In Silico-Based Vaccine Design Against Hepatopancreatic Microsporidiosis in Shrimp

Sk Injamamul Islam[1,2,3,*], Moslema Jahan Mou[4] and Saloa Sanjida[5]

[1]Department of Fisheries and Marine Bioscience, Faculty of Biological Science, Jashore University of Science and Technology, Bangladesh 7408, India
[2]Center of Excellence in Fish Infectious Diseases, Department of Veterinary Microbiology, Faculty of Veterinary Science, Chulalongkorn University, Bangkok 10330, Thailand
[3]The International Graduate Program of Veterinary Science and Technology, Department of Veterinary Microbiology, Faculty of Veterinary Science and Technology, Chulalongkorn University, Bangkok 10330, Thailand
[4]Department of Genetic Engineering and Biotechnology, Faculty of Life and Earth Science, University of Rajshahi, Bangladesh, India
[5]Department of Environmental Science and Technology, Faculty of Applied Science and Technology, Jashore University of Science and Technology, Bangladesh 7408, India

(*Corresponding author's e-mail: 6378506331@student.chula.ac.th)

Abstract

Enterocytozoon hepatopenaei (EHP) is a spore-forming intracellular parasite that produces an economically devastating disease in farmed shrimp called hepatopancreatic microsporidiosis (HPM), known to slow the growth of hosts in aquaculture. Spore wall proteins are implicated in the recognition of host cells in microsporidia. Many parasites have been demonstrated to detect heparin, a glycosaminoglycan (GAG) molecule present on cell surfaces, as a host. The sequencing and characterization of several EHP strains started to reveal information regarding pathogen biology, pathogenicity, and spore wall proteins must bind to the shrimp cell surface to infect it, and failure to do so results in ineffective infection. Despite these devastating complications, there is still no cure or vaccine for the parasites. In this regard, an immunoinformatics method was used to generate an epitope-based vaccine against this pathogen. The immunodominant T-cell and B-cell epitopes were identified using the spore wall proteins of EHP. The final constructed vaccine sequence was developed to be immunogenic, non-allergenic as well as have better solubility. Molecular dynamics simulation revealed significant binding stability and structural compactness. Finally, using *Escherichia coli* K12 as a model, codon optimization yielded ideal GC content and a higher CAI value, which was then included in the cloning vector pET2+ (a). Altogether, our outcomes imply that the proposed peptide vaccine may be a good option for HPM prophylaxis.

Keywords: *Enterocytozoon hepatopenaei*, Hepatopancreatic microsporidiosis, Vaccine, Epitopes, MD Simulation

Introduction

Litopenaeus vannamei, worldwide known as pacific white leg shrimp has been the most culture species for a decade [1]. This species is a prominent aquaculture crustacean, with a farm production scale of 5,446,216 tons and USD 32,191 million in 2019 [2]. However, disease plays a significant role in production losses. Hepatopancreatic microsporidiosis (HPM), a slow growth condition caused by an infection of the microsporidian parasite *Enterocytozoon hepatopenaei* (EHP), has been documented in Asia as a cause of mortality or slow growth disorder in shrimp farming. In Thailand, HPM was initially identified as a sign of a growth abnormality in the Black tiger shrimp (*Penaeus monodon*) [3]. So far, the primary hosts infected with EHP are recorded as Black tiger shrimp, Pacific white leg shrimp, and blue shrimp (*Litopenaeus stylirostris*), with large outbreaks documented in South Asian nations, Australia, and Venezuela [3-5].

EHP involves 2 life phases: An extracellular stage in the digestive tubule lumen with an active (mature) spore phase, and many intracellular spore-forming stages in the digestive epithelial cells of the

shrimp hepatopancreas [6]. It infects hepatic and pancreatic epithelial tubular epithelial cells of shrimp via its intracytoplasmic spores [7]. Furthermore, EHP established a unique invasion mechanism involving the polar tube and the spore wall [8]. GAGs (glycosaminoglycans) are glycoproteins that are capable of interacting with spore wall proteins (SWPs) during this first step of infection [9]. The functional characterization of SWPs contains 3 heparin-binding motifs and 1 Bin-amphiphysin-Rvs-2 domain which located in the electron-dense exospore and chitin dense endospore layers to make interaction with heparin and is responsible for the HPM disease in shrimp [10].

The prompt discovery of safe, efficient, uncomplicated, economical, dependable immune responses to the directed antigen emerges quickly and is made possible by the *in-silico* design of epitope vaccines against pathogens. Epitope-based vaccines have been successfully created in the postgenomic period to stimulate responsiveness against some of the worst human viruses, including Influenza, Nipah, Zika, Ebola, MARS-CoV, Rota, and others [11-15]. Lately, an in-silico technique was effective in predicting epitopes and multiepitopes with significant responsiveness against *Streptococcus agalactiae, Edwardsiella tarda*, and *Flavobacterium columnarie* separately [16-18]. Experts expect that in the coming days, computer-assisted techniques would be increasingly successful in controlling fish diseases [19,20]. As a result, the main objective of this research was to identify epitopes from the best antigenic spore wall proteins to fight infection against EHP.

Materials and methods

Protein selection and antigenicity prediction

NCBI database was used to retrieve Spore wall proteins (SWPs) of *Enterocytozoon hepatopenaei* (EHP) for antigen selection. There are 20 SWP and among them, 2 proteins possess the pseudo sequence and 5 proteins have the same sequence identity. Therefore, after core evaluation among these proteins, 3 proteins with high virulence and antigenicity were selected (Supplementary file). After retrieving all the SWPs, VirulentPred (http://bioinfo.icgeb.res.in/virulent/) and VaxiJen v2.0 server were used to predict the virulence and antigenicity score. The VirulentPred program predicts the sequences of virulent proteins to screen for them in proteomes using the SVM method [21]. Further, the selected 3 SWPs were examined for epitope-based vaccine design because of their direct role in pathogenesis. All the chosen protein sequences were obtained as FASTA files. VaxiJen v2.0 server was used to assess the protective antigens [22], AllerTop v2.0 was used for allergenicity score, and the ToxinPred server was used for toxicity. For each of the servers, a threshold value of 0.4 was chosen.

Epitopes prediction
Epitope prediction and evaluation of cytotoxic T-lymphocytes (CTLs)
CTLs are one of several types of immune system cells that can directly destroy other infectious cells [23]. They immediately enter the viral cell and contribute to the host's defensive response. The sequence of the chosen protein was entered into the EpiQuest™ 2014 (EpiQuest Suite EpiQuest-T) (http://www.epiquest.co.uk) software to predict CTLs epitope. EpiQuest is a unique software suitable for the analysis of linear protein sequences for the presence of B-cell and T-cell epitopes, area complexity (immunological, functional). It is based on new algorithms developed by Aptum Bio, and so far, is unparalleled by other software. The predicted epitopes were further assessed through the VaxiJen v2.0 [22], Immunogenicity [24], ToxinPred [25], and AllerTop v2.0 [26] servers. All of the forecasts were made using the default parameters of each server.
Epitopes of linear B-lymphocytes: Prediction and evaluation
To promote humoral or antibody-mediated immunity, B-cell epitopes are required. B-cells are made up of amino acid groups that bind with secreted antibodies and stimulate the immune system to fight infections [27]. As a result, the iBCE-EL server was used with default parameters to predict the linear B-lymphocyte (LBL) epitopes [28]. The anticipated epitopes were tested using VaxiJen v2.0, and AllerTop v2.0 servers.

