

## RESEARCH ARTICLE

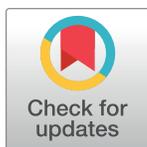
# Immunoinformatics and molecular dynamics approaches: Next generation vaccine design against West Nile virus

Md Tahsin Khan<sup>1</sup> <sup>✉</sup>, Rahatul Islam<sup>1</sup> <sup>✉</sup>, Tarhima Jahan Jerin<sup>2</sup> <sup>✉</sup>, Araf Mahmud<sup>1</sup>, Sahara Khatun<sup>1</sup> <sup>✉</sup>, Ahasanul Kobir<sup>3</sup>, Md Nahidul Islam<sup>4</sup>, Arzuba Akter<sup>5\*</sup>, Shakhinur Islam Mondal<sup>1\*</sup> <sup>✉</sup>

**1** Department of Genetic Engineering and Biotechnology, Shahjalal University of Science and Technology, Sylhet, Bangladesh, **2** Department of Biotechnology and Genetic Engineering, Mawlana Bhashani Science and Technology University, Tangail, Bangladesh, **3** UPMC Hillman Cancer Center, University of Pittsburgh, Pittsburgh, Pennsylvania, United States of America, **4** Department of Biochemistry, School of Natural Sciences, National University of Ireland Galway, Galway, Ireland, **5** Department of Biochemistry and Molecular Biology, Shahjalal University of Science and Technology, Sylhet, Bangladesh

 These authors contributed equally to this work.

\* [shakhin200-gen@sust.edu](mailto:shakhin200-gen@sust.edu) (SIM); [arzuba-bmb@sust.edu](mailto:arzuba-bmb@sust.edu) (AA)



## OPEN ACCESS

**Citation:** Khan MT, Islam R, Jerin TJ, Mahmud A, Khatun S, Kobir A, et al. (2021) Immunoinformatics and molecular dynamics approaches: Next generation vaccine design against West Nile virus. PLoS ONE 16(6): e0253393. <https://doi.org/10.1371/journal.pone.0253393>

**Editor:** Jianming Tang, University of Alabama at Birmingham, UNITED STATES

**Received:** December 17, 2020

**Accepted:** June 3, 2021

**Published:** June 17, 2021

**Copyright:** © 2021 Khan et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the manuscript and its [Supporting information](#) files.

**Funding:** The author(s) received no specific funding for this work.

**Competing interests:** The authors have declared that no competing interests exist.

## Abstract

West Nile Virus (WNV) is a life threatening flavivirus that causes significant morbidity and mortality worldwide. No preventive therapeutics including vaccines against WNV are available for human use. In this study, immunoinformatics approach was performed to design a multi epitope-based subunit vaccine against this deadly pathogen. Human (HLA) and Mice (H-2) allele specific potential T-cell and B-cell epitopes were shortlisted through a stringent procedure. Molecular docking showed selected epitopes that have stronger binding affinity with human TLR-4. Molecular dynamics simulation confirmed the stable nature of the docked complex. Furthermore, *in silico* cloning analysis ensures efficient expression of desired gene in the microbial system. Interestingly, previous studies showed that two of our selected epitopes have strong immune response against WNV. Therefore, selected epitopes could be strong vaccine candidates to prevent WNV infections in human. However, further *in vitro* and *in vivo* investigations could be strengthening the validation of the vaccine candidate against WNV.

## Introduction

West Nile virus (WNV) is a single-stranded RNA virus (family Flaviviridae; genus *Flavivirus*) [1,2] and its infection can cause physical, neurologic, and cognitive disabilities [3]. Generally, WNV infection is asymptomatic. A common febrile illness occurs in ≈20% of the population and a small portion (<1%) develop severe West Nile neuroinvasive disease (WNND); characterized by encephalitis, meningitis, and/or acute flaccid paralysis [4,5]. The virus was initially identified in Uganda but now prevalent almost in all continents except Antarctica [6,7]. Since it was first detected in New York City in 1999, it causes thousands of cases each year in the United States and around 7 million total human infections were reported in the continental US [1,8]. Between 1999 and 2019, Centers for Disease Control and Prevention (CDC) has

reported more than 48,000 affected cases and over 2,300 deaths in US due to WNV infection [9]. WNV infection, mainly caused by WNV lineage 1 strains, has been an endemic in Eastern Europe such as Romania and Russia for over 50 years. Between 2004 to 2010, mammalian infections by WNV lineage 2 strains were reported in Hungary, Austria, and Russia. Other countries such as Greece, Macedonia, Romania, Serbia, Turkey, and Ukraine reported WNV infections in mammal after 2010 [10]. Since last few decades, emergence of WNV infections had been also reported in Southeast Asia, especially in several districts of India such as Assam [11], Kerala [12], West Bengal [13], Tamil Nadu [14], Madhya Pradesh [15], Andhra Pradesh [16], and Pune [17]. The economic burden of hospitalized WNV patients in US is \$56 million, and short- or long- term costs can even exceed \$700,000 per patient [18,19]. Considering the overall clinical and economic impacts caused by WNV, currently it takes more concern and considered one of the most vital zoonotic diseases to the US population.

The genome of WNV virus is approximately 11kb that encodes three structural proteins: Capsid protein (C), Pre-membrane/membrane protein (prM/M), and envelope protein (E). WNV virus also encodes seven non-structural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5); however their role yet to be understood [20]. The heterodimer forms of the prM and E protein are the two main immunogens of this virus [21]. The E protein is the major surface component for receptor binding and viral entry into host cells through membrane fusion [22]. Capsid protein is involved in viral RNA replication and packaging [20,23]. Based on the molecular information of WNV and research advancements, so far, a total 23 versatile vaccines preparations have been developed that include 4 DNA, 4 recombinant vectors, 11 inactivated/killed, and 4 live/attenuated (<http://www.violinet.org/>). Some of those vaccine candidates are under pre-clinical studies and a very few are under clinical trials; however, none of the WNV vaccines has yet been approved to be used in human host. Currently, only the YFV-based vaccine (ChimeriVax-WN02) has been evaluated in Phase II trial [24]. Other human vaccines are in phase I clinical trial: WN/DEN4D30 (recombinant vaccine), VRC-WNV DNA017-00-VP (DNA vaccine), and WN-80E (Recombinant subunit vaccine) [25–27]. None of the studies yet developed a potential vaccine where all the structural and non-structural viral proteins can be considered as antigens. In fact, only one or two proteins of the virus, mostly PrM/E, were used as antigens for formulating these vaccines. Also, in case of live attenuated virus (LAV) strategy, there is associated risk of reversion to the original pathogenic form of the virus.

The growing advances in the field of bioinformatics open a new era to design next generation vaccines through the knowledge of immunoinformatics; the detailed protocol and immune responses have been widely reviewed in published literatures [28–31]. In recent years, reverse vaccinology approach has been utilized to design epitope based multi-subunit vaccines against different pathogens and cancers [32–41]. Sufficient cellular and humoral immune responses triggered by multi-epitope based vaccine against various pathogens including Epstein–Barr virus, Crimean-Congo Hemorrhagic Fever Virus and Chlamydia trachomatis [42–45]. These vaccines works successfully both *in vitro* and *in vivo* murine model and thereby reducing time, costs and repetition of error trails of traditional vaccine development. In this study, by utilizing immunoinformatics and reverse vaccinology approaches, we design a candidate multi-epitope chimeric vaccine against the WNV virus that is likely to be safe and immunogenic against the WNV infection.

## Methods

### Retrieval of protein sequences and physicochemical properties analysis

In this study, we adopted computational methods to predict the effective vaccine candidates against WNV Virus where UniProt ID- Q9Q6P4 was considered as the reference strain. All

the proteins of this strain including structural (Capsid protein, prM, small envelope protein M and envelope protein) and non-structural (NS1, NS2B, NS2A, NS4A, NS3, NS4B, and NS5) viral proteins were retrieved in FASTA format from UniProt (Universal Protein Resource) database (<http://www.uniprot.org/uniprot>). Further ProtParam tool at ExPASy server (<http://expasy.org/cgi-bin/protparam>) was used to characterize the functional physicochemical parameters of all the proteins of reference strain [46].

### Prediction of linear B cell epitope

Linear B cell epitopes were predicted by combining several B cell epitope prediction server including ABCpred (<http://crdd.osdd.net/raghava/abcpred/>) [47], Bepipred (<http://tools.iedb.org/bcell/>) [48], Emini Surface Accessibility Prediction tool of IEDB (Immune Epitope Database) (<http://tools.iedb.org/bcell/>) [49], Karplus and Schulz flexibility tool (<http://tools.iedb.org/bcell/>) [50] and Parker Hydrophilicity Prediction method of IEDB (<http://tools.iedb.org/bcell/>) [51]. Combination of all methods increases the accuracy of B cell epitope prediction to a greater extent.

