Turk. J. Fish.& Aquat. Sci. 19(2), 149-159
http://doi.org/10.4194/1303-2712-v19_2_07

RESEARCH PAPER

Turkish Journal of FISHERIES and AQUATIC SCIENCES

Design of Epitope Based Vaccine against Shrimp White Spot Syndrome Virus (WSSV) by Targeting the Envelope Proteins: an Immunoinformatic Approach

Farhana Momtaz[1], Javed Foysal[2,*], Mahbubur Rahman[3], Ravi Fotedar[4]

[1] Department of Microbiology, University of Chittagong, Chittagong 4331, Bangladesh.
[2] Department of Genetic Engineering and Biotechnology, Shahjalal University of Science and Technology, Sylhet-3114, Bangladesh.
[3] Department of Biotechnology, Bangabandhu Sheikh Mujibur Rahman Agricultural University, Gazipur 1706, Bangladesh.
[4] Department of Environment and Agriculture, School of Molecular and Life Sciences, Curtin University, WA6845, Australia.

Article History
Received 14 December 2017
Accepted 23 March 2018
First Online 25 March 2018

Corresponding Author
Tel.: +880.821 715542
E-mail: mjfoysal-geb@sust.edu

Keywords
WSSV
Envelope proteins
Epitope based vaccine
Immunoinformatics

Abstract

The shrimp white spot syndrome virus (WSSV) causes significant damage to aquaculture production worldwide but a vaccine, eliciting the immunogenicity of shrimps against WSSV has yet to be developed. Thus, a programmed immunoinformatics study was conducted to find out the potential immunogens based on genome-wide screening of WSSV envelope proteins. The measurements of the phylogenetic and evolutionary distances found the common geographical routes of three countries, where the proteins from other six countries were clustered together. Among all the four major envelope proteins i.e., VP19, VP24, VP26, and VP28; AAO69663.1 from VP26 showed the highest antigenicity and thus selected for further studies. The properties of the secondary and tertiary structure including the modelled 3D protein revealed that the protein had all the properties required for a protective immunogen. The peptide regions ranging from 99 to 115 and 98 to 106, representing the sequences "VTAPRTDPAGTGAENSN" and "TVTAPRTDP" were found to be most effective regions for B-cell linear and cytotoxic T lymphocyte (CTL), respectively. The CTL epitope also showed a strong and stable interaction with the MHC class I and class II molecules, reported to be found in fish. Therefore, the present epitope could be used as a potential vaccine candidate against WSSV.

Introduction

Aquaculture is one of the fastest growing sectors and shrimp culture reported to play a major role in boosting the seafood productions. Shrimp accounts for 15% of total seafood products with the market values of 102 billion dollars in 2008 (Karunasagar & Ababouch, 2012). However, shrimp white spot syndrome virus (WSSV) has been associated with major economic loss in shrimp farm industries worldwide, especially in South Asian countries (Vaseeharan, Jayakumar, & Ramasamy, 2003). Study reports revealed approximately 4-6 billons dollars economic loss in Asia, only in four years: from 1997 to 2001, and the damage still continues (Bir *et al*., 2017). The mortality of WSSV infections could reach to 100% within 10 days, resulting huge economic losses (Tuyen, Verreth, Vlak, & de Jong, 2014; Verbruggen *et al.*, 2016). Conventional strategies like improvement of environmental conditions, stocking of pathogen free juveniles, and use of broad spectrum drugs are currently applied to reduce damages but failed to produce positive outcomes (Ganjoor, 2015).

WSSV is a double stranded, rod shaped, enveloped DNA virus known to infect shrimp and other crustaceans (Sangamaheswaran & Jeyaseelan, 2001; Kua & Rashid, 2012). The genome of WSSV is completely sequenced and protein functions are revealed. Four envelope proteins (VP19, VP24, VP26, and VP28) having no homology with any other structural proteins from viruses, play crucial role in the viral pathogenesis (van Hulten, Goldbach, & Vlak, 2000; Tan & Shi, 2008; Zhou, Xu, Li, Qi, & Yang, 2009). Among these four, VP28 involved in viral attachment and penetration, VP24 and VP26 was initially regarded as nucleocapsid proteins; however, later also categorized as envelope proteins involved in DNA replication and help viruses in circumvention from the host immune responses (Tan & Shi, 2008; Tsai *et al.*, 2006; Wan, Xu, & Yang, 2008; Li, Li, Han, Xu, & Yang, 2015) Although VP19 categorized into envelope protein, its exact functional role and mechanism of assembly is yet unclear (Zhou, Xu, Li, Qi, & Yang, 2009).

Viral diseases are most difficult to control and vaccination is the only strategy to avert virus attack. No medicines yet to prove effective against WSSV; hence, finding ways to combat this shrimp pathogen now become to be indispensable. Oral vaccination with two structural proteins: VP19 and VP466, were unsuccessful in term of reducing mortality and envelope protein VP28 experimentally elicited the immune response by only some degrees (Ha *et al*, 2008; Syed Musthaq & Kwang, 2015). Another study with recombinant VP292 protein vaccine also showed unsatisfactory results *in-vivo* (mortality >50%) (Zhu, Du, Miao, Quan, & Xu, 2009). In addition, several herbal supplements also used into the feed as stimulants to defend against WSSV (Yogeeswaran *et al*., 2012). However, not a single study yet to claim more than 50% of shrimp survival in the laboratory conditions. Therefore, the development of new vaccines using a novel approach, i.e., searching an ideal immunogenic protein, would be the best alternative way to address this problem.

In-silico based design of epitope vaccines against viruses enable rapid discovery of safe, effective, effortless, inexpensive, reliable, and rapid development of immune responses against the directed antigen. In the post-genomic era, epitope based vaccines have been successfully designed to elicit immunogenicity against some deadliest human viruses like influenza, nipah, chikunguniya, zika, ebola, MARS-CoV, rota, etc. (Muñoz-Medina *et al*., 2015; Ali, Morshed, & Hassan, 2015; Kori, Sajjan, & Madagi, 2015; Mirza *et al*., 2017; Dash *et al.*, 2017; Shi *et al*., 2015; Ghosh, Chattopadhyay, Chawla-Sarkar, Nandy, & Nandy, 2012). However, in fish, the approach yet to be studied deeply because of proper knowledge on the differentiation between major histocompatibility complexes (MHC class I and II) and human leukocyte antigen (HLA) (Xu, Sun, Shi, Cheng, & Wang, 2011; Dijkstra, Grimholt, Leong, Koop, & Hashimoto, 2013; Grimholt, 2016). The experimental data of cord and tilapia reported to have both MHC class I and class II molecules for initiating immune responses against pathogens. Hence, the peptide that can target HLA-A*0201, HLA-B*3501, and HLA-B*3508 with excellent binding abilities, could be used as effective vaccines against selective fish pathogens (Grimholt, 2016; Mahendran, Jeyabaskar, Sitharaman, Michael, & Paul, 2016; Marana *et al*., 2017). Recently, *in-silico* approach was successfully used to predict epitopes, having strong immunogenicity against *Streptococcus agalactiae*, *Edwardsiella tarda*, and *Flavobacterium columnarie*: three pathogenic bacteria that cause streptococcosis, edwardsiellosis and columnaris in fish, respectively (Mahendran *et al.*, 2016; Pereira *et al*., 2013). Experts predict more effective control of fish diseases by computer aided tools in upcoming days (Madonia, Melchiorri, & Bonamano, 2017). Consequently, the principle aim of this study was to find out epitope from best antigenic protein to defend WSSV.