Development of an epitope vaccine

Vaccine was created by combining the chosen CTL and LBL epitopes with a suitable adjuvant and linking them with the proper linkers [29,30]. Since viral glycoproteins recognize TLR4, and adjuvants are essential for overcoming the constraints of translation and synthesis, the adjuvant used here was TLR4 agonist [31,32]. As a result, the adjuvant 50S ribosomal protein L7/L12 (NCBI ID: P9WHE3) was evaluated to boost vaccine candidate's immunogenicity. With the RS09 (APPHALS), the PADRE sequence (AKFVAAWTLKAAA) can break apart 2 b domains with weakly interacting over a wide range of peptide

lengths. The adjuvant RS09 is a synthetic agonist for the Toll-like receptor-4. It aids in the activation of both the innate and adaptive immune systems [33]. Innate immunity is activated by toll-like receptors, as well as antigen presentation by Antigen Presenting Cells (APCs). A PADRE sequence provides vaccine stability as well as adjuvant properties [34]. In contrast, the selected CTL was linked with the help of AAY linkers and the LBL was linked with the KK linker [27,29]. The AAY linker is a proteasome cleavage site that has been exploited to modify protein stability, decrease immunogenicity, and improve epitope presentation [35]. The bi-lysine KK linker helps maintain the separate immunogenic properties of the vaccine construct.

Structural analysis of vaccine

Physiochemistry of a protein describes its fundamental characteristics. Physicochemical properties of the vaccine were predicted by the ProtParam server to gain a comprehensive understanding of the vaccine's essential role [36]. The immunological properties was evaluated through VaxiJen v2.0 [22], Immunogenicity [24], AllerTop [26], and SOLpro [37] servers. SOPMA (Self-Optimized Prediction Method with Alignment) server identifies the 2-dimensional (2D) structural features of the construct, such as the alpha-helix, beta-turn, and random coils [38] and PSIPRED v4.0 server [39] with default parameters. SOPMA has a prediction accuracy above 80 % [38]. To further understand the vaccine's composition quality, 2D structural characteristics were retrieved and assessed.

Prediction and confirmation of tertiary structure

Constructed vaccine was submitted to the RaptorX server (http://raptorx.uchicago.edu/) [40]. Using a cutting-edge algorithm and a 3D structure, the RaptorX server produces the most precise structure of the protein and its activities [40]. The C-score, TM-score value, RMSD, and top 5 models of a particular protein sequence may all be predicted and determined using this web service. The generated 3D structure was saved as a PDB file, which was chosen based on the C-score. The C-score on the server ranges from -5 to 2, with a higher number indicating a more confident protein model. For refining of the vaccine structure, discovered 3D structure was uploaded to the GalaxyRefine (http://galaxy.seoklab.org/refine) online web-based server. CASP10 refining approach was used to operate this webserver [41]. The RMSD, energy score, and overall quality score are all available on the GalaxyRefine website. The improved structure was downloaded, and the chosen structure was determined using the energy scores of the lowest and maximum RMSD values. PyMOL v2.3.4 was used to show the refined and discovered structure [42]. Analyzing the final 3D structure, the Ramachandran plot score (vaccine structure validity) and Z-score value were used, which indicates the standard deviation from the mean value. PROCHECK server was used to analyze Ramachandran plot, which runs most allowed and disallowed amino acid regions, and ProSA-web to analyze Z-score plot [43].

Disulfide engineering of the designed vaccine

To move forward and begin docking analysis, a designed model must be stable. Disulfide bonds provide a geometrically stable protein structure. Disulfide by Design 2.0 [44] was used to assign such bonds for the designed vaccine.

Molecular docking

The binding interactions between modeled proteins and receptor molecules are revealed through molecular docking experiments. For this, we used the ClusPro v2.0 server, which is available at https://cluspro.bu.edu, to submit the refined vaccine model as a ligand and the TLR4 protein as an immunological receptor for molecular docking [45]. TLR4 receptor was chosen and downloaded from the PDB server (PDB ID: 4G8A). Separating the associated ligand from protein was the first step in preparing the receptor, followed by the removal of water and other chemicals. All of these procedures were carried out using the PyMOL v2.3.4 program [42]. Discovery Studio 2017 and PBDSum were used to investigate binding interactions and residues in the interacting surface.

MD simulation

Complex structure of the selected candidate compounds was evaluated using 50 nanoseconds (ns) molecular dynamic simulations (MDS) to evaluate their binding stability to the desired protein to the active site cavity of the protein [46]. MDS of receptor-ligand complex was performed using the 'Desmond v6.3 Program' in Schrödinger 2020-3 under Linux framework to evaluate the thermodynamic stability of the receptor-ligand complex [47]. To solve the system, a predetermined TIP3P water model was used, with an orthorhombic periodic boundary box form with a box distance of 10 Å assigned to both sides to retain a specific volume. Following the construction of the solvated system with protein in complex with the ligand,

system has been minimized using OPLS_2005 force field parameters in addition to the standard protocol introduced in the Desmond module [47]. Finally, the simulation was carried out for 50 ns, and RMSF, RMSD, and Protein secondary structure elements from the trajectories were analyzed to reveal the stability of the vaccine complex.

Codon adaptation and *in silico* cloning

To express a foreign gene in a host, codon optimization is necessary [48]. As a result, the construct was uploaded to the JCat service for codon adaptation (http://jcat.de/). Commonly used *E. coli* K12 was used as the host in this study, and the entire procedure was carried out avoiding the following 3 criteria: Sites of restriction enzyme cleavage, binding sites of prokaryotic ribosomes, and rho-independent transcription termination. The codon adaptation index (CAI) value and guanine-cytosine (GC) concentration of the modified sequence was used to evaluate this [48]. Lastly, the *in silico* cloning of the adapted nucleotide sequence into the pET28a (+) expression vector was performed using the modified nucleotide sequence. SnapGene v4.2 software was used to conduct the entire *in silico* cloning procedure [49].

Results and discussion

Results
Analysis of the proteins

Vaxijen Server predicts the antigenicity of the selected spore wall proteins (**Table 1**). Furthermore, the allergenicity and toxicity scores were analyzed for these proteins came out negative.

Table 1 Details of the antigenic vaccine candidate proteins of EHP.

Name of protein	Accession number	Antigenicity
spore wall protein 12	QLM03045.1	0.4965
spore wall protein 7	QLM03044.1	0.6572
spore wall protein 26	QLM03046.1	0.6332

Prediction of epitopes

The selected target proteins were screened for CTL and LBL epitopes. In total, 13 distinct CTL epitopes for spore wall protein 12, 19 for spore wall protein 7 and 17 for spore wall protein 26 were predicted by the software. A list of the top 2 CTL epitopes for each protein which was non-toxic, non-allergenic, non-toxic, and immunogenic, was compiled (**Table 2**).

Table 2 Final CTL epitopes.

Protein name	Epitopes	C-score	Immunogenicity	Allergenicity	Antigenicity	Toxicity
Spore wall protein 12	STTKIRETL	0.9644	Positive	Negative	0.5056	Negative
Spore wall protein 12	TTKIRETLK	0.5981	Positive	Negative	1.0918	Negative
Spore wall protein 7	MKPFGLFLN	0.0376	Positive	Negative	1.9994	Negative
Spore wall protein 7	KPFGLFLNI	0.8769	Positive	Negative	1.6815	Negative
Spore wall protein 26	QTRTTHEAS	0.0589	Positive	Negative	0.8759	Negative
Spore wall protein 26	QKTIFTIKV	0.9594	Positive	Negative	0.8888	Negative

Furthermore, 77 unique LBL epitopes were predicted for spore wall protein 12, 59 LBL epitopes for spore wall protein 7 and 91 epitopes for spore wall protein 26 based on their toxicity, immunogenicity, antigenicity, and non-allergenicity (**Table 3**). Among the top 2 proteins with the best probability,

antigenicity, allergenicity, and toxicity were chosen for final vaccine construction. The length of the epitope was selected at 12 (12-mer peptide).