ABCpred is a recurrent neural network (ANN) algorithm for predicting B-cell epitopes (random peptides) of maximum length of 20 residues. The server was trained on 700 B-cell epitopes and 700 non B-cell epitopes for predicting B-cell lymphomas. The predicted accuracy achieved by this server is 65.93% using recurrent neural network. We set the default threshold of 0.51 in ABCpred server for BCL prediction. The BepiPred-2.0 server, a Random Forest algorithm predicts B-cell epitopes by comparing a set of trained epitopes and non-epitope from crystal structures. Epitopes with threshold value  $>0.5$  are considered and colored in yellow on the graph. Emini Surface Accessibility tool calculates the surface accessible epitopes by using a formula  $S_n = (n+4+i) \cdot (0.37)^{-6}$  where  $n$  is the value of fractional surface probability,  $S_n$  refers the surface probability and values of  $i$  varies from 1 to 6. Peptide sequence with  $S_n$  greater than 1.0 indicates the probability of being as surface accessible epitopes. The output of Karplus and Schulz flexibility tool provides seven amino acids window lengths of epitope based on three scales for flexibility calculation. Parker Hydrophilicity Prediction tool predicts the hydrophilicity of peptides by calculating the retention times of peptides during high-performance liquid chromatography (HPLC) on a reversed-phase column. Hydrophilic epitope regions were predicted by a window of seven residues score. Again, predicted B cell epitopes were also screened through VaxiJen, AllerTOP, ToxinPred and TMHMM webserver for final candidate selection.

### Prediction of CD8+ T cell (CTL) epitope

T cell epitopes were identified by using a non-linear artificial neural networks-based server, NetCTL 1.2 server (<http://www.cbs.dtu.dk/services/NetCTL/>) [52] where HLA class I alleles are sub-grouped into 12 super-families (A1, A2, A3, A24, A26, B7, B8, B27, B39, B44, B58, B62). All of the amino acid sequences of all WNV viral proteins were screened against each of the HLA class I superfamily (a total of  $8 \times 12 = 96$  queries). The threshold values used for WNV CTL prediction was 0.75. All of the epitopes beyond the threshold value were then selected for further analysis.

VaxiJen server (<http://www.ddg-pharmfac.net/vaxijen/>) was used to analyze the antigenicity of the proteins [53]. Epitopes that cannot meet the threshold level of the VaxiJen server, which is 0.4, were discarded. AllerTOP was also employed to determine the allergenicity of the selected T-cell epitopes [54]. Additionally, ToxinPred webserver was used to identify the toxicity of the epitopes [55]. Immunogenicity of the predicted epitopes was examined by IEDB MHC I immunogenicity prediction tool (<http://tools.iedb.org/immunogenicity/>) [56]. Finally transmembrane topology of the predicted epitopes was determined by using the TMHMM

server (<http://www.cbs.dtu.dk/services/TMHMM/>) [57]. After analyzing the entire parameters NetMHC 4.0 server (<http://www.cbs.dtu.dk/services/NetMHC/>) was used for cross checking with the selected epitopes for mouse allele (H-2) binding [58,59]. Epitopes showing positive results were further cross-checked with IEDB class I immunogenicity server for Human HLA class-I binding. Finally, most HLA covering epitopes were selected as final epitopes and subjected to population coverage analysis.

### Prediction of CD4+ T cell (HTL) epitope

IEDB MHC II binding database (<http://tools.iedb.org/mhcii/>) [60] was used for prediction of Mouse H-2-I allele binding epitope from all the 10 proteins of WNV. IEDB recommended method was selected as the prediction method and a window length of 15 mer was selected for CD4+ T cell (HTL) epitope prediction at IEDB MHC II binding interface.

Further these epitopes were screened through VaxiJen, AllerTOP, ToxinPred and TMHMM webserver. After cross checking with these servers final selected epitopes were then screened through IFNepitope webserver (<http://crdd.osdd.net/raghava/ifnepitope/>) [61]. This server is used for the prediction of IFN- $\gamma$  inducing epitope in a set of peptide library. The amino acid sequences of all the selected epitopes from previous screening were submitted as query for IFN- $\gamma$  inducing peptide prediction. Finally, epitopes showing positive results were further cross-checked with IEDB class II immunogenicity server for Human HLA class-II binding. As like as CTL screening, most HLA class-II covering epitopes were selected as final HTL epitopes and subjected to population coverage analysis.

### Conservancy analysis

In case of *in silico* vaccinology approach, conservancy analysis is used to determine the degree of epitope distribution in the homologous protein set. We used the epitope conservancy analysis tool (<http://tools.iedb.org/conservancy/>) at the IEDB (Immune Epitope Database) for the analysis of conservancy pattern of the desired epitopes. The sequence identity threshold was set at 100%, and a set of 100 homologous sequences of the selected 10 antigenic protein were retrieved from the NCBI (National Center for Biotechnology Information) database with BLASTp tool against the WNV Virus, for using as homologous proteins set in the conservancy tool [62].

However, in South East Asia, Europe and USA closely related flaviviruses are circulating. These viruses does not protect each other; there can be a possibility of cross protective immunity as all these viruses shares the same family. Furthermore, as a rule in flaviviruses, virus neutralizing antibodies is necessary. That's why the conserved B cell epitopes of WNV virus were then again ran in the conservancy analysis tool to look for homology in the Dengue Virus (Type-1,2,3 and 4) and Japanese encephalitis viruses reference polyprotein to check if any of the B cell epitope is conserved and similar in those polyproteins or not. If our predicted epitopes are matched in these polyproteins, there may be a cross reaction possibility can be occurred.

### Population coverage analysis

For effective vaccination, a vaccine molecule must provide a broad-spectrum protection against different populations around the world. However extreme polymorphic behavior of MHC molecules (Near around 6000) causes different MHC derived pool/frequencies in individuals of different ethnicities/country. Thus, selecting multiple peptides with different HLA binding capacities can increase coverage of population around the world. To address this issue, IEDB population coverage tool (<http://tools.iedb.org/population/>) [63] was utilized to

calculate the individual fractions response to a given set of epitopes based on HLA genotypic frequencies. However, the program provides three different calculation options for population coverage analysis: (1) class I separate, (2) class II separate, and (3) class I and class II combined. We used all of these options for predicting the population coverage of the proposed CTL and HTL epitopes found from preceding analysis.

To do that, all parameters in IEDB population coverage interface kept at default (1), the projected population coverage (2), the average number of HLA combinations recognized by the populations (3), and minimum number of HLA combinations recognized by 90% of the population ( $PC_{90}$ ) were calculated. [63]. Additionally, the corresponding MHC allele of the CTL and HTL were derived from IEDB MHC-I and MHC II prediction server.

### Vaccine construction

For the construction of the final vaccine protein, predicted B cell epitopes and T cell epitopes were joined by suitable fusion protein linkers. Then Universal Pan HLA DR sequence (PADRE) sequence was attached to overcome the problems caused by highly polymorphic HLA class-2 alleles in the final vaccine construct [64]. C-terminal invasin sequence of *Yersinia* was also added at the C-terminal end of the construct. Finally, an Adjuvant peptide was also incorporated in the final vaccine protein. Adjuvants can interact with toll like receptors (TLRs) to stimulate robust immune-reaction [65]. Beta-defensin-3 was used as a toll like receptor agonist in the final vaccine construct. The linkers used in the final vaccine construct were EAAAK, GGGG, GPGP, KK and EGGE. Each of the different linkers has different role. However the EAAAK linker can act as rigid linkers in the final vaccine construct whereas GGGG, GPGP, KK and EGGE provide flexibility to fused the vaccine candidates consequently [66]. Thus utilized linkers can ensure effective separation of individual epitopes [67].

### Prediction of physicochemical and immunogenic properties

ProtParam, a tool available on ExPASy server (<http://expasy.org/cgi-bin/protpraram>) [46] was used to characterize the functional physicochemical parameters of our vaccine constructs. Various physicochemical properties of the vaccine proteins were resolved including isoelectric pH, aliphatic index, molecular weight, GRAVY values, instability index, hydrophobicity, and estimated half-life. Physicochemical parameters are analyzed by calculating the pKa values of each amino acid present in the protein. Further Protein-sol server was used to calculate the solubility of the two vaccine proteins in terms of surface distribution of charge, hydrophobicity and stability [68]. This server can detect those sequence features which has the strongest impact on protein solubility. The experimental dataset (PopAvrSol) in Protein-sol server had a population average of 0.45. So, any solubility scores larger than 0.45 were expected to have a higher solubility than the average soluble *Escherichia coli* (*E.coli*) proteins from the experimental solubility dataset and vice versa [68]. Additionally AlgPred (<http://crdd.osdd.net/raghava/algpred/>) was used for prediction of allergenicity of the vaccine construct [69].

