/nprot.2016.169
58. López-Blanco JR, Aliaga JI, Quintana-Ortí ES, Chacón P. IMODS: internal coordinates normal mode analysis server. *Nucleic Acids Res*. 2014;42:W271-W276. doi:10.1093/nar/gku339
59. Rapin N, Lund O, Bernaschi M, Castiglione F. Computational immunology meets bioinformatics: the use of prediction tools for molecular binding in the simulation of the immune system. *PLoS ONE*. 2010;5:e9862. doi:10.1371/journal.pone.0009862
60. Shepherd JG, Wang L, Reeves PR. Comparison of O-antigen gene clusters of *Escherichia coli* (*Shigella*) sonnei and Plesiomonas shigelloides O17: sonnei gained its current plasmid-borne O-antigen genes from P. Shigelloides in a recent event. *Infect Immun*. 2000;68:6056-6061. doi:10.1128/IAI.68.10.6056-6061.2000
61. Wen QF, Liu S, Dong C, Guo HX, Gao YZ, Guo FB. Geptop 2.0: an updated, more precise, and faster geptop server for identification of prokaryotic essential genes. *Front Microbiol*. 2019;10:1236. doi:10.3389/fmicb.2019.01236
62. Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, et al. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol*. 2019;37:420-423. doi:10.1038/s41587-019-0036-z
63. Vita R, Overton JA, Greenbaum JA, et al. The immune epitope database (IEDB) 3.0. *Nucleic Acids Res*. 2015;43:D405-D412. doi:10.1093/nar/gku938
64. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem*. 2009;31:455-461. doi:10.1002/jcc.21334
65. Lamiable A, Thévenet P, Rey J, Vavrusa M, Derreumaux P, Tufféry P. PEP-FOLD3: faster de novo structure prediction for linear peptides in solution and in complex. *Nucleic Acids Res*. 2016;44:W449-W454. doi:10.1093/nar/gkw329.
66. Dimitrov I, Bangov I, Flower DR, Doytchinova I., AllerTOP v.2—a server for in silico prediction of allergens. *J Mol Model*. 2014;20:2278. doi:10.1007/s00894-014-2278-5
67. Gupta S, Kapoor P, Chaudhary K, et al. In silico approach for predicting toxicity of peptides and proteins. *PLoS ONE*. 2013;8:e73957. doi:10.1371/journal.pone.0073957.
68. Buchan DWA, Jones DT. The PSIPRED protein analysis Workbench: 20 years on. *Nucleic Acids Res*. 2019;47:W402-W407. doi:10.1093/nar/gkz297
69. Debret G, Martel A, Cuniasse P. RASMOT-3D PRO: a 3D motif search webserver. *Nucleic Acids Res*. 2009;37:W459-464. doi:10.1093/nar/gkp304
70. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc*. 2015;10:845-858. doi:10.1038/nprot.2015.053
71. Heo L, Park H, Seok C. GalaxyRefine: protein structure refinement driven by side-chain repacking. *Nucleic Acids Res*. 2013;41:W384-W388. doi:10.1093/nar/gkt458
72. Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res*. 2007;35:w407–w410. doi:10.1093/nar/gkm290
73. Rapin N, Lund O, Castiglione F. Immune system simulation online. *Bioinformatics*. 2011;27:2013-2014. doi:10.1093/bioinformatics/btr335
74. Gerstner C, Dubnovitsky A, Sandin C, et al. Functional and structural characterization of a Novel HLA-DRB1*04:01-restricted α-enolase T cell epitope in rheumatoid arthritis. *Front Immunol*. 2016;7:494. doi:10.3389/fimmu.2016.00494
75. Hassan C, Chabrol E, Jahn L, et al. Naturally processed non-canonical HLA-A*02:01 presented peptides. *J Biol Chem*. 2015;290:2593-2603. doi:10.1074/jbc.M114.607028
76. Q S Medeiros PH, Ledwaba SE, Bolick DT, et al. A murine model of diarrhea, growth impairment and metabolic disturbances with Shigella flexneri infection and the role of zinc deficiency. *Gut Microbes*. 2019;10:615-630. doi:10.1080/19490976.2018.1564430
77. Rogawski McQuade ET, Scharf RJ, Svensen E, et al. Impact of *Shigella* infections and inflammation early in life on child growth and school-aged cognitive outcomes: findings from three birth cohorts over eight years. *Plos Negl Trop Dis*. 2022;16:e0010722. doi:10.1371/journal.pntd.0010722
78. Levine MM, Kotloff KL, Barry EM, Pasetti MF, Sztein MB. Clinical trials of *Shigella* vaccines: two steps forward and one step back on a long, hard road. *Nat Rev Microbiol*. 2007;5:540-553. doi:10.1038/nrmicro1662
79. van der Put RMF, Smitsman C, de Haan A, et al. The first-in-human synthetic glycan-based conjugate vaccine candidate against Shigella. *ACS Cent Sci*. 2022;8:449-460. doi:10.1021/acscentsci.1c01479

80. Kanampalliwar AM. Reverse vaccinology and its applications. *Methods Mol Biol* 2020;2131:1-16. doi:10.1007/978-1-0716-0389-5_1
81. Koebnik R, Locher KP, Van Gelder P. Structure and function of bacterial outer membrane proteins: barrels in a nutshell. *Mol Microbiol*. 2000;37:239-253. doi:10.1046/j.1365-2958.2000.01983.x
82. Ray A, Biswas T. Porin of Shigella dysenteriae enhances Toll-like receptors 2 and 6 of mouse peritoneal B-2 cells and induces the expression of immunoglobulin M, immunoglobulin G2a and immunoglobulin A. *Immunology*. 2005;114:94-100. doi:10.1111/j.1365-2567.2004.02002.x
83. Chanin RB, Nickerson KP, Llanos-Chea A, et al. Shigella flexneri Adherence Factor Expression in In Vivo-Like Conditions. *mSphere* 2019;4(6):e00751-19.