Table 3 Final LBL epitopes.

Protein name	Sequence	Probability	Antigenicity	Allergenicity	Toxicity
Spore wall protein 12	PKKISTTKIRET	0.8016	0.5266	Negative	Negative
Spore wall protein 12	KKISTTKIRETL	0.7737	0.6101	Negative	Negative
Spore wall protein 7	NKKLDKKDKDAY	0.7852	0.7303	Negative	Negative
Spore wall protein 7	DKKDKDAYPVES	0.7520	1.0012	Negative	Negative
Spore wall protein 26	KVQVNLVLDPQE	0.7527	0.8566	Negative	Negative
Spore wall protein 26	NEDLFKYKVQVN	0.7194	0.9211	Negative	Negative

The structure of a vaccine and its fundamental characteristics

Vaccine was created utilizing epitopes from 2 distinct classes (CTL and LBL) that had previously been chosen. AAY and KK linkers were used to link epitopes together. To enhance immunogenicity, an adjuvant was applied before the construct. An adjuvant was attached to the CTL epitope with the RS09 (APPHALS) and PADRE sequence as a linker to activate TLR4 by ribosomal protein the 50S/L12 as an agonist. The final vaccination had a length of 303 amino acids (**Figure 1**).

MAKLSTDELLDAFKEMTLLELSDFVKKFEETFEVTAAAPVAVAAAGAAPAGAAVEAAEEQSEFDVILEAAGDKKIGVIKVVREI
VSGLGLKEAKDLVDGAPKPLLEKVAKEAADEAKAKLEAAGATVTVKAPPHALSAKFVAAWTLKAAASTTKIRETLAAYTTKIRE
TLKAAYMKPFGLFLNAAYKPFGLFLNIAAYQTRTTHEASAAYQKTIFTIKVKKPKKISTTKIRETKKKKISTTKIRETLKKNKKLDKK
DKDAYKKDKKDKDAYPVESKKKVQVNLVLDPQEKKNEDLFKYKVQVN

Figure 1 Constructed vaccine sequence.

### Immunological assessment and physicochemical characteristics

Table 4 shows the physicochemical parameters of the vaccine construct. The construct was discovered to have a molecular weight of 33,209.78 Da. Other features such as the theoretical isoelectric point (pI) of 9.59, the chemical formula of $C_{1507}H_{2477}N_{391}O_{440}S_3$, the instability index of 16.41, the aliphatic index of 87.43, and the grand average of hydropathicity of −0.350 were also present. The construct physicochemical properties and immunological efficacy were also assessed. For example, the construct antigenicity was 0.5717, whereas its immunogenicity was .92208. Moreover, with a score of 0.89019 out of 1, the vaccine proved soluble (**Table 4**). Alpha-helix, beta-strand, and random coils were examined utilizing 2 distinct servers (SOPMA and PSIPRED) as secondary structural characteristics (**Table 5**). On the other hand, the PSIPRED server anticipated the features as 44.88 % α-helix, 13.86 % β-strand, and 41.25 % random coils (**Table 5**) (**Figure 2**).

Table 4 Characteristics of the construct in terms of antigenicity, allergy, and physicochemical properties.

Characteristics	Finding	Remark
Number of amino acids	303	Suitable
Molecular weight	33,209.78	Average
Theoretical pI	9.59	Base
Chemical formula	$C_{1507}H_{2477}N_{391}O_{440}S_3$	-
Instability index of vaccine	16.41	Stable
Aliphatic index of vaccine	87.43	Thermostable
Grand average of hydropathicity (GRAVY)	−0.350	Hydrophilic
Antigenicity	0.5717	Antigenic
Immunogenicity	Positive	Immunogenic
Allergenicity	No	Non-allergen
Solubility	0.89019	Soluble

**Table 5** The vaccine's secondary structural characteristics.

| Characters	SOPMA		PSIPRED server	
	AA	%	AA	%
Alpha helix	162	53.47	136	44.88
Beta strand	18	5.94	42	13.86
Random coil	78	25.74	125	41.25
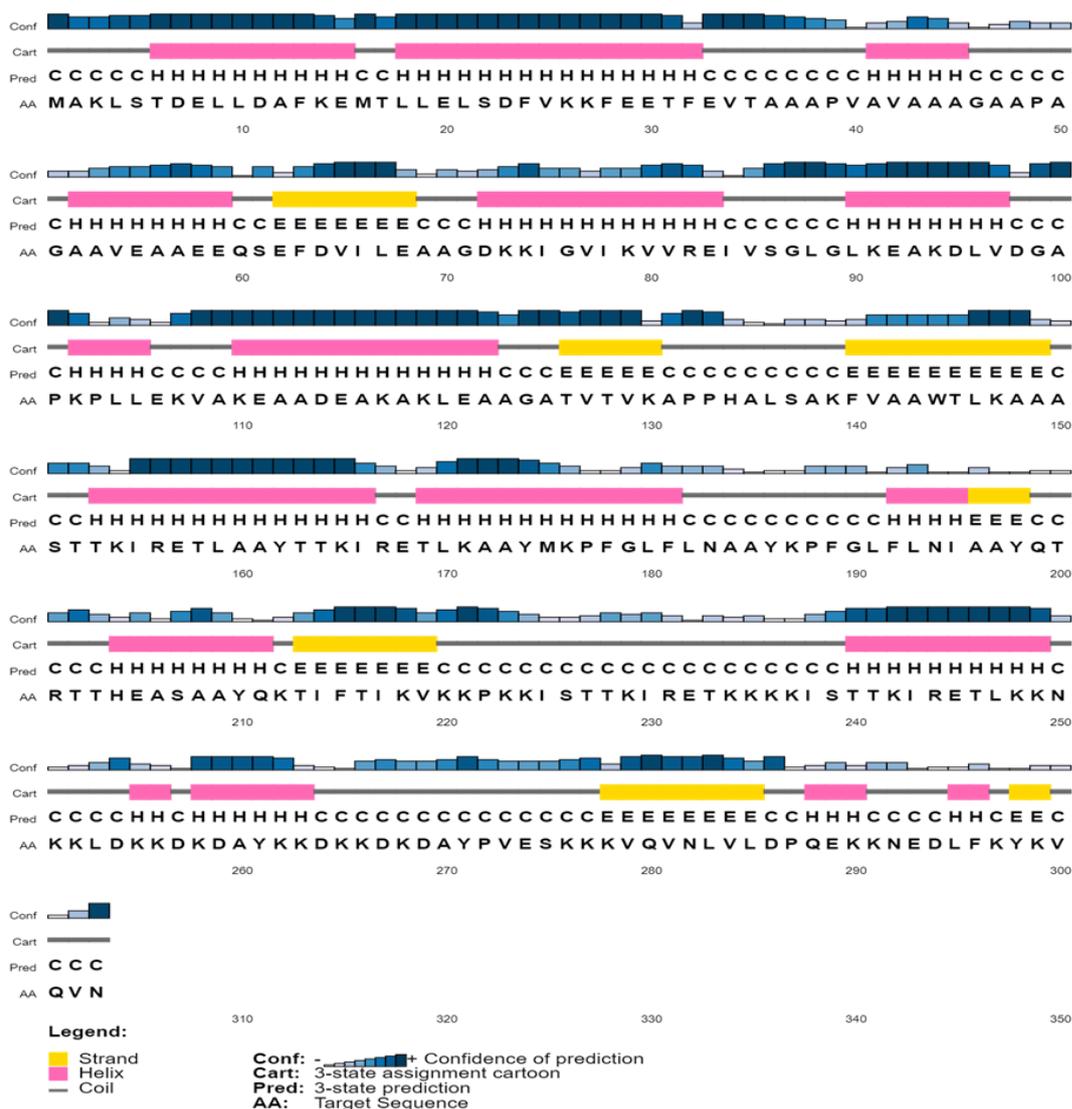
Figure 2 Prediction of the secondary structure of designed multi-epitope vaccines using PSI-PRED.