### Secondary structure prediction

Secondary structure is the general two-dimensional form of local segments of a protein or nucleic acid sequence. In proteins, secondary structure can be defined by the hydrogen bond patterns between backbone amino and carboxyl groups of proteins which is found in an atomic-resolution structure.

PSIPRED (<http://bioinf.cs.ucl.ac.uk/psipred/>) [70] an artificial neural network was utilized to predict the alpha helix, beta sheet and coil structure of the vaccine construct. PSIPRED used PSI-BLAST (Position-Specific Iterated-BLAST) for secondary structure prediction. PSIPRED

predicts the secondary structure of protein by two feed forward neural network processes, a) First network and b) Second network. The first network is utilized for initial prediction where the second network refines the structure obtained from the first prediction [70].

### Tertiary structure prediction, refinement, and validation

The tertiary structure of multi-epitope vaccine subunit vaccine was predicted by using Robetta (<http://rosetta.bakerlab.org/>), a protein structure prediction online server. Robetta is continually evaluated through CAMEO (continuous automated model evaluation) for tertiary structure prediction. Also, Robetta server utilizes custom sequence alignments for homology modeling, constraints, local fragments and so on. However, it uses the PDB100 template database and a co-evolution based model database (MDB) which is updated weekly.

On the other side, protein model structure produced by different protein structure prediction strategies mostly relies on similarity between the input and available template structure of PDB. So, the whole protein structure needs to be refined for improving the template-based protein model structure beyond the precision level. Therefore, ModRefiner server (<https://zhanglab.ccmb.med.umich.edu/ModRefiner/>) was used to enhance the accuracy of the predicted 3D modeled structure [71]. Furthermore, PROCHECK (<https://servicesn.mbi.ucla.edu/PROCHECK/>) [72] was used to analyze residual geometry of the refined vaccine construct and predict the improved stereochemical quality of the construct for validation.

### Discontinuous B-cell epitope prediction

Discontinuous or conformational epitopes are those that consist of multiple, distinct segments from the primary amino-acid sequence. Conformational epitopes residues are distantly placed in the sequence and brought together by physicochemical folding of the Protein's tertiary structure. [73]. Most of the acknowledged epitopes in a protein sequences are discontinuous [74]. So, analysis of discontinuous epitope in the final vaccine construct is prerequisite.

ElliPro (<http://tools.iedb.org/ellipro/>) [75] a discontinuous B cell epitope prediction server at IEDB was used for discontinuous/conformational B cell epitope prediction in the final vaccine construct. ElliPro scored each output epitope based on PI (Protrusion Index) value of epitope residues. The PI value of epitope residues is the criteria for discontinuous epitopes identification in a protein. Tertiary structure of the refined vaccine model was provided in the ElliPro server for conformational epitope prediction.

### Molecular docking of the vaccine candidate with immune receptor

Molecular docking is an *in silico* approach that can evaluate the binding correlation between a ligand and a receptor molecule [76]. We assessed PatchDock server (<https://bioinfo3d.cs.tau.ac.il/PatchDock/>) for analyzing protein-protein docking interaction to determine the binding affinity of our designed peptide vaccine with immune receptor TLR-3, 4, 7 and 8 (PDB ID: 1ZIW, 4G8A, 5gmf and 3W3J respectively) [77]. PDB files of vaccine protein and all the TLRs receptors were uploaded to the PatchDock servers for docking interaction. Later, FireDock output refinement of PatchDock server was used for refining the complexes.

### Molecular dynamics simulation

Molecular dynamics simulation is recognized as an effective method for molecular analysis of biological systems [9,78,79]. In previous studies it was used to check the stability of different protein complexes [80–82]. However we used the Linux-based GROMACS [83] software for molecular dynamics simulation to understand the structural properties and interaction

between TLR-4 and predicted vaccine protein at microscopic level. GROMACS can be performed using a variety of force fields, such as GROMOS, OPLSS, AMBER, and CHARMM [84]. In our study, the simulation was executed with GROMACS 2020.1 package with the force field as CHARMM36 [85]. The protein solvation was executed with SPC water model in a cubic box. About 83842 water molecules were added to the system and 25 CL molecules were added to neutralize the system. Meanwhile the system was equilibrated for 100ps.

The simulation of docked vaccine was performed to study interaction pattern of vaccine and changes in the complex after molecular docking with respective ligand. Additionally, the molecular dynamics simulation was carried out for 10ns long.

We analyzed the RMSD (Root Mean Square Deviation of the atomic position), RMSF (Root Mean Square Fluctuations) and Rg (Radius Gyration) of the vaccine-receptor complex. The RMSD, RMSF and Rg can be used to obtain information about our Biosystems. The RMSD includes the calculation concerning the reference structure (backbone) of the average atom location in a molecule. This method is used to calculate the average changes in atom displacement for evaluating the conformational shifts and Biosystems stability. The average deviation of a particle (for example, a protein residue) from the reference location (usually the average particle location) over time is calculated with root-mean-square fluctuation (RMSF). RMSF therefore analyzes the structural portions which fluctuate more (or less) from their mean structure. The gyration radius (Rg) represents the compactness of the structure. The lower degree of fluctuation with its simulation stability indicates the higher compactness and rigidity of the device.

### ***In Silico* development of adenovirus based vaccine**

To construct an adenovirus based vaccine, Kozak consensus sequence containing codon for methionine was added at the N-terminal end of vaccine protein. Codon optimization was done in GenScript server for *Homo sapiens* expression system. Four restriction sites (*BglII*, *EcoRV*, *PmeI* and *PacI*) were avoided from the optimized cDNA sequence. GC content, CAI values were again evaluated, and only highly efficient cDNA sequence was inserted in pAd-Track-CMV shuttle vector [86] in between *BglIII* and *EcoRV* restriction site under strong CMV promoter through SnapGene v5.1 [87] restriction cloning module. RNAfold server was also employed to analyze the translation efficiency and thermodynamic stability of expressed mRNA sequence [88].

## **Results**

### **Retrieval of protein sequences**

The amino acid sequences of structural and non-structural proteins of WNV virus (UniProt ID- Q9Q6P4) were retrieved from UniProt Database. The physiochemical properties (i.e. Molecular weight, Aliphatic Index, Theoretical pI, Grand average of hydropathicity and Instability index) of each protein were analyzed using ExPASy ProtParam tools (Table 1).