Materials and Methods

Retrieval of Protein Sequences from Data Bank

The present study was designed according to the published data of four major envelope and nucleocapsid proteins of WSSV, responsible for pathogenicity (van Hulten *et al*., 2000; Tan & Shi, 2008; Zhou *et al*., 2009; Tang & Hew, 2007). Sequences of VP19, VP24, VP26 and VP28 proteins of WSSV were downloaded from National Centre for Biotechnology Information (NCBI) database (https://www.ncbi.nlm.nih.gov/) in FASTA format (Table 1). Present study collected all the available (15 from VP19, 6 from VP24, 6 from VP26 and 40 from VP28 protein) full length (not partial or whole genome) envelope protein sequences of WSSV from nine different countries i.e., India, Thailand, China, Egypt, Mexico, Iran, Vietnam, South Korea, and Bangladesh; over the period of 18 years (from 2000 to 2017) (Table 1). The sequences were then subjected to different bioinformatics tools to predict antigenicity, secondary, and 3D structures as well as epitopes discovery and validation.

Selection of Antigenic Proteins

Any molecule that elicit the immune responses can be defined as antigen. To find out the best potential antigen, all of the 67 proteins were subjected to in-silico antigenicity analysis in web based VaxiJen (v2.0) server (http://ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html). VaxiJen server could predict antigenic proteins with approximately 80-90% of accuracy (Doytchinova & Flower, 2007). Non-antigenic proteins were excluded from the study and antigenic proteins were stratified according to orders (Table 1). The highest antigenic protein in VaxiJen was selected for further studies i.e., prediction of secondary and tertiary structure, modelling, protein-peptide docking as well as other criteria's and parameters to be considered as vaccine candidate.

Multiple Alignment and Phylogenetic Analyses

The downloaded sequences were aligned in MEGA 7.0 by muscle alignment tool. Muscle alignment is the better tool to align protein coding sequences and also to predict the phylogenetic relationships among the taxa or species (Kumar, Stecher, & Tamura, 2016). Proteins from the same families having similar isolation history (country, date) and VaxiJen scores were excluded from the phylogeny and evolutionary distance study. Phylogenetic tree was constructed using MEGA 7.0 by neighbour joining method in 1000 bootstrap replicates in the branches. The evolution divergence among the sequences were calculated as p-distances with transition and trans-version substitutions and complete deletion

Table 1. VaxiJen scores of the four major envelope proteins (VP19, VP24, VP26 and VP28) of WSSV in term of antigenicity

Protein	Accession number	Geographical origin	Submitted year	VaxiJen score	Remarks
VP19	ABG75925.1	India	2011	0.5860	Antigenic
VP19	AAP79433.1	South Korea	2005	0.5854	Antigenic
VP19	ACT78467.1	South Korea	2009	0.5854	Antigenic
VP19	AAO69661.1	China	2003	0.5817	Antigenic
VP19	ABZ80614.1	India	2008	0.5854	Antigenic
VP19	CAI79042.1	Mexico	2016	0.5765	Antigenic
VP19	BAP16759.1	Iran	2014	0.5854	Antigenic
VP19	ANN78124.1	Vietnam	2002	0.5854	Antigenic
VP19	AAW67477.1	India	2005	0.5854	Antigenic
VP19	ABS59293.1	India	2007	0.5854	Antigenic
VP19	ADE45323.1	India	2010	0.5854	Antigenic
VP19	AFO63166.1	India	2012	0.5854	Antigenic
VP19	ABI74630.1	India	2006	0.5792	Antigenic
VP19	ALB00655.1	Egypt	2015	0.5854	Antigenic
VP19	AAK77851.1	Thailand	2005	0.5701	Antigenic
VP24	ABG75922.1	India	2011	0.5757	Antigenic
VP24	ABA10826.1	Iran	2005	0.5194	Antigenic
VP24	AAO69622.1	China	2003	0.5792	Antigenic
VP24	AAK77700.1	Thailand	2005	0.5757	Antigenic
VP24	ABI74631.1	India	2006	0.5757	Antigenic
VP24	AAL33006.1	China	2014	0.5757	Antigenic
VP26	ABG75924.1	India	2016	0.7570	Antigenic
VP26	AAO69663.1	China	2003	0.7601	Antigenic
VP26	CAD83838.1	China	2003	0.7570	Antigenic
VP26	ABP52056.1	India	2007	0.7570	Antigenic
VP26	AAK77822.1	Thailand	2005	0.7570	Antigenic
VP26	AAF29806.1	Thailand	2000	0.7570	Antigenic
VP28	ABG75923.1	India	2011	0.5668	Antigenic
VP28	AKS25342.1	India	2015	0.5854	Antigenic
VP28	ADJ36912.1	India	2010	0.5798	Antigenic
VP28	ADJ36911.1	India	2010	0.5798	Antigenic
VP28	ADJ36910.1	India	2010	0.5798	Antigenic
VP28	ADJ36909.1	India	2010	0.5798	Antigenic
VP28	ADJ36908.1	India	2010	0.5768	Antigenic
VP28	ADJ36907.1	India	2010	0.5768	Antigenic
VP28	ADJ36906.1	India	2010	0.5768	Antigenic
VP28	ADJ36905.1	India	2010	0.5768	Antigenic
VP28	ADJ36904.1	India	2010	0.5768	Antigenic
VP28	ADJ36903.1	India	2010	0.5791	Antigenic
VP28	ADJ36902.1	India	2010	0.5768	Antigenic
VP28	ADJ36901.1	India	2010	0.5768	Antigenic
VP28	AAP87278.1	South Korea	2005	0.5668	Antigenic
VP28	AAT81591.1	China	2005	0.5668	Antigenic
VP28	ACT78468	South Korea	2009	0.5527	Antigenic
VP28	AAY40276.1	India	2005	0.5854	Antigenic
VP28	AAY40275.1	India	2005	0.5668	Antigenic
VP28	AAY40274.1	India	2005	0.5668	Antigenic
VP28	AAR12965	India	2005	0.5447	Antigenic
VP28	AAM27199	China	2002	0.5915	Antigenic
VP28	ABZ80613	India	2008	0.5668	Antigenic
VP28	ALB00646	Egypt	2015	0.5757	Antigenic
VP28	AJC52394	Bangladesh	2015	0.5694	Antigenic
VP28	ALK87217	Mexico	2016	0.5668	Antigenic
VP28	AAW67476	India	2005	0.5668	Antigenic
VP28	CAD83839.1	Vietnam	2005	0.5668	Antigenic
VP28	AFO63167.1	India	2012	0.5854	Antigenic
VP28	ABP52057.1	India	2007	0.5854	Antigenic
VP28	AAK77670.1	Thailand	2005	0.5668	Antigenic
VP28	AAF29807.1	Vietnam	2000	0.5668	Antigenic
VP28	ATE50874.1	Bangladesh	2017	0.5854	Antigenic
VP28	ADE45322.1	India	2010	0.5668	Antigenic
VP28	AFU90301.1	Vietnam	2012	0.6273	Antigenic
VP28	AFU90300.1	Vietnam	2012	0.5567	Antigenic
VP28	AFU90299.1	Vietnam	2012	0.5854	Antigenic
VP28	ABP01348.1	Thailand	2007	0.5641	Antigenic
VP28	ABI74633.1	India	2006	0.5668	Antigenic
VP28	ABS00973.1	India	2007	0.5951	Antigenic

settings in MEGA 7.0.