Refinement and confirmation of the tertiary structure

The top 5 homology models were built using RaptorX as the best template. Out of the 5 models, the one with the lowest C-score (−4.73) was selected. The 3D representation of the produced vaccine is shown in **Figure 3**. Before refinement, the Ramachandran plot of the vaccine revealed that 77.1 % of residues were in the most favorable zone, 20.4 % in the additional allowed region, and 0.7 % in the disallowed region. However, after refinement, 80.7 % of the residues in the most favorable region were found in the Ramachandran plot, and 16.8 % in additional allowed regions, while 1.8 % in disallowed regions were seen in **Figure 4(B)**. The Z score was −5.32 of the crude model whereas Z-score was –5.65 of the refine model (**Figure 4(D)**).
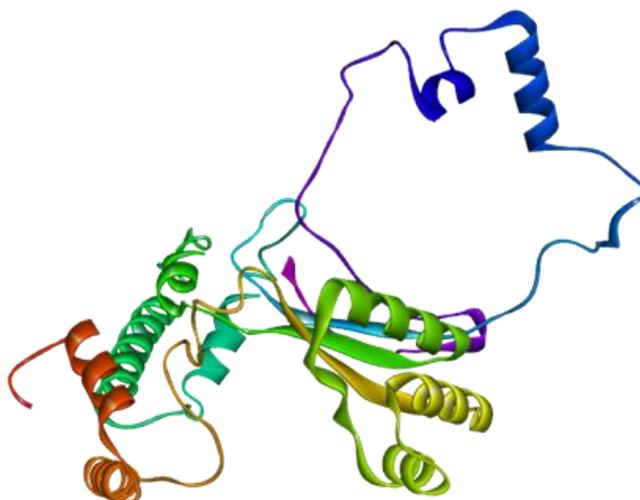
Figure 3 3D structure of the designed vaccine construct.
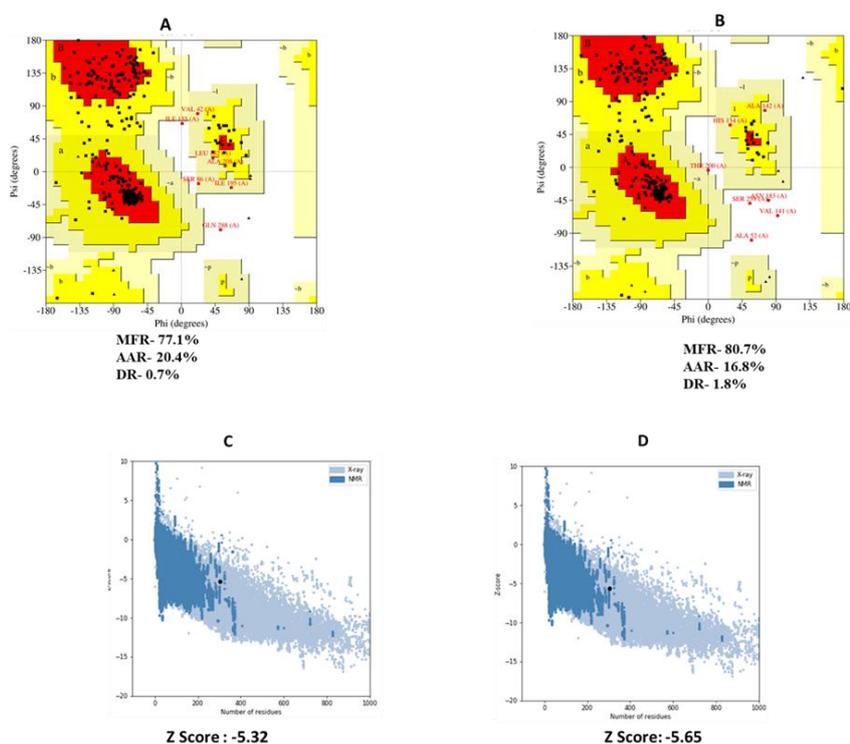
Figure 4 (A-B) Analysis of Ramachandran plot PROCHECK server. The MFR, AAR, GAR, and DR were represented as the most favored, additional allowed, generously al-lowed, and disallowed regions of the vaccine. (C-D) 3-D structure validation with a Z-score by Pro-SA server.

Vaccine disulfide engineering

To stabilize the vaccine design, disulfide engineering was employed. In the case of the vaccine, the DbD2 server found that there were 30 pairs of amino acids with the potential to make disulfide bonds. With other parameters such as energy and the chi3 value taken into account, 2 pairs of mutations with cysteine were recommended. Being larger chi3 means more electrons, and thus higher polarizability, which creates a stronger intermolecular bond between receptor and vaccine. Thus, Ser37-CYS47 and LEU89-CYS155

were the residue pairs with the highest number of mutations. Energy and chi3 have approved values of less than 4.88 and 102.19: −119.49, correspondingly.