### **Linear B Cell (BCL) epitope prediction**

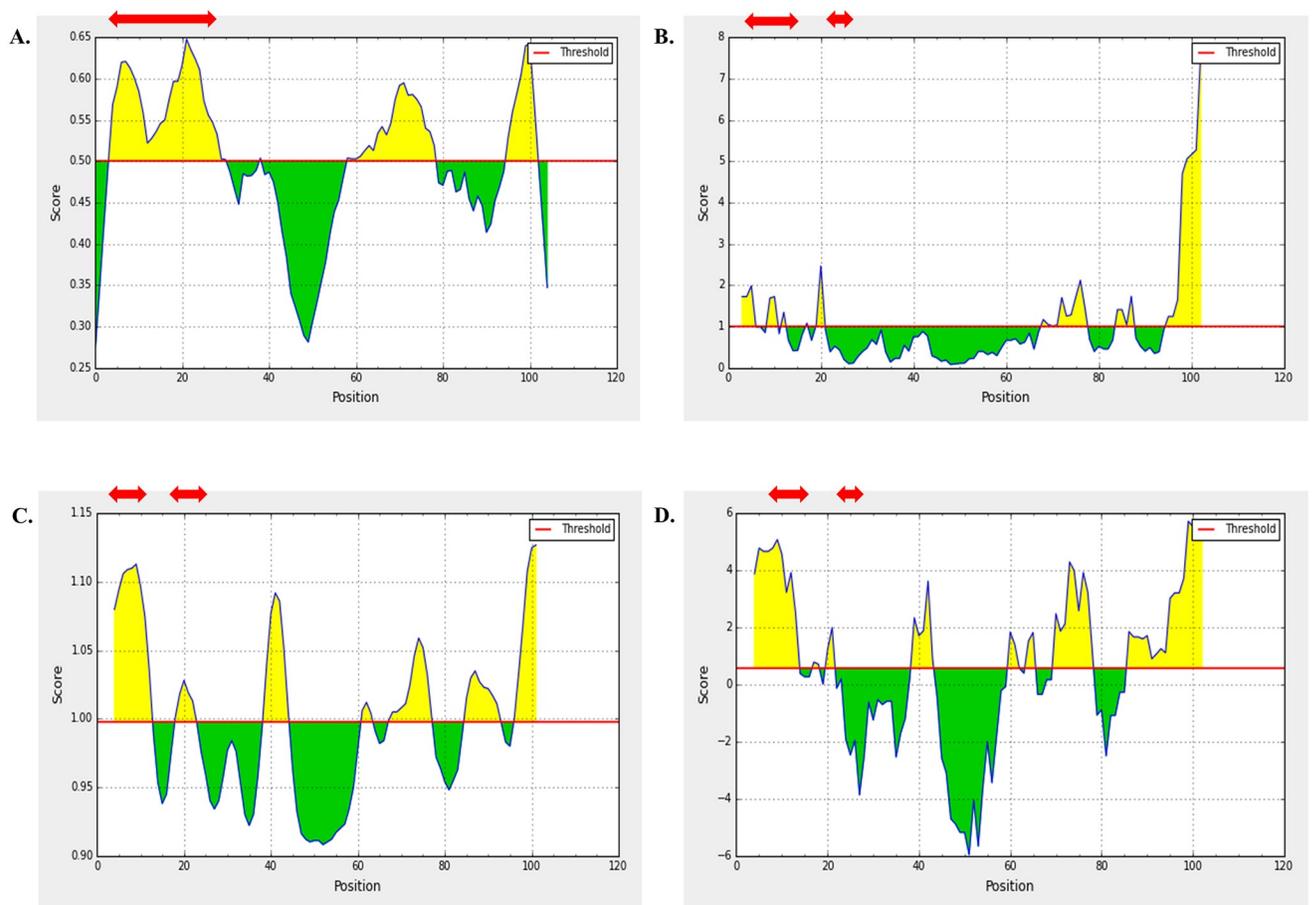
Linear B cell epitopes were screened only from the three structural proteins. Several B cell epitopes prediction server including ABCpred, Bepipred, and Emini Surface Accessibility Prediction tool of IEDB were used for identification of potential B cell epitopes from WNV structural proteins. Linear B cell epitopes of 16-mer predicted through ABCpred server were cross-checked with BepiPred 2.0 and Emini Surface Accessibility Prediction tool of IEDB server and overlapped epitopes were shortlisted (S1 File). Moreover, we analyzed the antigenicity, non-toxicity, and non-allergenicity of the overlapped peptides. Based on the best score of the

Table 1. Physicochemical parameters of structural and non-structural protein of West Nile virus.

Protein Name	Number of amino acids	Molecular weight (Da)	Theoretical pI	Instability index	Aliphatic index	Grand average of hydropathicity
Capsid C	105	11718.20	12.31	46.07	92.95	-0.209
Protein prM	167	18465.41	9.10	45.03	88.08	0.037
Envelope	501	53620.34	7.66	23.78	84.45	0.058
NS1	352	39749.80	5.73	45.87	73.10	-0.575
NS2A	231	25391.57	9.67	42.24	133.90	0.790
NS2B	131	14466.79	4.22	28.53	96.11	0.369
NS3	619	68995.37	7.32	31.42	76.56	-0.457
NS4A	126	13685.60	6.70	33.78	120.00	0.660
NS4B	255	27588.36	8.93	30.82	109.02	0.430
NS5	905	103683.06	8.63	36.25	72.87	-0.580

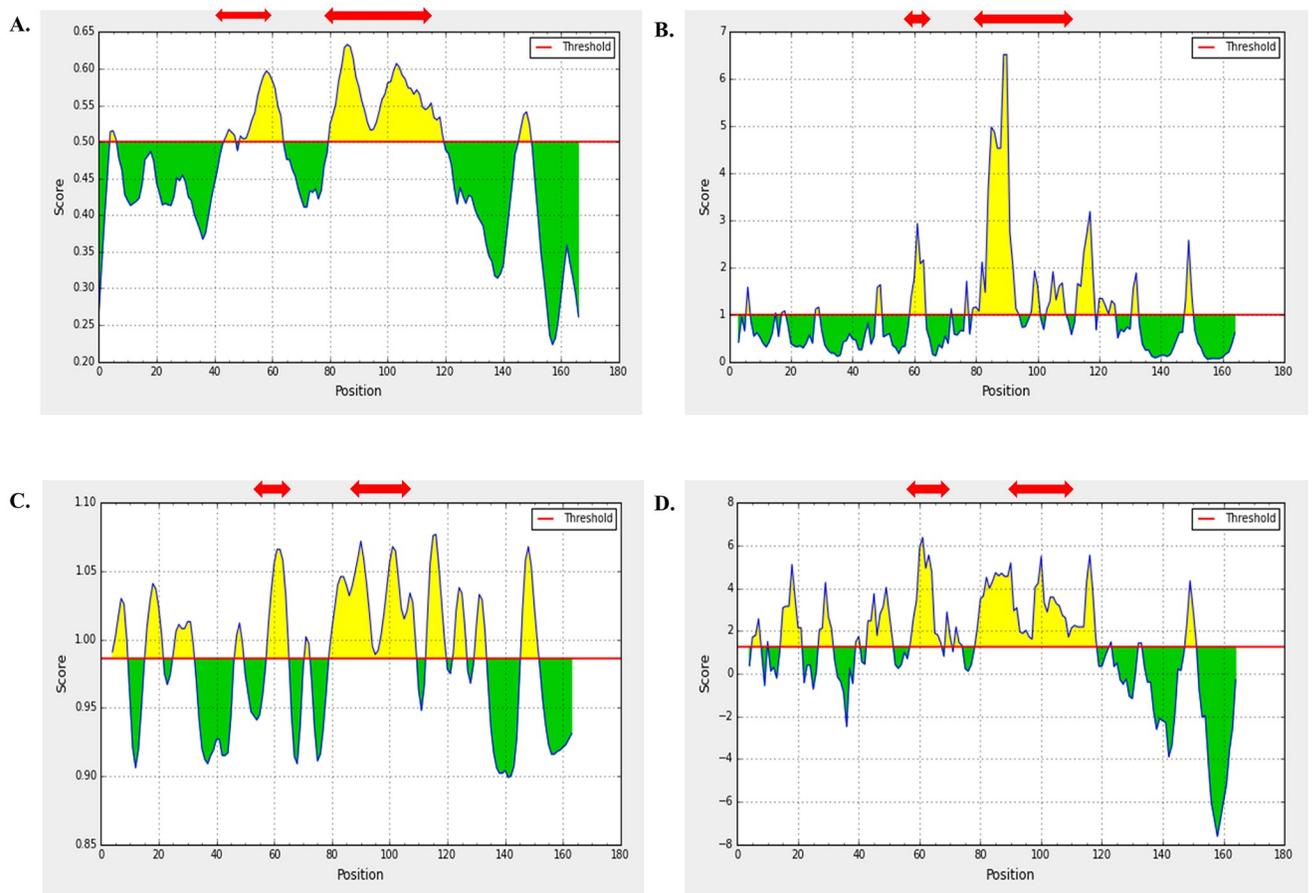
<https://doi.org/10.1371/journal.pone.0253393.t001>

analysis, two epitopes from each protein was selected as the final B cell epitopes. As these proteins are located in the outer surface of the virus and thus highly mutable, selection of two epitopes were considered. Finally, these 6 epitopes were further analyzed through Karplus & Schulz Flexibility and Parker Hydrophilicity Prediction tool at IEDB. All of the predicted peptide fragments were obtained to have satisfactory surface accessibility and flexibility (Figs 1–3).



**Fig 1. B cell Epitope prediction of Capsid C protein at IEDB linear epitope prediction tool.** Predicted peptide regions above the threshold values (yellow region) are shown in red arrow. Methods used for prediction are enlisted as: (A) BepiPred 2.0, (B) Emini Surface Accessibility tool, (C) Karplus & Schulz Flexibility tool, and (D) Parker Hydrophilicity tool.

<https://doi.org/10.1371/journal.pone.0253393.g001>



**Fig 2. B cell Epitope prediction of prM protein at IEDB linear epitope prediction tool.** Predicted peptide regions above the threshold values (yellow region) are shown in red arrow. Methods used for prediction are enlisted as: (A) BepiPred 2.0, (B) Emini Surface Accessibility tool, (C) Karplus & Schulz Flexibility tool, and (D) Parker Hydrophilicity tool.