Evaluating the Structure of the Best Antigenic Protein

The secondary and tertiary structure gives a lot of information about the protein to be used as vaccine. The secondary structure of the highest antigenic protein was determined by ProtParam (http://web.expasy.org/protparam/), self-optimized prediction method for alignment (SOMPA) (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html), and Jpred (http://www.compbio.dundee.ac.uk/jpred/index.html). The three-dimensional (3D) structure was visualized by web based RaptorX (http://raptorx.uchicago.edu/StructurePrediction/predict/). RaptorX is very popular tool for reliable 3D model of proteins and is ranked first in the past couple of years (Källberg *et al*., 2012). ProtParam and Jpred can calculate various parameters like coil, helices, molecular weight (MW), atomic and amino acid compositions, theoretical pI, stability, hydrophilicity etc., that are very crucial in selecting the antigenic protein. All the generated data were analysed using standard methods for protein structure analysis (Shi *et al*., 2015; Mahendran *et al*., 2016; Patel, Koringa, Reddy, Nathani, & Joshi, 2015).

B-cell Linear Epitope Prediction

The B-cell epitopes were predicted by the method of Kolaskar and Tongaonkar 1990. These epitope prediction tool is available at the Immune Epitope Database (IEDB) (http://tools.iedb.org/main/bcell/) server. The method has been widely applied to predict epitopes from large number of proteins with more than 75% accuracy and hence, the efficiency is higher than other currently existing tools (Shi *et al*., 2015). In addition, Karplus and Schulz method on IEDB was applied to calculate flexibility, hydrophilicity and surface accessibility of the predicted epitope with default settings.

Prediction of Cytotoxic T-cell (CTL) Epitopes

The cytotoxic T lymphocyte (CTL) epitopes that can bind MHC class I and class II molecules were predicted using CTLPred (http://crdd.osdd.net/raghava/ctlpred/) (Bhasin & Raghava, 2004). Artificial neural network (ANN) and stabilized matrix method implementation (SMM) were used to predict the peptide binding ability to MHC molecules. Previously reported alleles in the vertebrates were considered prior to predict epitopes. The length of the epitope was selected at 9 (9-mer peptide) because of more positive (>75%) binding to MHC-I and MHC-II molecules (Haug *et al*., 2016). The epitope with highest antigenicity was then screened for binding scores to those MHC class I and class II molecules that are commonly found in fish.

Modelling and Molecular Docking of the Peptide to MHC Molecules

The modelling of the 9-mer epitope (TVTAPRTDP) was performed by PEP-FOLD3 (http://bioserv.rpbs.univ-paris-diderot.fr/services/PEP-FOLD3/) web based tool (Lamiable *et al*., 2016). T-epitope scores of 9-mer peptide sequences were calculated using online portal T-epitope designer (http://www.bioinformation.net/script/hla_search.cgi) for all the MHC class I and class II molecules those have been reported to be involved in the immunogenicity of fish in earlier studies (Mahendran *et al.*, 2016; Haug *et al.*, 2016). Molecular docking was performed by CABS-dock (http://biocomp.chem.uw.edu.pl/CABSdock) using HLA-A0201 (PDB code 1HLA) and HLA-B3508 (PDB code 1ZHL) with default settings (Ciemny, Kurcinski, Kozak, Kolinski, & Kmiecik, 2012; Kurcinski, Jamroz, Blaszczyk, Kolinski, & Kmiecik, 2015). Pymol (v2.0.5) was used to visualize 3D structure of docked protein-peptide. This molecular graphics system is an excellent visualization platform to analyse computational drugs (Yuan, Chan, & Hu, 2017).

Evaluation of Allergenicity of the Peptide

The selection of non-allergen protein and epitope was one of our prime aims. As the vaccine will be directly applied to the crustaceans and fish, the popular consumable vertebrates and also there are possibilities of human interactions i.e., handling vaccine and feeding fish. AllerTop (http://www.ddg-pharmfac.net/AllerTOP/) web based system allows assessment of allergenicity and allergic cross reactivity in proteins (Dimitrov, Bangov, Flower, & Doytchinova, 2014). AllerTop search all the known allergen protein and accurately find matches with high sensitivity and specificity.

Results

Analysis of the Envelope Protein Sequences of WSSV

Analysis of all four envelope protein sequences (VP19, VP24, VP26 and VP28) from nine different countries revealed two significant findings: (1) Proteins from other six countries (China, Mexico, Iran, Egypt, South Korea, and Thailand) clustered closely to India, Vietnam and Bangladesh; therefore, these three countries could be the common evolutionary origin for global spread of WSSV (Figure 1) (2) Nodes of VP19 and VP28 were poorly while VP24 and VP26 nodes were well supported by the published data. Evolutionary divergence analysis of all 67 sequences revealed four outcomes: (1) ABI74630.1 (India, 2006) from VP19 showed significant p-distances (0.037-0.028), followed

by both AAO69661.1 (China, 2003) and AAK77851.1 (Thailand, 2005) with the value of 0.009 (supplementary table), (2) ABA10826.1 from Iran under VP24 protein family showed significant p-distances (0.053), followed by AO69662.1 (0.005) from China, (3) only AAO69663.1 from China under VP26 proteins had evolutionary distances to other five VP26 proteins (0.005), (4) AFU90301.1 (Vietnam, 2012) from VP28 had the highest evolutionary distances (0.015-0.020) to other VP28 followed by ALB00646.1 (Egypt), AJC52394.1 (Bangladesh) (0.010-0.015), and ADJ36912.1 (India) (0.005-0.010), respectively.