Molecular docking research

To predict their binding affinity and interactions, the vaccine (ligand) and TLR4 (receptor) were docked. ClusPro v2.0 server produced 10 docked complexes in various positions. The complex with the lowest energy score was chosen and the binding posture with functional interactions from among them. Thus model 1 met the inclination criterion. So, it was chosen as the best vaccine–TLR4 complex, with a −967.5 energy score. Binding interactions and residues implicated in active site residues were investigated in the chosen complex. A total of 9 hydrogen bonds were found on the interaction surface. The interacting residues in the CHB from the vaccine were Ile84-Thr125; Ala50-Asn57; Ala50-Thr36; Lys3-Leu98 (**Figure 5(A)**) and Gln80-Glu62; Gln80-Lys107; Thr36-Ala50; Thr125-Ile84 and Leu598-Lys3 (**Figure 5(B)**). Moreover, associated TLR4 active site residues are shown in **Figure 5**.
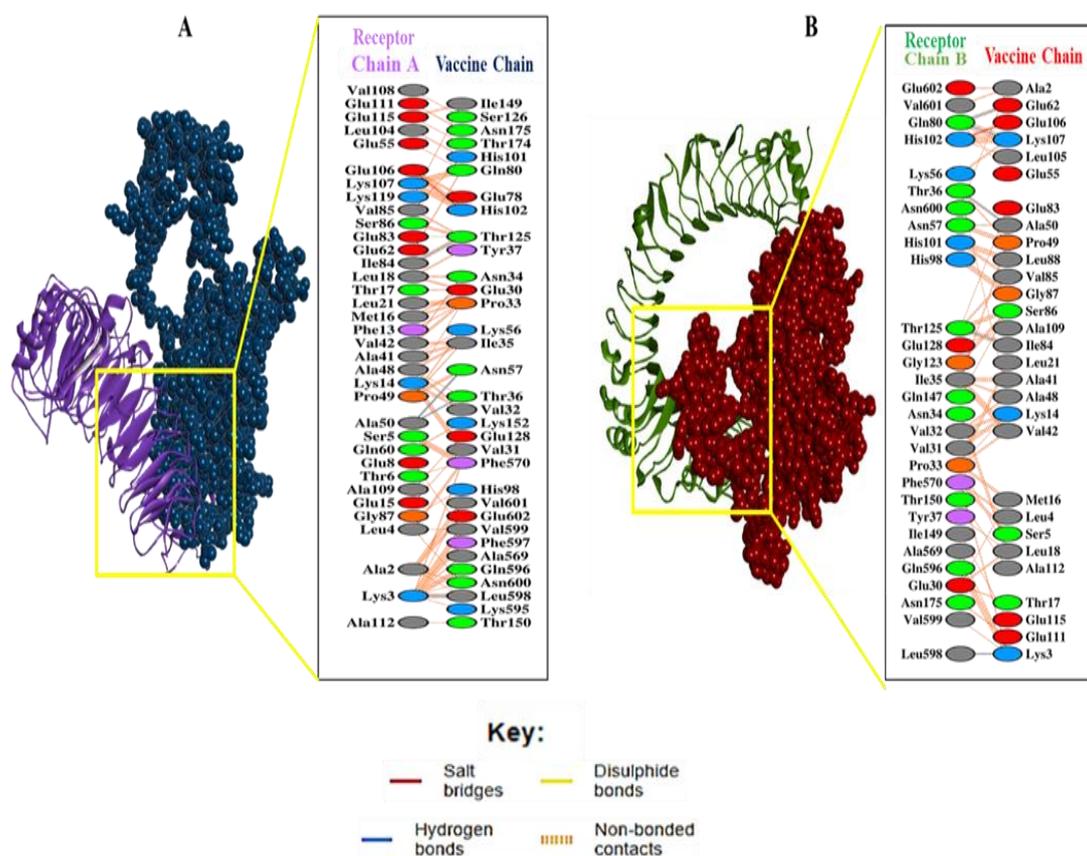

Figure 5 Analysis of EBV-immune receptor binding conformation and interaction. (A) Intermolecular binding mode and residue-level chemical interactions of MEBV-TLR4 receptor chain (A) and (B) Intermolecular binding mode and residue-level chemical interactions of MEBV-TLR4 receptor chain (B).

MD simulation

The root mean square deviation (RMSD) of both the vaccine complex and vaccine was calculated. The average RMSD value for the vaccine complex was 4.86Å, which demonstrates the structural stability during the interaction. From **Figure 6**, it was observed that the vaccine complex has the initial increase of RMSD descriptors till 10 ns, and after that, it showed stability till 25 ns. A lower degree of fluctuation was observed from 7 - 8 ns, which may be responsible for structural integrity and/or to allow firm binding. Furthermore, the root mean square fluctuation (RMSF) score was used to assess protein flexibility across amino acid residues. The RMSF profile of the vaccine complex indicates maximum amino acid residues from complexes that an RMSF profile below 4.0 Å and greater change were observed for fewer residues. **Figure 7** shows the stability and stiffness of the vaccine complex.
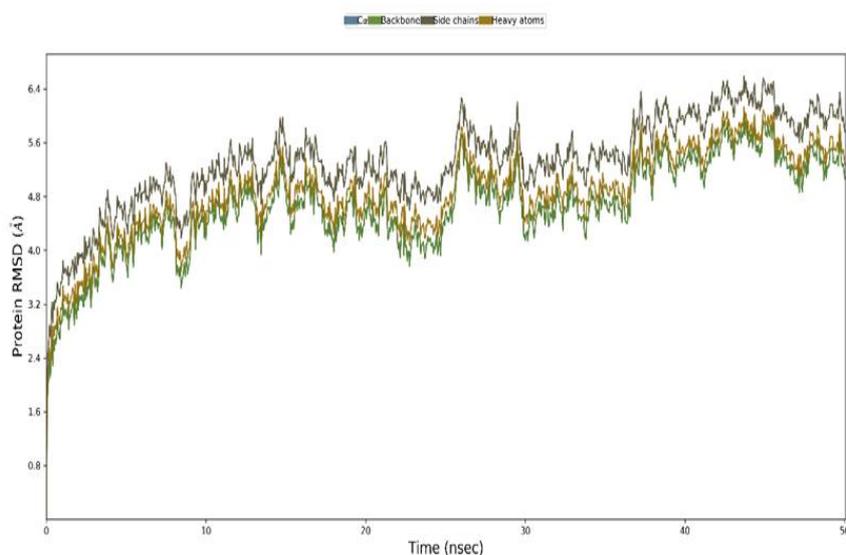
Figure 6 Simulation of the epitope vaccine complex at the molecular level. The backbone atoms of the complexes were plotted using the RMSD method
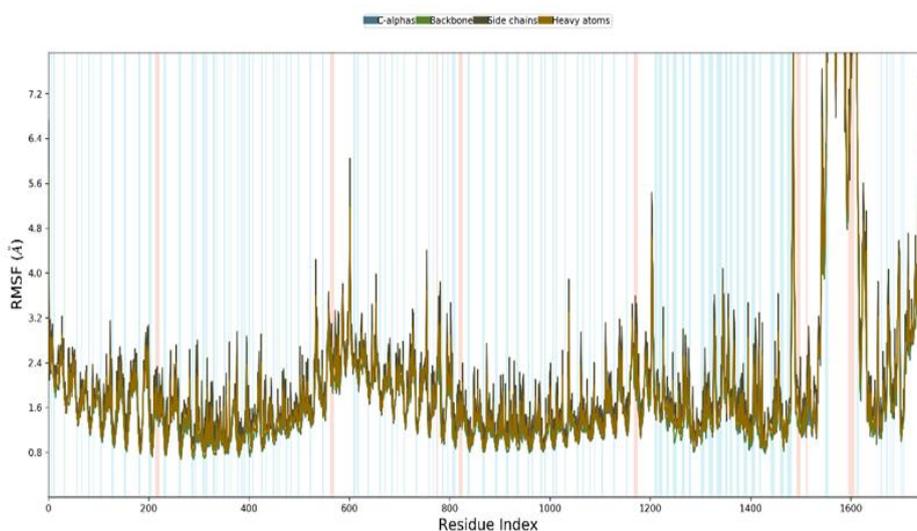
Figure 7 Simulation of the epitope vaccine complex at the molecular level; The multi-epitope docked vaccine candidate RMSF plot; Red and blue backgrounds emphasize the alpha-helical and beta-strand sections, respectively; These areas are defined by helices or strands that last for 70 % of the simulation time.