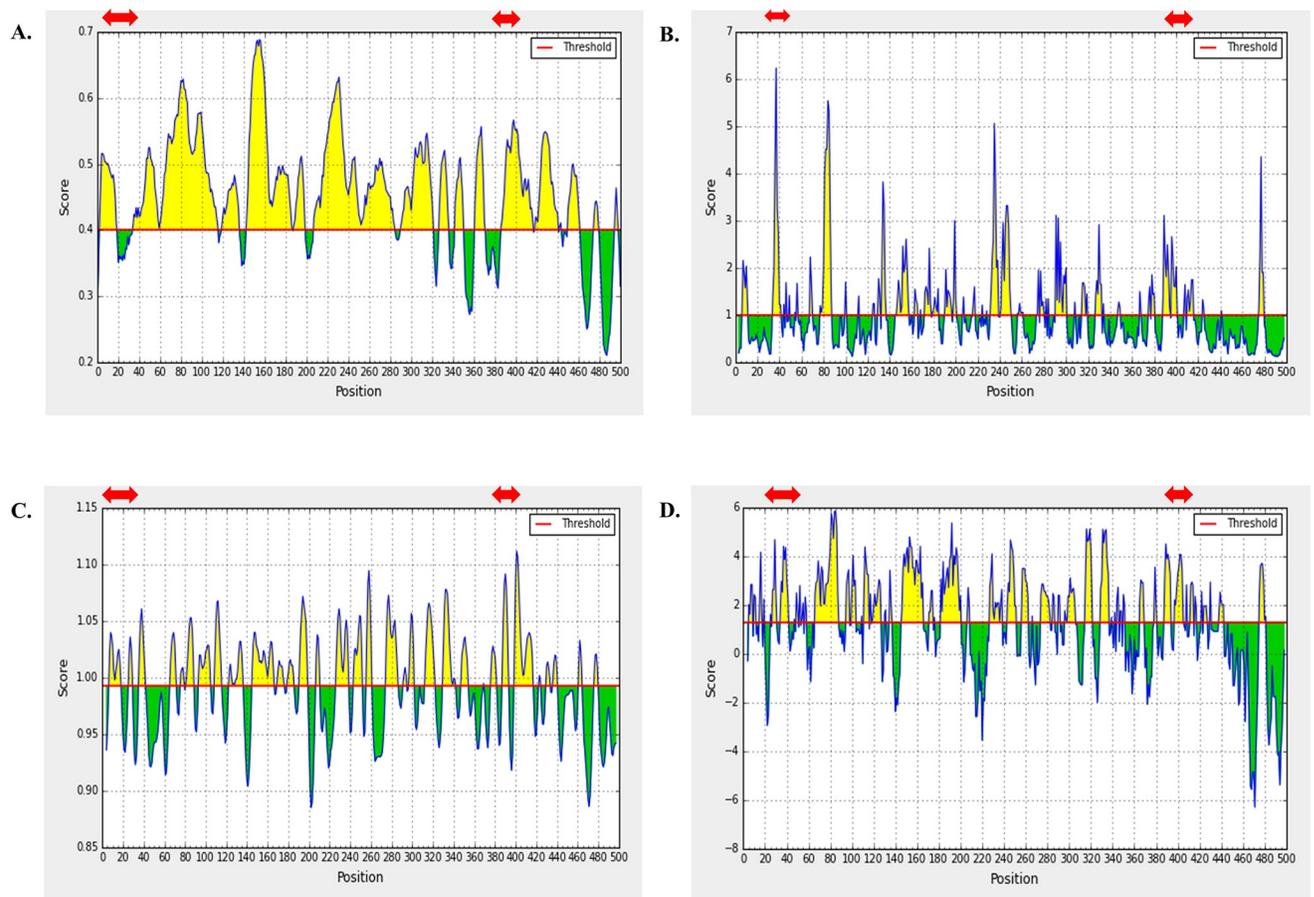
<https://doi.org/10.1371/journal.pone.0253393.g002>

### CD8+ T cell (CTL) epitope prediction

NetCTL 1.2 server was used to predict the T cell epitopes of corresponding WNV proteins (both structural and non-structural). Full screening of epitope is provided in Supplementary [S2 File](#). Using 75% threshold value, NetCTL 1.2 server predicted a total of 791 epitopes for all 10 proteins against 12 MHC supertypes of humans. Among those epitopes, 88 were predicted as antigenic, immunogenic, non-allergenic, non-toxic and showed outside topology of trans-membrane helices. After cross checking in NetMHC 4.0 server a total of 32 epitopes were revealed that had capacity to bind to different mouse alleles; indicating those epitopes able to illicit immunogenicity in both humans and mice. However, NS4B protein failed to generate any common epitope. One epitope from each protein was then shortlisted on the basis of respective human HLA alleles (epitopes which can bind to most HLA alleles under low percentile rank e.g. IEDB percentile rank  $\leq 1$ ) for additional analysis ([S2 File](#)).

### CD4+ T cell (HTL) epitope prediction

After CD8+ T cell prediction, IDEB MHC II binding database predicted a total of 497 epitopes from all 10 proteins of WNV that can bind at least one mouse H-2-I allele. Among these epitopes, only 124 were selected that predicted to be antigenic, non-allergenic, non-toxic, and



**Fig 3. B cell Epitope prediction of Envelope protein at IEDB linear epitope prediction tool.** Predicted peptide region above the threshold values (yellow region) are shown in red arrow. Methods used for prediction are enlisted as: (A) BepiPred 2.0, (B) Emini Surface Accessibility tool, (C) Karplus & Schulz Flexibility tool, and (D) Parker Hydrophilicity tool.

<https://doi.org/10.1371/journal.pone.0253393.g003>

showed outside topology. Furthermore, IFNepitope server predicted 77 epitopes to be IFN- $\gamma$  positive. Therefore, these 77 epitopes were also capable of inducing Th1 humoral responses. However, we found no IFN- $\gamma$  inducing epitope from NS1 protein. Full screening is provided in supplementary [S3 File](#). Based on respective human HLA alleles (e.g. epitopes which can bind to most HLA alleles under low percentile rank [IEDB percentile rank <10]), at most one IFN- $\gamma$  positive epitope from each protein was shortlisted for further analysis ([S3 File](#)).

### Conservancy analysis of shortlisted epitopes

Conserved epitopes against multiple serovars are prerequisite to design a universal broad-spectrum vaccine. Hence, each of the CTL, HTL, and BCL epitopes predicted from previous pipelines were screened with respective 100 homologous protein sequences of all other serovars of WNV. IEDB conservation analysis tool was used for conservancy analysis of epitopes against other serovars of WNV. At 90% sequence identity threshold, all the CTL, HTL, and BCL epitopes showed 100% conservancy ([S1–S3 Files](#)). The dataset for 100 homologous protein sequences of all other serovars of WNV can be found in [S4 File](#).

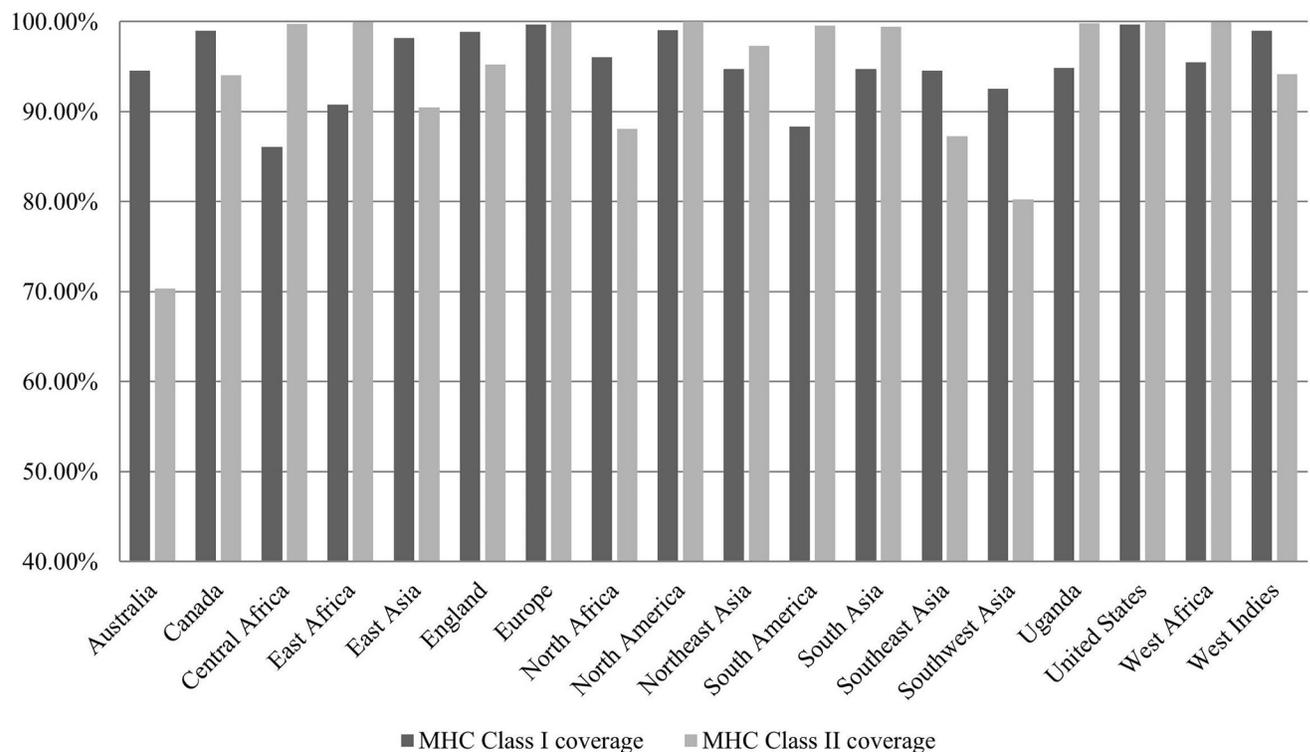
To investigate whether the conserved B cell epitopes (as predicted by the conservancy analysis) have similar sequence homology IEDB conservancy analysis tool was applied and