Antigenic Features of the Envelope Proteins

The antigenicity values of all 67 full length envelope proteins were calculated by auto cross covariance (ACC) transformation of protein using VaxiJen 2.0. All of the proteins showed a varied number of antigenicity ranging from 0.5194 (ABA10826.1, Iran_2005, VP24) to 0.7601 (AAO69663.1, China_2003,

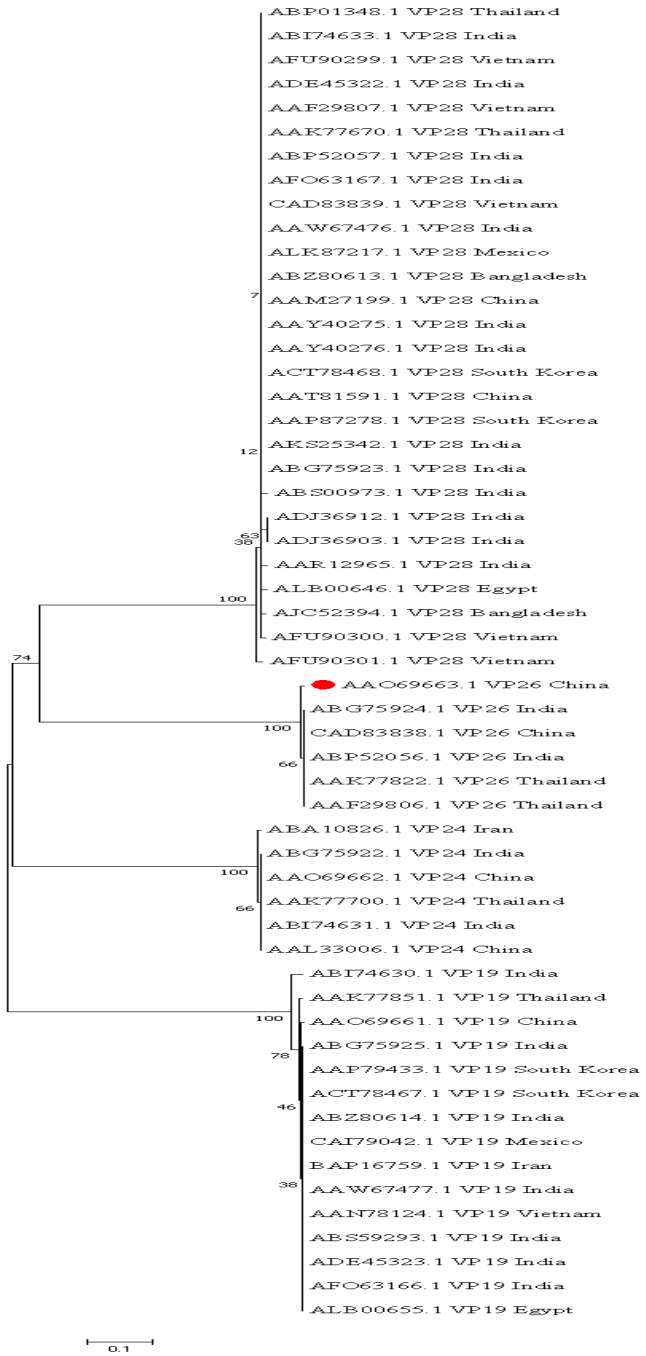

Figure 1. Phylogenetic relationship of all four envelope proteins sequences of WSSV. The percentages of the replicate in trees in which the associated taxa clustered together by the bootstrap test (1000 replicates) are shown next to the branches. The tree was constructed using neighbour-joining method in MEGA7. AAO69663.1 clustered separately from other VP26 proteins with significant bootstrap values of 99 and the only protein of VP26 family having evolutionary distances (0.009) in between them (VP26 cluster).

VP26) (Table 1). All of the five proteins from VP26 except AAO69663.1, exhibited similar antigenicity value of 0.7570. AAO69663.1 was the highest antigenic protein with the value of 0.7601. Antigenicity values of other three protein groups (VP19, VP24 and VP28) were varied in between 0.5 to 0.6. The highest antigenic protein, AAO69663.1 from VP26 was selected for further studies.

Evaluating the Secondary and Tertiary Structure

ProtPerm and Jpred secondary structure prediction tools revealed many characteristic features of the highest antigenic protein AAO69663.1. In the 204 amino acids, isoleucine (Ile) was found to be abundant in the chain (12.3%) followed by valine (8.8%), serine (8.3%), threonine (7.8%), asparagine (7.8%), alanine (7.4%), and lysine (7.4%). The protein contained total number of 3171 atoms, 59 alpha helices, 51 beta strands, 16 beta run with the atomic formula of $C_{970}H_{1623}N_{267}O_{299}S_{12}$ and molecular weight (MW) of 22194.89 (Table 2). Stability index indicated that the protein was stable and had a long half-life in mammals, yeast and *E. coli*. Blue and green peaks after 100 positions showed the richness of alpha helices and stand: a properties required to select the best epitopes for effective vaccines (Figure 2). The tertiary structure of the protein specified a single domain having 98% similarity with the protein (PDEB code: 2dem) from shrimp (Figure 4a). In general, a protein with >100 residues and >50 GDT can be concluded as a good indicator; so, overall uGDT value of 154 and GDT of 75 specified that the modelled protein covered most of the amino acids and large portions of the protein were flawless. The 3D structure also revealed only 44 positions in the protein as disordered, that means, only 19% of the proteins did not form stable tertiary structure.

B-cell and CTL Epitopes Prediction

The prediction of the B-cell epitopes in the target antigen based on Kolaskar and Tongaonkar, 1990 method on IEDB: one of the key steps in the epitope-based vaccine design. All of the 10 B-cell linear epitopes were identified in the VP26 envelope protein of WSSV, had different amino acids length, varying from 1 to 27. The conservancy level of these predicted peptides were

Table 2. Characteristics of secondary structure of the highest antigenic protein (AAO69663.1) of WSSV

Criteria	Score
Number of amino acids	204
Major amino acids	Ile 12.3%, Val 8.8%, Ser 8.3%, Thr 7.8%, Asn 7.8%, Ala 7.4%, Lys 7.4%
Molecular weight	22194.89
Total number of atom	3171
Formula	$C_{970}H_{1623}N_{267}O_{299}S_{12}$
Theoretical pI	9.35
Extinction coefficient	4470
Estimated half-life	30 hrs (mammal); >20 hrs (Yeast); >10 hrs (*E coli*)
Instability index	23.22 (indicates stable protein)
Aliphatic index	103.68
GRAVY	0.112
Alpha helix	59 (28.92%)
Extended strand	51 (25.00%)
Beta turn	16 (7.84%)
Random coil	78 (38.24%)

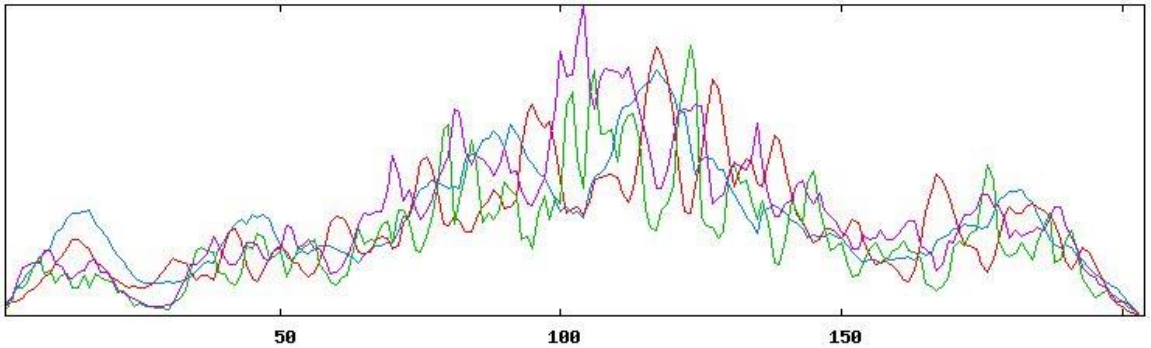

Figure 2. Secondary structure of the highest antigenic envelope protein AAO69663.1 (VP26) of WSSV. Here helices are indicated by blue, extended strands, beta turns and coils are specified by red, green and yellow, respectively.