Codon adaptation and in silico cloning

To improve the translation efficiency of the vaccine design, the codons according to the *E. coli* K12 on the JCat service were adjusted. The nucleotide sequences created by the peptide vaccine construct (303 AA residues) totaled 558 lengths (**Figure 8**). Furthermore, the modified nucleotide sequence has a GC content of 45.87 % and a CAI value of 1.0, respectively. To insert the changed sequence into the pET28a (+) vector, the XhoI and BamHI restriction sites as the start and end cut points, respectively were employed. The improved vaccine design was cloned into the pET28a (+) cloning vector using the SnapGene software (**Figure 9**).

```
ATGGCTATCCGTGACTAAGCTTAAGAATAACTGGACGACTGCAAAATCGA    50
AAACCTGTCTACCCGTGCTCGTGCTTGCGCTCGTGCTCCGGCTGGTGCTT   100
AAGCTCGTGACCGTCTGTCTATCGAACGTGGTAAAAAATGGGTTTAAGTT   150
TGGAACGTTGTTGTTCGTGCTCGTTGCGAAGGTCCGACCCTGAAAGTTCA   200
GGAACGTGAATCTAAACAGTAAGCTCGTGCTCGTACCTAAGCTCCGCAGC   250
TGCAGCTGTGCGCTGGTCTGTCTGCTCACACCAACGGTACCGCTCTGACC   300
CAGATGGAACTGTCTGCTAACGAACCGTGGTTCTAAGCTCTGAACCTGGG   350
TTTCAAATGCGCTACCACCACCCCGGAACGTGCTACCAAACACCTGCACA   400
AATAAAACCAGAAATACACCAACGGTACCAAAAAAAACACCCAGATGGAA   450
CTGAAAAAAAAATAAAAACGTAAATCTTAAAAAGAAAAAGAACGTGCTAC   500
CGTTAAAAAATCTCAGTAATGCTAACCGCGTAAAAAAGAAATCTAATAAT   550
CTCAGTAA
```

Figure 8 Codon adaptation of EBV to *E. coli* K12 strain.
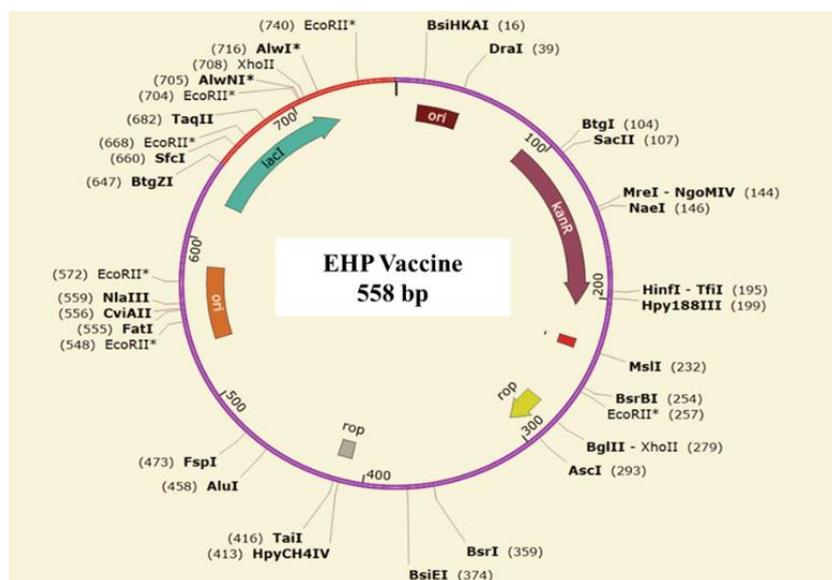

Figure 9 The proposed vaccine was cloned into the pET-28a (+) vector *in silico*.

Discussion

The current diabolical emergence of infectious from Hepatopancreatic microsporidiosis (HPM) poses a danger to the worldwide aquaculture shrimp industry, which influences to use of an immunoinformatics method to build this multi-epitope vaccine against this disease. The vaccination based on the most virulent and antigenic spore wall proteins (SWTs) of *Enterocytozoon hepatopenaei* (EHP) displayed outstanding relevance as predicted by immunoinformatics, proving this study to be reliable. A vaccine protects against infectious illnesses safely and effectively [50]. Vaccines have the ability to provide acquired immunity against infectious diseases [51]. As a result of this study, the designed vaccine based on epitopes would provide a strong immune response to HPM. EHP infection and transmission are difficult to control and prevent in the absence of an effective vaccine. Furthermore, to regulate the current situation, effective immunization has yet to be produced. As a result, a novel vaccine development strategy is critical to finding a solution to the current economically threatening aquaculture problem. Because the selected SWTs of EHP are important for immunological invasion and transmission, the goal of the study was to develop an epitope vaccination that targeted the most virulent and antigenic proteins. To enable cellular and humoral immune systems to recognize this protein, all the virulent and antigenic SWTs selected through *in silico* screening were evaluated for their antigenic region. The first step was identifying all possible CTL and LBL epitopes. Because the linkers below corresponded to the top epitopes, vaccines were created with 2 antigenic epitopes, CTL and LBL. These epitopes were used in vaccine development as an important component that improves the stability, folding, and transcriptional regulation of our peptide vaccine [52]. Adjuvants are attached to CTL epitopes with RSO9 and PADREE sequences as linkers, which make the vaccine more

stable and durable, as well as enhance cellular and humoral immunogenic responses [53]. A total of 303 amino acid residues were found in the vaccine construction. An essential characteristic of a recombinant vaccine is its solubility, a type of physicochemical property [54]. A solubility assessing tool was used to determine whether the vaccine construct was solvable inside the host *E. coli*, and the results showed that it was solvable. The nature of the constructed vaccine, as indicated by the theoretical PI value, was basic. As recommended by server tools, the stability index of the vaccine sequence indicates that it is stable following synthesis. The GRAVY value and aliphatic index, on the other hand, indicated that the vaccine was hydrophobic and thermostable, respectively. According to the prediction of physicochemical properties and scores on all parameters, there is a high probability for this vaccine to be a valid candidate against EHP infection. The detected models were revised and the best model (based on the lowest energy score) was chosen after the 3D structure prediction (based on c-score). The Z-score (–5.65) and superior features of the most favored, acceptable, and prohibited areas for the Ramachandran plot in the validation test of the 3D structure were observed. It was suggested by the lowest energy score of Vaccine-TLR4 complexes, respectively, from a molecular docking suggested that the vaccine could have infection-inhibiting activity and might interact tightly with these receptors chain A and chain B. The molecular dynamics simulation is a potentially useful tool for understanding how proteins function and how their structure is derived. Anatomical movement can be stimulated by protein dynamic simulations as a function of time. MD simulation of the vaccine candidate was performed for 50 ns and analyzed the results used the RMSD and RMSF scores. When comparing distinct atomic conformations of a molecular system, the RMSD value is employed. Significant flexibility and departure of vaccine candidates from receptor structure were determined using the RMSD value, whereas the displacement of particular vaccine candidate's atoms from receptor structure was determined using RMSF of the complex structure. The calculated average RMSD and RMSF value was 4.86 Å and 4.0 Å, respectively. The fluctuation was not observed to be larger in the vaccine section, however, it smoothed out after 5 ns, suggesting that the modeled vaccine and receptor are stable. To optimize the multi-epitope vaccine production, the MD simulation was done to evaluate the stability of the vaccine candidate with the receptor, in which codon optimization was done for the stability of the construct vaccine within the host. Eventually, the codon was adjusted, and *in silico* cloning of the intended vaccine candidate into the *E. coli* K12 expression host pET28a (+) vector was successful.

Conclusions

EHP spore wall proteins and structure are extensively studied, yet controlling disease from it still lacking. A range of computational techniques was used in this work to find possible B and T-cell epitopes in EHP virulent and antigenic spore wall proteins, which were finally stitched into an epitope vaccine. Recently developed vaccine possesses the immuno-dominant qualities that are sought. Significantly, it was capable of binding to the immunological receptors and induces a substantial immune response in favor of EHP infection. Based on the findings of this study, the preparation of a vaccine against the etiological mediator of the EHP epidemic in shrimp must start with the vaccine candidate. In addition, the possible epitopes discovered in this study can be employed or re-modeled in future research. Nevertheless, more wet lab analysis is needed to show that our designed vaccine is an effective prevention against HPM disease.