84. Jin Q. Genome sequence of *Shigella flexneri* 2a: insights into pathogenicity through comparison with genomes of *Escherichia coli* K12 and O157. *Nucleic Acids Res*. 2002;30:4432-4441. doi:10.1093/nar/gkf566
85. Gouma S, Vermeire T, Van Gucht S, et al. Differences in antigenic sites and other functional regions between genotype A and G mumps virus surface proteins. *Sci Rep*. 2018;8:13337. doi:10.1038/s41598-018-31630-z
86. Clarkson KA, Talaat KR, Alaimo C, et al. Immune response characterization in a human challenge study with a Shigella flexneri 2a bioconjugate vaccine. *EBioMedicine*. 2021;66:103308. doi:10.1016/j.ebiom.2021.103308
87. Jehl SP, Nogueira CV, Zhang X, Starnbach MN. IFNγ inhibits the cytosolic replication of *Shigella flexneri* via the cytoplasmic RNA sensor RIG-I. *PLoS Pathog* 2012;8(8):e1002809. doi:10.1371/journal.ppat.1002809.
88. Jennison AV, Verma NK. *Shigella flexneri* infection: pathogenesis and vaccine development. *FEMS Microbiol Rev* 2004;28(1):43-58. doi:10.1016/j.femsre.2003.07.002.
89. Jaimes JA, Andre NM, Millet JK, Whittaker GR. Structural modeling of 2019-novel coronavirus (nCoV) spike protein reveals a proteolytically-sensitive activation loop as a distinguishing feature compared to SARS-CoV and related SARS-like coronaviruses. *bioRxiv*. 2020.
90. Srivastava S, Kamthania M, Kumar Pandey R, et al. Design of novel multi-epitope vaccines against severe acute respiratory syndrome validated through multistage molecular interaction and dynamics. *J Biomol Struct Dyn*. 2019;37:4345-4360. doi :10.1080/07391102.2018.1548977
91. Jafari Najaf Abadi MH, Abdi Abyaneh F, Zare N, et al. In silico design and immunoinformatics analysis of a chimeric vaccine construct based on Salmonella pathogenesis factors. *Microb Pathog*. 2023;180:106130. doi:10.1016/j.micpath.2023.106130
92. Shahbazi S, Badmasti F, Habibi M, et al. In silico and in vivo investigations of the immunoreactivity of *Klebsiella pneumoniae* OmpA Protein as a Vaccine Candidate. *Iran Biomed J*. 2024;28:156-167. doi:10.61186/ibj.4023
93. Shi J, Zhu Y, Yin Z, et al. In silico designed novel multi-epitope mRNA vaccines against *Brucella* by targeting extracellular protein BtuB and LptD. *Sci Rep*. 2024;14:7278. doi:10.1038/s41598-024-57793-6
94. Asadinezhad M, Khoshnood S, Asadollahi P, et al. Development of innovative multi-epitope mRNA vaccine against Pseudomonas aeruginosa using in silico approaches. *Brief Bioinform*. 2023;25:bbad502. doi:10.1093/bib/bbad502
95. Wang H, Lin Q, Liu M, et al. Molecular docking and site-directed mutagenesis of GH49 family dextranase for the preparation of high-degree polymerization isomaltooligosaccharide. *Biomolecules*. 2023;13:300. doi:10.3390/biom13020300
96. Liu MA. A comparison of plasmid DNA and mRNA as vaccine technologies. *Vaccines*. 2019;7:37. doi:10.3390/vaccines7020037.
97. Zhang H, Zhang L, Lin A, et al. Algorithm for optimized mRNA design improves stability and immunogenicity. *Nature*. 2023;621:396-403. doi:10.1038/s41586-023-06127-z