compared against the reference polyproteins of Dengue and Japanese encephalitis viruses (NCBI accession no: NP\_059433.1, NP\_056776.2, YP\_001621843.1, NP\_073286.1 and UniProtKB ID: P0DOH7). Similarity in epitopes indicate possible cross-reaction. None of these epitopes showed conservancy against the reference proteome of both Dengue and Japanese encephalitis viruses (S1 File); indicating that no cross-reactivity of this epitope with other arboviruses. The dataset for referral polyprotein of Dengue and Zika viruses can be found in S4 File.

### World population coverage

MHC-I and MHC-II prediction server at IEDB predicted the corresponding allele for each of the proposed nine CD8+ (CTL) and nine CD4+ (HTL) epitopes. The output was given in units of IC50 nM. Peptides with IC50 values of <50, <500 and <5000 nM are considered to have high, intermediate and low affinity respectively. Here, alleles with IC50 value <500 were selected as optimum binders for the population coverage analysis.

Population coverage was calculated with corresponding HLA alleles (IC50 <500 nm) of selected nine CTL and nine HTL epitopes (one epitope from each protein) through population coverage tool at IEDB interface. We observed 98.55% (CTLs) and 98.99% (HTLs) of world population coverage. A strong immune response requires activation of both CD4+ and CD8+ T cells. Population coverage analysis of the combined MHC-I and MHC-II epitopes revealed an average of 99.99% coverage of the population worldwide. Whereas, the Europe, England, North America, United States, and West Africa showed the highest (100%) population coverage of the combined MHC-I and MHC-II epitopes (Fig 4).



**Fig 4. Population coverage analysis of finalized epitopes (MHC-I and MHC-II).** Europe, England, North America, United States, and West Africa showed the highest (100%) population coverage.

<https://doi.org/10.1371/journal.pone.0253393.g004>

## Vaccine construction

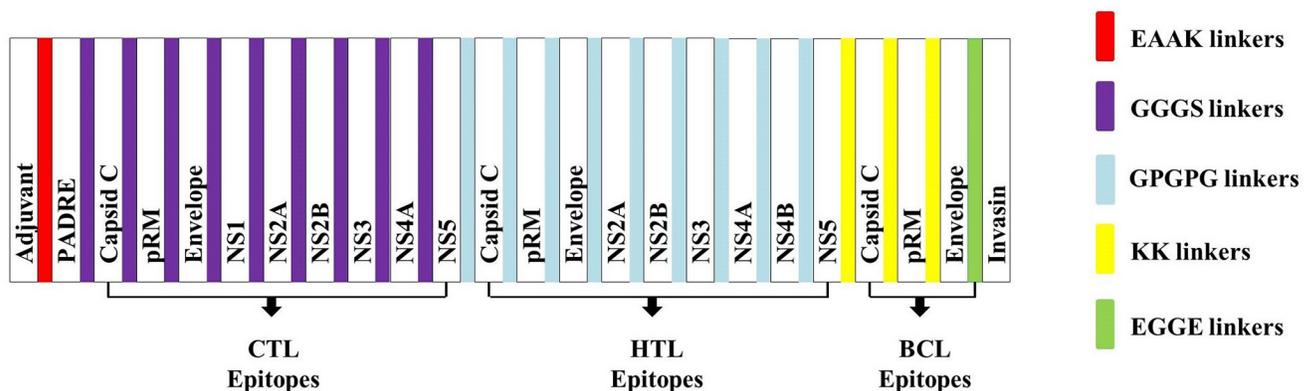
The chimeric vaccine was constructed with final nine CTL, nine HTL, and six B cell epitopes (BCL). Epitopes were joined based on genomic arrangement of their corresponding proteins in virus. GGGG linkers were added to separate each CTL epitope from other CTL epitopes, whereas GPGPG linkers were used to separate each HTL epitope from other HTL epitopes. KK linker was added to separate each BCL epitope from other BCL epitopes. Beta defensin-3 (GIINTLQKYYCRVRGGRCVLSCLPKKEEQIGKCSTRGRKCCRRKK), a TLR4 agonist, was added at N-terminal end of the construct followed by EAAAK linker and PADRE sequence (AKFVAAWTLKAAA). GGGG linker was used to separate PADRE sequence from CTL. The C-terminal invasin sequence (TAKSKKFPSYTATYQF) of *Yersinia* was added at the C-terminal end of the construct which was separated from the BCL epitope by an EGGE linker. These linkers would ensure maximal immunity and efficient presentation of epitopes in the body while PADRE sequence would induce CTL responses and ensure maximal MHC-II allele coverage. The final construct of multi-subunit vaccine comprises 488 amino acid residues and is represented in Fig 5 and S5 File.

## Immunogenic and physiochemical profiles

The final vaccine construct was found antigenic (VaxiJen score 0.7) in nature. Physiochemical analysis revealed that the construct was highly stable (Instability index 29.55), basic in nature (theoretical PI 10.21), and highly thermostable (AI value 84.36). The estimated half-life was found 30 hours in mammalian reticulocytes (*in vitro*), more than 20 hours in case of yeast cell (*in vivo*) and over 10 hours if expressed in *E. coli* (*in vivo*). Molecular weight (50.914 kDa) of the construct was in favorable range. The recombinant vaccine protein was also predicted to be soluble (0.551) in *E. coli* heterologous expression system. Furthermore, AlgPred predicted the construct to be non-allergenic in nature (Table 2).

## Secondary and tertiary structure prediction, refinement and validation

Secondary structure was predicted on the basis of the amino acid sequence of the final vaccine construct using PSIPRED. It was revealed that the predicted structure of the vaccine protein has 103 alpha helix (21%), 145 sheet (29%), and 240 (50%) coil regions.



**Fig 5. Schematic representation of WNV vaccine construct:** A 488 amino acids long multi-epitope vaccine sequence consisting an adjuvant at N-terminal end linked with a multi-epitope sequence with the help of EAAAK linker (Red) followed by PADRE sequence. CTL epitopes were linked through GGGG linkers (blue). HTL epitopes were linked through GPGPG linkers (Sky). BCL epitopes were linked through KK linkers (Yellow). Invasin molecule linked through EGGE linker (Green).

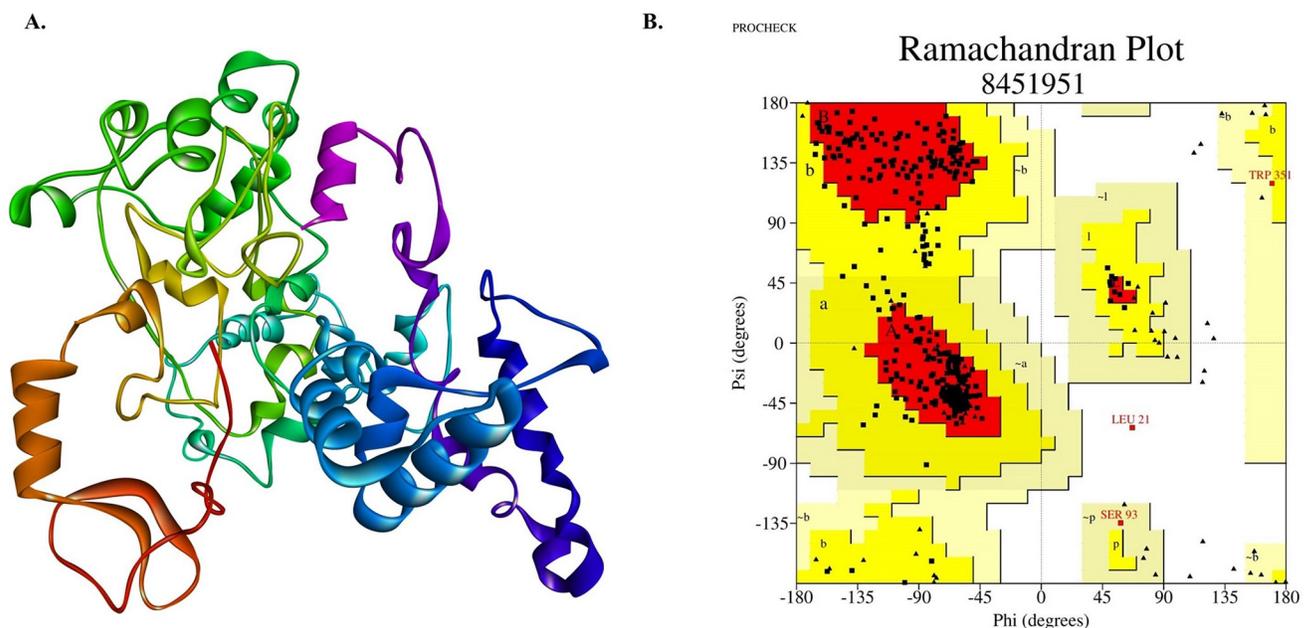
<https://doi.org/10.1371/journal.pone.0253393.g005>

Table 2. Physicochemical, immunogenic and solubility properties of vaccine protein.