Acknowledgments

References

[1]　N Ghaffari, A Sanchez-Flores, R Doan, KD Garcia-Orozco, PL Chen, A Ochoa-Leyva, AA Lopez-Zavala, JS Carrasco, C Hong and LG Brieba. Novel transcriptome assembly and improved annotation of the whiteleg shrimp (Litopenaeus vannamei), a dominant crustacean in global seafood mariculture. *Sci. Rep.* 2014; **4**, 7081.

[2]　U Lele and S Goswami. *The food and agriculture organization of the United Nations*. Food and Agriculture Organization, Québec, Canada, 2021, p. 630-706.

[3]　K Chayaburakul, G Nash, P Pratanpipat, S Sriurairatana and B Withyachumnarnkul. Multiple pathogens found in growth-retarded black tiger shrimp Penaeus monodon cultivated in Thailand. *Dis. Aquat. Organisms* 2004; **60**, 89-96.

[4]　KF Tang, CR Pantoja, RM Redman, JE Han, LH Tran and DV Lightner. Development of in situ hybridization and PCR assays for the detection of Enterocytozoon hepatopenaei (EHP), a microsporidian parasite infecting penaeid shrimp. *J. Invertebr. Pathol.* 2015; **130**, 37-41.

[5]   DJ Aldama-Cano, P Sanguanrut, N Munkongwongsiri, JC Ibarra-Gámez, O Itsathitphaisarn, R Vanichviriyakit, TW Flegel, K Sritunyalucksana and S Thitamadee. Bioassay for spore polar tube extrusion of shrimp Enterocytozoon hepatopenaei (EHP). *Aquaculture* 2018; **490**, 156-61.

[6]   T Chaijarasphong, N Munkongwongsiri, GD Stentiford, DJ Aldama-Cano, K Thansa, TW Flegel, K Sritunyalucksana and O Itsathitphaisarn. The shrimp microsporidian Enterocytozoon hepatopenaei (EHP): Biology, pathology, diagnostics and control. *J. Invertebr. Pathol.* 2021; **186**, 107458.

[7]   KFJ Tang, JE Han, LF Aranguren, B White-Noble, MM Schmidt, P Piamsomboon, E Risdiana and B Hanggono. Dense populations of the microsporidian Enterocytozoon hepatopenaei (EHP) in feces of penaeus vannamei exhibiting white feces syndrome and pathways of their transmission to healthy shrimp. *J. Invertebr. Pathol.* 2016; **140**, 1-7.

[8]   Q Lv, L Wang, Y Fan, X Meng, K Liu, B Zhou, J Chen, G Pan, M Long and Z Zhou. Identification and characterization a novel polar tube protein (NbPTP6) from the microsporidian Nosema bombycis. *Parasites Vectors.* 2020; **13**, 475.

[9]   JR Hayman, TR Southern and TE Nash. Role of sulfated glycans in adherence of the microsporidian Encephalitozoon intestinalis to host cells *in vitro*. *Infect. Immun.* 2005; **73**, 841-8.

[10] P Jaroenlak, DW Boakye, R Vanichviriyakit, BAP Williams, K Sritunyalucksana and O Itsathitphaisarn. Identification, characterization and heparin binding capacity of a spore-wall, virulence protein from the shrimp microsporidian, Enterocytozoon hepatopenaei (EHP). *Parasites Vectors* 2018; **11**, 177.

[11] JE Muñoz-Medina, CJ Sánchez-Vallejo, A Méndez-Tenorio, IE Monroy-Muñoz, J Angeles-Martínez, AS Coy-Arechavaleta, CE Santacruz-Tinoco, J González-Ibarra, YM Anguiano-Hernández and CR González-Bonilla. *In Silico* identification of highly conserved epitopes of influenza A H1N1, H2N2, H3N2, and H5N1 with diagnostic and vaccination potential. *BioMed Res. Int.* 2015; **2015**, 813047.

[12] MT Ali, MM Morshed and F Hassan. A computational approach for designing a universal epitope-based peptide vaccine against nipah virus. *Interdiscipl. Sci.* 2015; **7**, 177-85.

[13] S Anwar, JT Mourosi, MF Khan and MJ Hosen. Prediction of epitope-based peptide vaccine against the chikungunya virus by immuno-informatics approach. *Curr. Pharmaceut. Biotechnol.* 2020; **21**, 325-40.

[14] R Dash, R Das, M Junaid, MF Akash, A Islam and SZ Hosen. *In silico*-based vaccine design against Ebola virus glycoprotein. *Comput. Biol. Chem. Adv. Appl.* 2017; **10**, 11-28.

[15] J Shi, J Zhang, S Li, J Sun, Y Teng, M Wu, J Li, Y Li, N Hu, H Wang and H Yunzhang. Epitope-based vaccine target screening against highly pathogenic MERS-CoV: An *in silico* approach applied to emerging infectious diseases. *PLoS One* 2015; **10**, e0144475.

[16] R Mahendran, S Jeyabaskar, G Sitharaman, RD Michael and AV Paul. Computer-aided vaccine designing approach against fish pathogens edwardsiella tarda and flavobacterium columnare using bioinformatics softwares. *Drug Des. Dev. Ther.* 2016; **10**, 1703-14.

[17] UP Pereira, SC Soares, J Blom, CA Leal, RT Ramos, LC Guimarães, LC Oliveira, SS Almeida, SS Hassan and AR Santos. *In silico* prediction of conserved vaccine targets in streptococcus agalactiae strains isolated from fish, cattle, and human samples. *Genet. Mol. Res.* 2013; **12**, 2902-12.

[18] A Pumchan, S Krobthong, S Roytrakul, O Sawatdichaikul, H Kondo, I Hirono, N Areechon and S Unajak. Novel chimeric multiepitope vaccine for streptococcosis disease in nile tilapia (Oreochromis niloticus Linn.). *Sci. Rep.* 2020; **10**, 603.

[19] A Madonia, C Melchiorri, S Bonamano, M Marcelli, C Bulfon, F Castiglione, M Galeotti, D Volpatti, F Mosca, T Pietro-Giorgio and N Romano. Computational modeling of immune system of the fish for a more effective vaccination in aquaculture. *Bioinformatics* 2017; **33**, 3065-71.

[20] A Joshi, DC Pathak, MA Mannan and V Kaushik. *In-silico* designing of epitope-based vaccine against the seven banded grouper nervous necrosis virus affecting fish species. *Netw. Model. Anal. Health Informat. Biol.* 2021; **10**, 37.

[21] A Garg and D Gupta. VirulentPred: A SVM based prediction method for virulent proteins in bacterial pathogens. *BMC Bioinformatics* 2008; **9**, 62.

[22] IA Doytchinova and DR Flower. VaxiJen: A server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics* 2007; **8**, 4.

[23] B Farhood, M Najafi and K Mortezaee. CD8 + cytotoxic T lymphocytes in cancer immunotherapy: A review. *J. Cell. Physiol.* 2019; **234**, 8509-21.

[24] JJ Calis, M Maybeno, JA Greenbaum, D Weiskopf, ADD Silva, A Sette, C Keşmir and B Peters. Properties of MHC class I presented peptides that enhance immunogenicity. *PLoS Comput. Biol.* 2013; **9**, e1003266.

[25] S Gupta, P Kapoor, K Chaudhary, A Gautam, R Kumar and GP Raghava. *In silico* approach for predicting toxicity of peptides and proteins. *PLoS One* 2013; **8**, e73957.

[26] I Dimitrov, DR Flower and I Doytchinova. AllerTOP--a server for *in silico* prediction of allergens. *BMC Bioinformatics* 2013; **14**, S4.