ranging from 76%-95%. In these prediction, the epitope from position 99 to 115 showed highest antigenicity (Figure 3) with maximum value of 2.273 for position 106 and minimum value of 0.204 for position 99 while average value was 1.851. Among the 17 amino acids in the epitope, 13 had values of more than 1.0 and six had more than 2.0. Hydrophilicity of the predicted peptides showed that 15 out of 17 amino acids had positive values ranging from 5.714 to 2.871, recommended hydrophilic nature of the peptides. The cytotoxic T lymphocyte (CTL) epitopes are potential peptides, used as vaccines for the control of diseases of human and animals for prolong time. CTLPred was applied to predict the best epitopes based on MHC class I binding and protesomal c-terminal cleavage by directing artificial neural network (ANN). CTLPred identified top five epitopes, all having scores of more than 0.98. Among these five, only the epitope ^{98}TVTAPRTDP^{106} scored 1.000 and positioned on the top of five (Table 3). The epitope was rich in tryptophan (3T) and proline (2P): the two antigenic proteins those elicit immune responses. The sequences position of CTL epitope, ^{98}TVTAPRTDP^{106} had 89% similarity match with the highest antigenic B-cell linear epitope.

Modelling of the Peptide and Docking to MHC Molecules

The sticks presentation of ^{98}TVTAPRTDP^{106} peptide modelled by PEP-FOLD3 is shown in Figure 4b. CABS-Dock docking results showed a strong binding of current epitope to MHC class I (HLA_A*0201_1HLA) molecule, having all strong hydrogen bonds (<3°A). The visualisation tool Pymol showed that all threonine (3T) and valine formed stable hydrogen bonds with the MHC class I molecule (Figure 4c). The predicted epitope bound more strongly to the MHC class II molecule (HLA_B*3508_1ZHL) with 11 hydrogen bonds where eight of them formed stable hydrogen bonds (<3°A). Like epitope binding to MHC class I (HLA_A*0201_1HLA) molecule, here also threonine (3T) made strong and stable hydrogen bonds (Figure 4d).

Allergenicity of the Epitope

Non-allergenic peptides are fundamental requirement to be considered as efficient and safe vaccines. AllerTop results showed that all the five epitopes were non-allergen and thus safe from any other sensitive reactions.

Discussion

For the last two decades, the outbreak of WSSV has caused massive damage to the aquaculture industry. Several studies on the development of a subunit vaccine against WSSV could not produce satisfactory results, whereas some epitope-based vaccines designed for many pathogenic microbes produced outstanding results in preclinical studies. Therefore, using an immunoinformatic approach, the main aim of the present study was to identify the potential T-cell epitope(s) from the envelope and nucleocapsid proteins of WSSV that could be used as promising vaccines. We

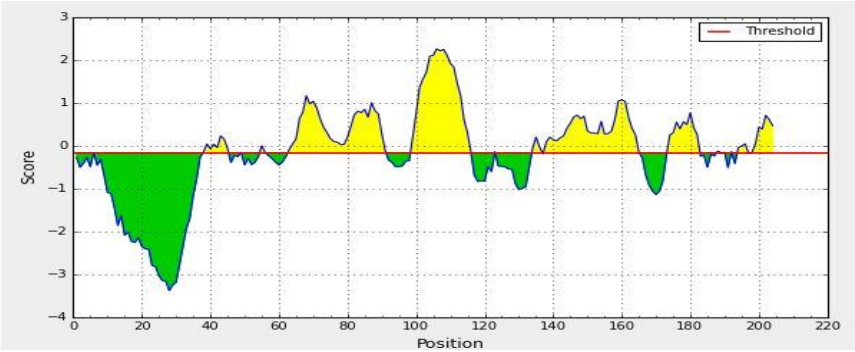

Figure 3. The threshold level (antigenicity) of top B-cell linear epitope. While most of the epitopes were unsuccessful in term of satisfactory threshold level of >1, the top B-cell epitope crossed the threshold level of 2.0. Threshold level was calculated in web based IDEB and NetCTL tool.

Table 3. Predicted five best epitopes in AAO60663.1 (VP26) protein with their position and antigenic score

Epitope	Start position	Score
TVTAPRTDP	98	1.000
DKDMKDVSA	84	0.990
MTLKILNTT	116	0.990
TMSNTYFSS	148	0.990
NTRVGRSVV	35	0.980

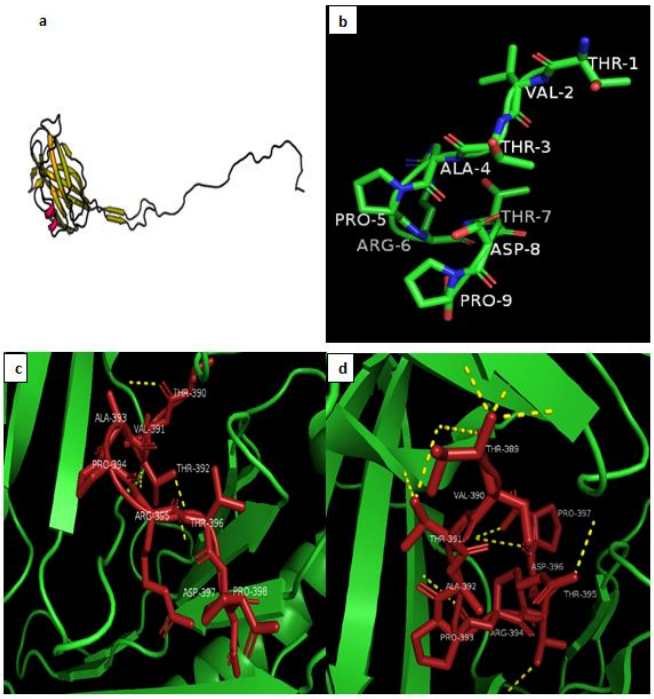

Figure 4. Predicted three-dimensional (3D) structure of AAO69663.1 (VP26) (**4a**); Modelled 3D structure of the ^{98}TVTAPRTDP^{106} epitope using PEP-FOLD3 (**4b**); Molecular docking of the predicted epitope to MHC class I (HLA A*0201; PDB code 1HLA) (**4c**) and class II (HLA B*3508; PDB code 1ZHL) molecule (**4d**). Docking was performed in CABS-dock and results were visualized using Pymol visualization tool. The epitope ^{98}TVTAPRTDP^{106} bound strongly to the binding groove of the MHC class I and class II molecule with strong clustering density (>55%) and covered many elements (>150)