Parameter	Value	Comment
Molecular Weight (kDa)	50.914	Suitable
Theoretical PI	10.21	Basic in nature
Estimated Half Life	30 h	Mammalian reticulocyte ( <i>in vitro</i> )
	>20 h	Yeast ( <i>in vivo</i> )
	>10 h	<i>E. coli</i> ( <i>in vivo</i> )
Instability Index	29.55 (stable)	Stable
Aliphatic Index	84.36	Thermostable
GRAVY	0.170	Slightly hydrophobic
VaxiJen Score	0.7042	Antigenic
AlgPred	-	Non-allergic
Protein-Sol	0.551	Highly soluble in <i>E. coli</i>

<https://doi.org/10.1371/journal.pone.0253393.t002>

ROBETTA server generated 5 models (3D) of vaccine protein using comparative modeling. Among those, pre-refinement quality analysis through PROCHECK tool revealed model no. 5 as the best candidate as it had highest number of residues in most favored region (83.1%). To improve the standard of predicted 3D modeled structure beyond the accuracy, refinement was performed using ModRefiner. Refined structure showed RMSD value of 1.268 and TM-score 0.9755 to initial model (model no. 5). Post-refinement quality analysis through PROCHECK tool revealed improved Ramachandran plot as 88.1% residues are in most favored regions, 11.1% in additionally allowed region, 0.5% in generously allowed region, and only 0.3% in disallowed region (Fig 6 and S1 Table).



**Fig 6. Tertiary structure and Ramachandran plot assessment of the refined vaccine protein.** a) Model-5 from Robetta server (Homology modelling server) after refinement; b) Ramachandran plot analysis of refined vaccine protein by PROCHECK server. Refined model contains 88.1% residues in most favored region, 5.3% residues in additionally allowed region, 0.5% in generously allowed region and only 0.3% in disallowed region.

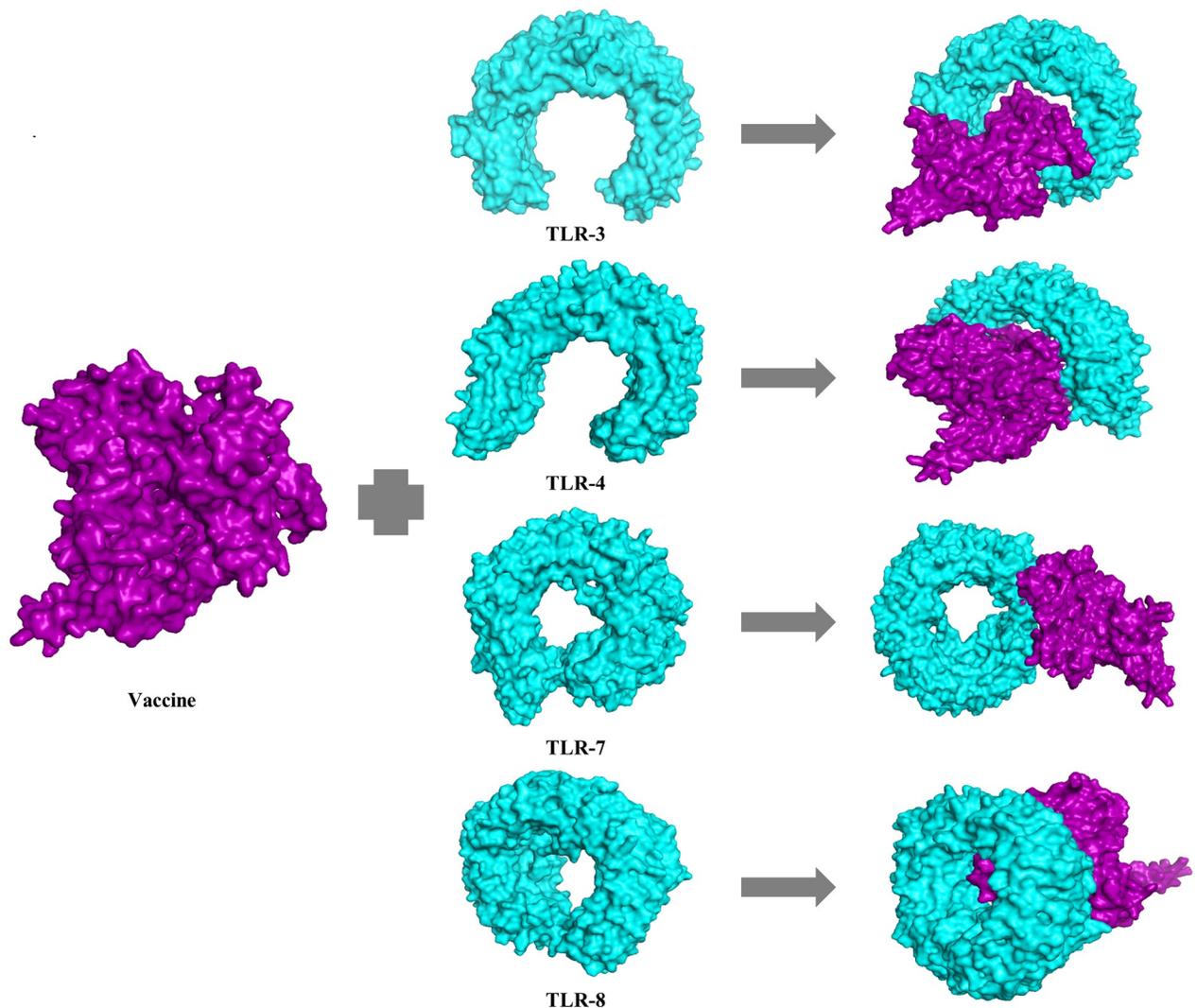
<https://doi.org/10.1371/journal.pone.0253393.g006>

### Conformational B cell epitope prediction

ElliPro at IEDB was utilized for identifying the discontinuous B-cell epitopes. The PDB structure of the vaccine protein was used as an input for the prediction of conformational epitopes in the tertiary structure of the vaccine construct. ElliPro identified seven epitopes on the three-dimensional structure of the vaccine construct where highest predicted score was 0.735 with residual position of 359–386 (S1 Fig and S6 File).

### Docking analysis of the vaccine construct with toll like receptor (TLR)

Vaccine construct was docked with human TLR- 3, 4, 7, and 8 receptors (Fig 7). PatchDock webservers generated a total of 20 best protein-ligand complexes of vaccine protein and all the different TLRs as output along with respective free binding energy. Patchdock server ranked top ten docked models based on the protein's surface as well as electrostatic complementarity. FireDock web tool was further utilized to refine and rescore the docked complexes. After