[27] Z Nain, F Abdulla, MM Rahman, MM Karim, MSA Khan, SB Sayed, S Mahmud, SMR Rahman, MM Sheam, Z Haque and UK Adhikari. Proteome-wide screening for designing a multi-epitope vaccine against emerging pathogen Elizabethkingia anophelis using immunoinformatic approaches. *J. Biomol. Struct. Dynam.* 2020; **38**, 4850-67.

[28] B Manavalan, RG Govindaraj, TH Shin, MO Kim and G Lee. iBCE-EL: A new ensemble learning framework for improved linear b-cell epitope prediction. *Front. Immunol.* 2018; **9**, 1695.

[29] H Dorosti, M Eslami, M Negahdaripour, MB Ghoshoon, A Gholami, R Heidari, A Dehshahri, N Erfani, N Nezafat and Y Ghasemi. Vaccinomics approach for developing multi-epitope peptide pneumococcal vaccine. *J. Biomol. Struct. Dynam.* 2019; **37**, 3524-35.

[30] Z Nain, MM Karim, MK Sen and UK Adhikari. Structural basis and designing of peptide vaccine using PE-PGRS family protein of mycobacterium ulcerans-an integrated vaccinomics approach. *Mol. Immunol.* 2020; **120**, 146-63.

[31] J Olejnik, AJ Hume and E Mühlberger. Toll-like receptor 4 in acute viral infection: Too much of a good thing. *PLoS Pathogens* 2018; **14**, e1007390.

[32] RK Pandey, TK Bhatt and VK Prajapati. Novel immunoinformatics approaches to design multi-epitope subunit vaccine for malaria by investigating anopheles salivary protein. *Sci. Rep.* 2018; **8**, 1125.

[33] A Shanmugam, S Rajoria, AL George, A Mittelman, R Suriano and RK Tiwari. Synthetic toll like receptor-4 (TLR-4) agonist peptides as a novel class of adjuvants. *PLoS One* 2012; **7**, e30839.

[34] M Hasan, S Islam, S Chakraborty, AH Mustafa, KF Azim, ZF Joy, MN Hossain, SH Foysal and MN Hasan. Contriving a chimeric polyvalent vaccine to prevent infections caused by herpes simplex virus (type-1 and type-2): An exploratory immunoinformatic approach. *J. Biomol. Struct. Dynam.* 2020; **38**, 2898-915.

[35] GS Abdellrazeq, LM Fry, MM Elnaggar, JP Bannantine, DA Schneider, WM Chamberlin, AHA Mahmoud, KT Park, V Hulubei and WC Davis, Simultaneous cognate epitope recognition by bovine CD4 and CD8 T cells is essential for primary expansion of antigen-specific cytotoxic T-cells following ex vivo stimulation with a candidate Mycobacterium avium subsp. paratuberculosis peptide vaccine. *Vaccine* 2020; **38**, 2016-25.

[36] M Wilkins, E Gasteiger, A Bairoch, JC Sanchez, K Williams, R Appel and D Hochstrasser. *Protein identification and analysis tools in the ExPASy server*. The Proteomics Protocols Handbook. Humana Press, New Jersey, 2008, p. 571-607.

[37] CN Magnan, M Zeller, MA Kayala, A Vigil, A Randall, PL Felgner and P Baldi. High-throughput prediction of protein antigenicity using protein microarray data. *Bioinformatics* 2010; **26**, 2936-43.

[38] C Geourjon and G Deléage SOPMA: Significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Comput. Appl. Biosci.* 1995; **11**, 681-4.

[39] DW Buchan, F Minneci, TC Nugent, K Bryson and DT Jones. Scalable web services for the PSIPRED protein analysis workbench. *Nucleic Acids Res.* 2013; **41**, W349-57.

[40] J Xu, M McPartlon and J Li. Improved protein structure prediction by deep learning irrespective of co-evolution information. *Nat. Mach. Intell.* 2021; **3**, 601-9.

[41] T Nugent, D Cozzetto and DT Jones. Evaluation of predictions in the CASP10 model refinement category. *Proteins* 2014; **82**, 98-111.

[42] WL DeLano. PyMOL: An open-source molecular graphics tool. *Adv. Biol. Chem.* 2002; **8**, 1-9.

[43] M Wiederstein and MJ Sippl. ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* 2007; **35**, W407-10.

[44] DB Craig and AA Dombkowski. Disulfide by design 2.0: A web-based tool for disulfide engineering in proteins. *BMC Bioinformatics* 2013; **14**, 346.

[45] D Kozakov, DR Hall, B Xia, KA Porter, D Padhorny, C Yueh, D Beglov and S Vajda: The ClusPro web server for protein-protein docking. *Nat. Protocol.* 2017; **12**, 255-78.

[46] S Pokhrel, TA Bouback, A Samad, SM Nur, R Alam, M Abdullah-Al-Mamun, Z Nain, RR Imon, MEK Talukder, MMI Tareq, MS Hossen, TM Karpiński, F Ahammad, I Qadri and MS Rahman.Spike protein recognizer receptor ACE2 targeted identification of potential natural antiviral drug candidates against SARS-CoV-2. *Int. J. Biol. Macromol.* 2021; **191**, 1114-25.

[47] TA Bouback, S Pokhrel, A Albeshri, AM Aljohani, A Samad, R Alam, MS Hossen, K Al-Ghamdi, MEK Talukder, F Ahammad, I Qadri and J Simal-Gandara. Pharmacophore-based virtual screening,

quantum mechanics calculations, and molecular dynamics simulation approaches identified potential natural antiviral drug candidates against MERS-CoV S1-NTD. *Molecules* 2021; **26**, 4961.

[48] A Grote, K Hiller, M Scheer, R Münch, B Nörtemann, DC Hempel and D Jahn. JCat: A novel tool to adapt codon usage of a target gene to its potential expression host. *Nucleic Acids Res.* 2005; **33**, W526-31.

[49] MF Goldberg, EK Roeske, LN Ward, T Pengo, T Dileepan, DI Kotov and MK Jenkins. Salmonella persist in activated macrophages in T cell-sparse granulomas but are contained by surrounding CXCR3 ligand-positioned Th1 cells. *Immunity* 2018; **49**, 1090-102.

[50] W Li, MD Joshi, S Singhania, KH Ramsey and AK Murthy. Peptide vaccine: Progress and challenges. *Vaccines* 2014; **2**, 515-36.

[51] KF Bol, EH Aarntzen, JM Pots, MAO Nordkamp, MWVD Rakt, NM Scharenborg, AJD Boer, TGV Oorschot, SA Croockewit and WA Blokx. Prophylactic vaccines are potent activators of monocyte-derived dendritic cells and drive effective anti-tumor responses in melanoma patients at the cost of toxicity. *Canc. Immunother.* 2016; **65**, 327-39.

[52] S Shamriz, H Ofoghi and N Moazami. Effect of linker length and residues on the structure and stability of a fusion protein with malaria vaccine application. *Comput. Biol. Med.* 2016; **76**, 24-9.

[53] SR Bonam, CD Partidos, SKM Halmuthur and S Muller. An overview of novel adjuvants designed for improving vaccine efficacy. *Trends Pharmacol. Sci.* 2017; **38**, 771-93.

[54] N Khatoon, RK Pandey and VK Prajapati. Exploring leishmania secretory proteins to design B and T cell multi-epitope subunit vaccine using immunoinformatics approach. *Sci. Rep.* 2017; **7**, 8285.