analysed the sequence database of the four major proteins of WSSV deeply, over a period of 18 years for the precise prediction of the preeminent epitope(s). The sequence investigations revealed three common sources of origin: India, Vietnam, and Bangladesh, among nine different countries. Hence, we can conclude that WSSV has spread worldwide from South Asia. Secondly, only the VP24 and VP26 proteins contained significant bootstrap values (>70%), indicating that these taxa are well-supported by published data to be considered as vaccines. Thirdly, the evolutionary distances varied according to the geographical area and the year of collection while, the antigenicity values were the same, regardless of the country. The antigenicity scores of the ten VP19 envelope protein from India, South Korea, Iran, and Vietnam carried the same value of 0.5854. Similarly, only one out of the six VP26 proteins, AAO69663.1, had a different antigenicity score and so was for the other two. The AAO69663.1 protein from China was positioned in a completely separate cluster with a high bootstrap value (99%), which is required to avoid unexpected effects from unknown functions (Sheikh, Kahveci, Ranka, & Burleigh, 2017; Katsura, Kumar, & Nei, 2017). Theoretically, a vaccine should have two classes of antigenic epitopes in-order to elicit immunogenicity: the B-cell linear epitopes and the cytotoxic T lymphocyte (CTL) epitopes or T-cell epitopes (Sumo *et al*., 2016). In the present study, the B-cell epitope ^{99}VTAPRTDPAGTGAENSN^{115} showed 98% conservancy among all the WSSV data present at NCBI. Therefore, the greater conservancy is required to generate a broad spectrum protection against other microbes, a criterion that is always expected to be present in the selected epitopes (Shi *et al*., 2015). The prediction of T-cell antigens is one of the landmark inventions in designing epitope-based vaccines. The CTL epitopes play a central role in generating protective immune responses against diverse microbial infections (Esser *et al*., 2003). During the initial phase, the CTL calculation tool was solely used for the design of vaccines against human pathogens. Nowadays, however, immunoinformatics is also applied toward the CTL epitopes for designing vaccines against the pathogens of fish (Barbedo, 2014). Published data showed that the designed CTL epitopes are successful in neutralizing the pathogens of fish such as *Aeromonas* sp., *Edwardsiella tarda*, and *Flavobacterium columnare* (Mahendran *et al*., 2016; Marana *et al*., 2017). Among the top five epitopes predicted in the current study, only the epitope ^{98}TVTAPRTDP^{106}, having all the necessary properties of a probable candidate vaccine against WSSV scored a value of 1.000, in terms of antigenicity. Molecular docking, a tool that is extensively used to predict and visualize the stability of protein-peptide interactions, showed the interacting ability of the T-cell receptors to the binding trench of the MHC molecules (class I and class II) inside the cell (Anderson, 2003; Luo *et al*., 2016). In this study, the docking results showed that the current peptide was perfectly nestled into the binding grooves of the MHC class I and class II molecules by stable hydrogen bond. Earlier studies have claimed

the presence of four strong hydrogen bonds, ensuring a stable interaction between the MHC molecules and the peptide (Mahendran *et al*., 2016). Here, we found four and eight durable hydrogen bonds (<3Å) and hence, we can claim the present protein-peptide interactions as stable and rigid. The present epitope was found to be rich in threonine (3T) and proline (2P) residues within its 9-mer peptide. For several years, both the antigenic amino acids, threonine, and proline have been recommended for the preparation of vaccines (Kumar, Hussain, Sharma, Mehrotra, & Gissmann, 2015). Threonine is an essential, stable, polar amino acid that is used in the biosynthesis of proteins and is present on the surfaces of proteins. Threonine is a much stable amino acid, has antigenic determinants for generating immune responses, and together with serine contributes 67% of the hydrogen bonds, where 80% of interactions are with the DNA backbone (Luscombe, Laskowski, & Thornton, 2001). In addition, it has the potential to stabilize any dimer through hydrogen bonds and ensure a tighter association of small molecules into the transmembrane helices of proteins (Smith *et al*., 2002). *In-vivo* proline-rich antigens have also been proven to be effective vaccines for defending *Coccidiodes immitis* in mice (Peng, Orsborn, Orbach, & Galgiani, 1999). One experiment has revealed that the hydrogen bonds formed by proline are likely to deliver stability to the loops and capping regions of proteins (Deepak & Sankararamakrishnan, 2016). In the present study, the current epitope fulfilled all the criteria for the use as a future vaccine that could elicit protective neutralizing antibodies as well as cellular immune responses against WSSV. In examining the health and immunological conditions, further laboratory trials are needed in crayfishes before a commercial application.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

*Farhana Momtaz carried out sequence retrieval, alignment, phylogenetic evolutionary divergence analysis, and determination of antigenicity of major envelope proteins. *Md Javed Foysal designed the work and participated in the determination of primary and secondary structures, B cell and cytotoxic T cell epitopes prediction, modelling of the peptide, molecular docking of the peptide to proteins, also drafting of the manuscript; Md Mahbubur Rhaman and Ravi Fotedar were responsible for reviewing of the draft manuscript, analysis of data and critical points.
* Both authors contributed equally to this study.

Acknowledgement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Ali, M. T., Morshed, M. M., & Hassan, F. (2015). A computational approach for designing a universal epitope-based peptide vaccine against Nipah virus. *Interdisciplinary Sciences: Computational Life Sciences*, *7*(2), 177–185. https://dx.doi.org/10.1007/s12539-015-0023-0

Anderson, A. C. (2003). The process of structure-based drug design. *Chemistry and Biology*, *10*, 787–797. https://dx.doi.org/10.1016/j

Barbedo, J. G. A. (2014). Computer-aided disease diagnosis in aquaculture : Current state and perspectives for the future. *Revista Innover*, *1*(19), 19–32.

Bhasin, M., & Raghava, G. P. S. (2004). Prediction of CTL epitopes using QM, SVM and ANN techniques. *Vaccine*, *22*(23–24), 3195–3204. https://dx.doi.org/10.1016/j.vaccine.2004.02.005

Bir, J., Howlader, P., Ray, S., Sultana, S., Khalil, S. M. I., & Banu, G. R. (2017). A critical review on white spot syndrome virus (WSSV): A potential threat to shrimp farming in Bangladesh and some Asian countries. *International Journal of Microbiology and Mycology*, *6*(1), 39–48.

Ciemny, M., Kurcinski, M., Kozak, K., Kolinski, A., & Kmiecik, S. (2012). Highly flexible protein-peptide docking using CABS-Dock. *Methods in Molecular Biology*, *857*, 375–398. https://dx.doi.org/10.1007/978-1-61779-588-6_17

Dash, R., Das, R., Junaid, M., Akash, M. F. C., Islam, A., & Hosen, S. M. Z. (2017). In silico-based vaccine design against Ebola virus glycoprotein. *Advances and Applications in Bioinformatics and Chemistry*, *10*(1), 11–28. https://dx.doi.org/10.2147/AABC.S115859

Deepak, R. N. V. K., & Sankararamakrishnan, R. (2016). Unconventional N-H...N hydrogen bonds involving proline backbone nitrogen in protein structures. *Biophysical Journal*, *110*(9), 1967–1979. https://dx.doi.org/10.1016/j.bpj.2016.03.034

Dijkstra, J. M., Grimholt, U., Leong, J., Koop, B. F., & Hashimoto, K. (2013). Comprehensive analysis of MHC class II genes in teleost fish genomes reveals dispensability of the peptide-loading DM system in a large part of vertebrates. *BMC Evolutionary Biology*, *13*(1), 260. https://dx.doi.org/10.1186/1471-2148-13-260

Dimitrov, I., Bangov, I., Flower, D. R., & Doytchinova, I. (2014). AllerTOP v.2—a server for in silico prediction of allergens. *Journal of Molecular Modeling*, *20*(6), 2278. https://dx.doi.org/10.1007/s00894-014-2278-5

Doytchinova, I. A., & Flower, D. R. (2007). VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics*, *8*(1), 4. https://dx.doi.org/10.1186/1471-2105-8-4

Esser, M. T., Marchese, R. D., Kierstead, L. S., Tussey, L. G., Wang, F., Chirmule, N., & Washabaugh, M. W. (2003). Memory T cells and vaccines. *Vaccine*, *21*, 419–430.

Ganjoor, M. (2015). A short review on infectious viruses in cultural shrimps (Penaeidae family). *Fisheries and Aquaculture Journal*, *6*(3). https://dx.doi.org/10.4172/2150-3508.1000136

Ghosh, A., Chattopadhyay, S., Chawla-Sarkar, M., Nandy, P., & Nandy, A. (2012). In silico study of rotavirus VP7 surface

accessible conserved regions for antiviral drug/vaccine design. *PLoS ONE*, *7*(7). https://dx.doi.org/10.1371/journal.pone.0040749

Grimholt, U. (2016). MHC and evolution in Teleosts. *Biology*, *5*(1), 6. https://dx.doi.org/10.3390/biology5010006

Ha, Yu-Mi, Gong, S., Nguyen, T., Ra, C. H., Kim, K., & Nam, Y. (2008). Vaccination of shrimp (Penaeus chinensis) against white spot syndrome virus (WSSV). *Journal of Microbiology and Biotechnology*, *18*(5), 964–967.

Haug, B. E., Camilio, K. A., Eliassen, L. T., Stensen, W., Svendsen, J. S., Berg, K., … Rekdal. (2016). Discovery of a 9-mer cationic peptide (LTX-315) as a potential first in class oncolytic peptide. *Journal of Medicinal Chemistry*, *59*(7), 2918–2927. https://dx.doi.org/10.1021/acs.jmedchem.5b02025

Källberg, M., Wang, H., Wang, S., Peng, J., Wang, Z., Lu, H., & Xu, J. (2012). Template-based protein structure modeling using the RaptorX web server. *Nature Protocols*, *7*(8), 1511–1522. https://dx.doi.org/10.1038/nprot.2012.085

Karunasagar, I., & Ababouch, L. (2012). Shrimp viral diseases, import risk assessment and international trade. *Indian Journal of Virology*, *23*(2), 141–148. https://dx.doi.org/10.1007/s13337-012-0081-4

Katsura, Y., Jr, C. E. S., Kumar, S., & Nei, M. (2017). The reliability and stability of an inferred phylogenetic tree from empirical data. *Molecular Biology and Evolution*, *34*(3), 718–723. https://dx.doi.org/10.1093/molbev/msw272

Kori, P., Sajjan, S. S., & Madagi, S. B. (2015). In silico prediction of epitopes for Chikungunya viral strains. *Journal of Pharmaceutical Investigation*, *45*(6), 579–591. https://dx.doi.org/10.1007/s40005-015-0205-0

Kua, B. C., & Rashid, N. M. (2012). Pathology associated with white spot virus (WSV) infection in wild broodstock of tiger prawns (Penaeus monodon). *Tropical Life Sciences Research*, *23*(1), 87–92.

Kumar, A., Hussain, S., Sharma, G., Mehrotra, R., & Gissmann, L. (2015). Identification and validation of immunogenic potential of India specific HPV-16 variant constructs : In-silico & in-vivo insight to vaccine development. *Scientific Reports*, (March), 1–9. https://dx.doi.org/10.1038/srep15751

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, *33*(7), 1870–1874. https://dx.doi.org/10.1093/molbev/msw054

Kurcinski, M., Jamroz, M., Blaszczyk, M., Kolinski, A., & Kmiecik, S. (2015). CABS-dock web server for the flexible docking of peptides to proteins without prior knowledge of the binding site. *Nucleic Acids Research*, *43*(W1), W419–W424. https://dx.doi.org/10.1093/nar/gkv456

Lamiable, A., Thévenet, P., Rey, J., Vavrusa, M., Derreumaux, P., & Tufféry, P. (2016). PEP-FOLD3: Faster de novo structure prediction for linear peptides in solution and in complex. *Nucleic Acids Research*, *44*(W1), W449–W454. https://dx.doi.org/10.1093/nar/gkw329

Li, Z., Li, F., Han, Y., Xu, L., & Yang, F. (2015). VP24 is a chitin-binding protein involved in white spot syndrome virus infection. *Journal of Virology*, *90*(October), JVI.02357-15. https://dx.doi.org/10.1128/JVI.02357-15

Luo, H., Du, T., Zhou, P., Yang, L., Mei, H., Ng, H., & Rock, L. (2016). Molecular docking to identify associations between drugs and class I human leukocyte antigens for predicting idiosyncratic drug reactions. *Combined Chemistry High Throughput Screening*, *18*(3), 1–18.

Luscombe, N. M., Laskowski, R. A., & Thornton, J. M. (2001). Amino acid – base interactions : A three-dimensional analysis of protein – DNA interactions at an atomic level. *Nucleic Acids Research*, *29*(13), 2860–2874.

Madonia, A., Melchiorri, C., Bonamano, S., Marcelli, M., Bulfon, C., Castiglione, F., & Romano, N. (2017). Computational modeling of immune system of the fish for a more effective vaccination in aquaculture. *Bioinformatics*, *33*(19), 3065–3071. https://dx.doi.org/10.1093/bioinformatics/xxxxx

Mahendran, R., Jeyabaskar, S., Sitharaman, G., Michael, R. D., & Paul, A. V. (2016). Computer-aided vaccine designing approach against fish pathogens Edwardsiella tarda and Flavobacterium columnare using bioinformatics softwares. *Drug Design, Development and Therapy*, *10*, 1703–1714. https://dx.doi.org/10.2147/DDDT.S95691

Marana, M. H., Von Gersdorff Jørgensen, L., Skov, J., Chettri, J. K., Mattsson, A. H., Dalsgaard, I., & Buchmann, K. (2017). Subunit vaccine candidates against Aeromonas salmonicida in rainbow trout Oncorhynchus mykiss. *PLoS ONE*, *12*(2), 1–15. https://dx.doi.org/10.1371/journal.pone.0171944

Mirza, M. U., Rafique, S., Ali, A., Munir, M., Ikram, N., Manan, A., & Idrees, M. (2017). Towards peptide vaccines against Zika virus: Immunoinformatics combined with molecular dynamics simulations to predict antigenic epitopes of Zika viral proteins. *Scientific Reports*, *7*(October), 44633. https://dx.doi.org/10.1038/srep44633

Muñoz-Medina, J. E., Sánchez-Vallejo, C. J., Méndez-Tenorio, A., Monroy-Muñoz, I. E., Angeles-Martínez, J., Santos Coy-Arechavaleta, A., & Díaz-Quiñonez, J. A. (2015). In silico identification of highly conserved epitopes of influenza A H1N1, H2N2, H3N2, and H5N1 with diagnostic and vaccination potential. *BioMed Research International*, *2015*, ID813047. https://dx.doi.org/10.1155/2015/813047

Patel, S. M., Koringa, P. G., Reddy, B. B., Nathani, N. M., & Joshi, C. G. (2015). In silico analysis of consequences of non-synonymous SNPs of Slc11a2 gene in Indian bovines. *Genomics Data*, *5*, 72–79. https://dx.doi.org/10.1016/j.gdata.2015.05.015

Peng, T., Orsborn, K. I., Orbach, M. J., & Galgiani, J. N. (1999). Proline-rich vaccine candidate antigen of Coccidioides immitis: Conservation among isolates and differential expression with spherule maturation. *The Journal of Infectious Diseases*, *179*(November), 518–521.

Pereira, U. P., Soares, S. C., Blom, J., Leal, C. A. G., Ramos, R. T. J., Guimarães, L. C., & Figueiredo, H. C. P. (2013). In silico prediction of conserved vaccine targets in Streptococcus agalactiae strains isolated from fish, cattle, and human samples. *Genetics and Molecular Research*, *12*(3), 2902–2912. https://dx.doi.org/10.4238/2013.August.12.6

Sangamaheswaran, P., & Jeyaseelan, M. J. P. (2001). White spot viral disease in penaeid shrimp – A review. *NAGA: The ICLARM Quarterly*, *24*(December), 16–22.

Sheikh, S. I., Kahveci, T., Ranka, S., & Burleigh, J. G. (2017). Stability analysis of phylogenetic trees. *Bioinformatics*, *29*(2), 166–174. https://dx.doi.org/10.1093/bioinformatics/bts657

Shi, J., Zhang, J., Li, S., Sun, J., Teng, Y., Wu, M., & Hu, Y. (2015). Epitope-based vaccine target screening against highly pathogenic MERS-CoV: An in silico approach applied to emerging infectious diseases. *PLoS ONE*, *10*(12), 1–16.

https://dx.doi.org/10.1371/journal.pone.0144475

Smith, S. O., Eilers, M., Song, D., Crocker, E., Ying, W., & Groesbeek, M. (2002). Implications of threonine hydrogen bonding in the glycophorin a transmembrane helix dimer. *Biophysical Journal*, *82*(5), 2476–2486. https://dx.doi.org/10.1016/S0006-3495(02)75590-2

Sumo, U., Tambunan, F., Ruth, F., Sipahutar, P., Parikesit, A. A., & Kerami, D. (2016). Vaccine design for H5N1 based on B- and T-cell epitope predictions. *Bioinformatics and Biology Insights*, *10*, 27–35. https://dx.doi.org/10.4137/BBI.S38378

Syed Musthaq, S. K., & Kwang, J. (2015). Evolution of specific immunity in shrimp - a vaccination perspective against white spot syndrome virus. *Developmental and Comparative Immunology*, *48*(2), 342–353. https://dx.doi.org/10.1016/j.dci.2014.07.016

Tan, Y., & Shi, Z. (2008). Proteomic analyses of the shrimp white spot syndrome virus. *Virologica Sinica*, *23*(3), 157–166. https://dx.doi.org/10.1007/s12250-008-2924-0

Tang, X., & Hew, C. L. (2007). Expression, purification and crystallization of two major envelope proteins from white spot syndrome virus. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications*, *63*(7), 624–626. https://dx.doi.org/10.1107/S1744309107029351

Tsai, J., Wang, H., Leu, J., Wang, H., Zhuang, Y., Walker, P. J., & Lo, C. (2006). Identification of the nucleocapsid , tegument, and envelope proteins of the shrimp white spot syndrome virus virion. *Journal of Virology*, *80*(6), 3021–3029. https://dx.doi.org/10.1128/JVI.80.6.3021

Tuyen, N. X., Verreth, J., Vlak, J. M., & de Jong, M. C. M. (2014). Horizontal transmission dynamics of white spot syndrome virus by cohabitation trials in juvenile Penaeus monodon and P. vannamei. *Preventive Veterinary Medicine*, *117*(1), 286–294. https://dx.doi.org/10.1016/j.prevetmed.2014.08.007

van Hulten, M. C., Goldbach, R. W., & Vlak, J. M. (2000). Three functionally diverged major structural proteins of white spot syndrome virus evolved by gene duplication. *The Journal of General Virology*, *81*, 2525–2529.

Vaseeharan, B., Jayakumar, R., & Ramasamy, P. (2003). PCR-based detection of white spot syndrome virus in cultured and captured crustaceans in India. *Letters in Applied Microbiology*, *37*(6), 443–447. https://dx.doi.org/10.1046/j.1472-765X.2003.01428.x

Verbruggen, B., Bickley, L. K., van Aerle, R., Bateman, K. S., Stentiford, G. D., Santos, E. M., & Tyler, C. R. (2016). Molecular mechanisms of white spot syndrome virus infection and perspectives on treatments. *Viruses*, *8*(1), 1–29. https://dx.doi.org/10.3390/v8010023

Wan, Q., Xu, L., & Yang, F. (2008). VP26 of white spot syndrome virus functions as a linker protein between the envelope and nucleocapsid of virions by binding with VP51. *Journal of Virology*, *82*(24), 12598–601. https://dx.doi.org/10.1128/JVI.01732-08

Xu, T., Sun, Y., Shi, G., Cheng, Y., & Wang, R. (2011). Characterization of the major histocompatibility complex class II genes in miiuy croaker. *PLoS ONE*, *6*(8). https://dx.doi.org/10.1371/journal.pone.0023823

Yogeeswaran, A., Velmurugan, S., Punitha, S. M. J., Babu, M. M., Selvaraj, T., Kumaran, T., & Citarasu, T. (2012). Protection of Penaeus monodon against white spot syndrome virus by inactivated vaccine with herbal immunostimulants. *Fish and Shellfish Immunology*, *32*(6), 1058–1067. https://dx.doi.org/10.1016/j.fsi.2012.02.029

Yuan, S., Chan, H. C. S., & Hu, Z. (2017). Using PyMOL as a platform for computational drug design. *WIREs Computational Molecular Science*, *7*(2), 1–10. https://dx.doi.org/10.1002/wcms.1298

Zhou, Q., Xu, L., Li, H., Qi, Y. P., & Yang, F. (2009). Four major envelope proteins of white spot syndrome virus bind to form a complex. *Journal of Virology*, *83*(9), 4709–4712. https://dx.doi.org/10.1128/JVI.02360-08

Zhu, F., Du, H., Miao, Z. G., Quan, H. Z., & Xu, Z. R. (2009). Protection of Procambarus clarkii against white spot syndrome virus using inactivated WSSV. *Fish & Shellfish Immunology*, *26*(5), 685–90. https://dx.doi.org/10.1016/j.fsi.2009.02.022