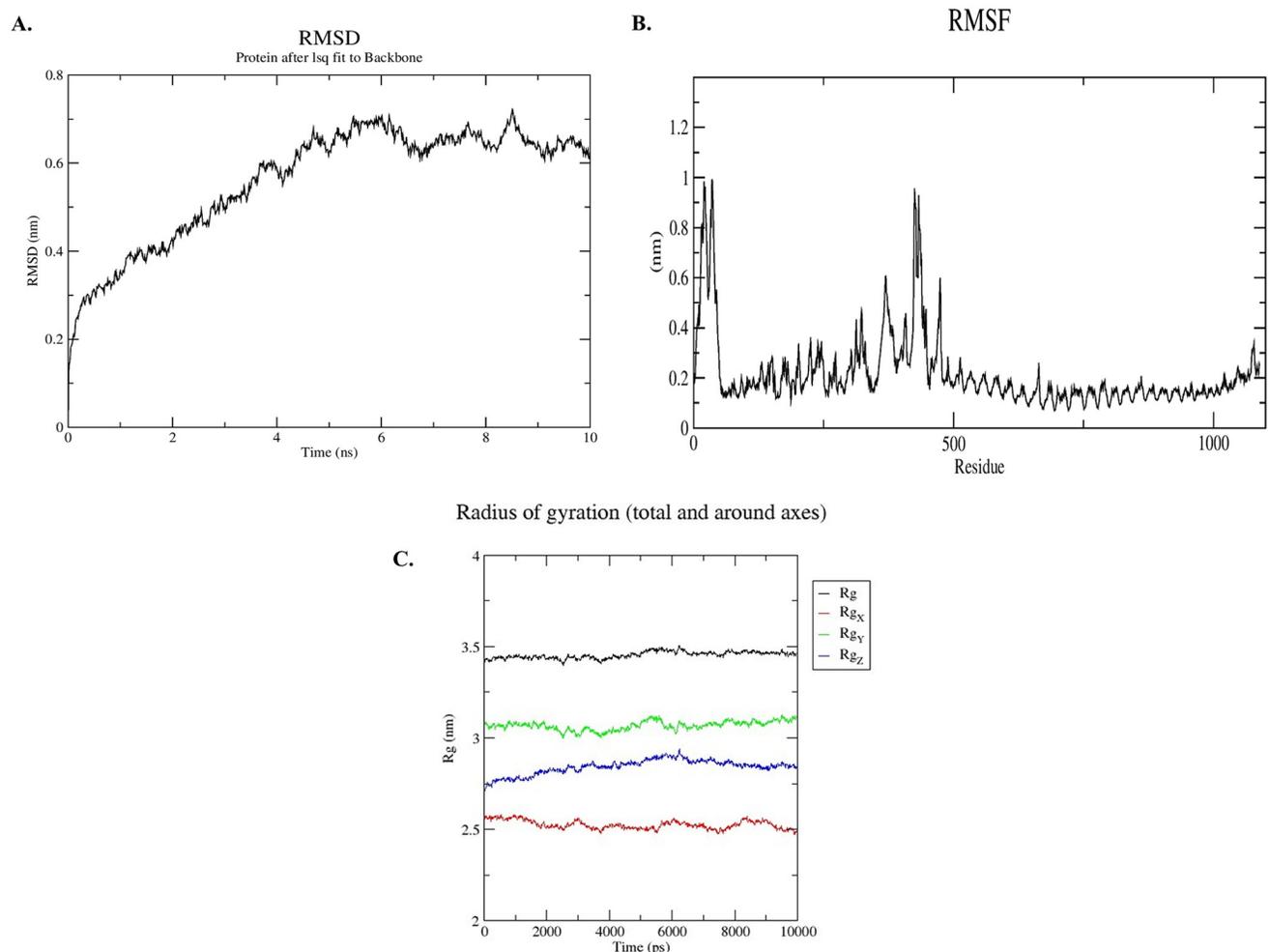
**Fig 7. Ligand-receptor docked complex of vaccine protein with human toll like receptor (TLR)-3, 4, 7 and 8.** Vaccine protein is represented as purple color while TLR-3, 4, 7, 8 as aqua color. The global energy of vaccine protein with other toll like receptors were found -22.08 for TLR3, -26.96 for TLR4, -0.06 for TLR7 and -6.29 for TLR respectively.

<https://doi.org/10.1371/journal.pone.0253393.g007>

refinement, solution 2 was ranked as the best model for TLR-4 and vaccine complex whereas solution 1 for TLR-3, solution 5 for TLR-7 and solution 1 for TLR-8 and vaccine complex. However, the best docked complex was found to be TLR-4 and vaccine complex with global energy of -26.96 which was selected for further molecular dynamics simulation. Other parameters like atomic contact energy (ACE), electrostatic interactions, and Van Der Waals associations were 14.10, 28.65, and -42.76 respectively. The global energy of vaccine protein with other TLRs were -22.08 for TLR3, -0.06 for TLR7, and -6.29 for TLR8. The lowest binding indicates that the docked complex signifies the highest binding affinity between vaccine construct and TLR-4.

### Molecular dynamics simulation of receptor ligand complex

The molecular dynamics simulation was performed for 10ns timescale and RMSD, RMSF, and Rg plots were drawn to check the flexibility and stability of the vaccine-receptor complex. The RMSD plot showed significant results highlighting the stability of the vaccine-receptor complex (Fig 8A). The RMSD of the protein fitting to its backbone showed deviations initially



**Fig 8. Molecular dynamics simulation of the vaccine protein with human toll like receptor-4.** (A) RMSD-Root Mean Square Deviation of docked complex shows very minor deviation reflecting the stable interaction between vaccine protein and TLR-4 molecule. (B) RMSF-Root Mean Square Fluctuation plot of docked protein complex remained fairly flat reflecting the flexibility of side chain of docked protein complex. (C) Rg plot of docked complex shows a relatively flat curve suggesting stable vaccine-receptor complex.

<https://doi.org/10.1371/journal.pone.0253393.g008>

which was later stabilized after 5ns. The vaccine-receptor complex showed deviation of less than 0.2 nm. The deviations higher than that correlate to instability. The vaccine-receptor showed less deviation after 5ns mark indicates the stability of the complex.

Root means square fluctuations (RMSF) helps to explain the protein area that fluctuate throughout the simulation. The flexibility of each residue is therefore measured in order to provide a deeper insight into the degree to which protein binding influences complex's flexibility. The RMSF plot showed very less fluctuations of the atoms in the complex (Fig 8B); suggesting that the amino acid residues of the protein remained stable throughout the simulation. This strengthens the fact that the complex was stable after the docking of the vaccine on to the receptor.

As the radius of gyration shows the compactness of the complex structure, there should be fewer movements or deviation of the positions of the residues of the complex in the simulation trajectory in order to deem the complex to be stable. The Rg plot (Fig 8C) showed less deviation throughout the simulation process and the relatively flat curve suggesting the compactness and stability of the complex.

The molecular dynamics simulation results showed significant structural flexibility and stability of the vaccine-receptor complex; thus implies a positive response by the immune receptor.

### ***In silico* cloning in adenovirus vector and mRNA secondary structure analysis**

GenScript Codon Optimization tool (<https://www.genscript.com/>) was utilized to generate reversely transcribed nucleotide sequences of recombinant vaccine protein for human cell expression system. GC content of obtained cDNA (1467 bp) was 62.25% whereas CAI value was calculated as 0.92. These values above indicate efficient expression of optimized codon in human host cell. This sequence was inserted as GOI under strong CMV promoter in pAd-Track-CMV vector through SnapGene v5.1 software. *BglIII* restriction site was added at the 5' end followed by Kozak sequence which ensures maximal expression of the desired gene. Stop codon TAA was added at 3' end of the optimized codon followed by *EcoRV* restriction site. Stop codon would ensure termination of gene translation (S2A Fig). Total size of recombinant plasmid contained 10688 base pairs.

mRNA secondary structure was predicted through aforementioned RNA fold server. The minimum free energy -624.40 kcal/mol indicates thermodynamic stability of mRNA structure. Furthermore, first 10 nucleotides of mRNA secondary structure were free from forming any pseudoknot or long stable hairpin thereby ensuring efficient translation initiation from mRNA structure (S2B Fig).

## **Discussion**

West Nile Virus is recognized as a reemerging global pathogen and now endemic in African continent and subcontinent of Asia and Europe [89,90]. The virus is also endemic of encephalitis and meningitis in humans in United States [91]. Along with other deadliest mosquito borne virus such as Chikungunya virus and Zika virus, WNV is enlisted as the B category pathogen by National Institute of Allergy and Infectious Diseases (NIAID) [92]. However, there is no FDA-approved therapeutic treatment available for the diseases caused by WNV and even no commercial vaccine has developed so far [93]. While economically insignificant market and high cost-to-benefit ratio hinder the development and commercialization of WNV vaccine for human use [94], researchers have conducted several studies to find out immunodominant epitopes [7,95–97].

Recent advances in computational biology, immunoinformatics, and reverse vaccinology can facilitate in designing safe and efficient vaccines in a time and cost-effective manner [29,98]. Biological data such as genomics and proteomics are being used to mine potential epitopes, design long lasting immunogenic subunit vaccine and laboratory validation of that candidate [99,100]. Therefore, the present study was undertaken to screen out the potential immunodominant B cell as well as CTL and HTL epitopes from all different structural and non-structural proteins of WNV virus through immunoinformatics approach to design a safe, chimeric multi-subunit vaccine against the human host.

In this study, Linear B cell epitopes were screened only from structural proteins (E, PrM, and Capsid C) of WNV. Six epitopes, two from each protein (E<sub>30-45</sub>: CVTIMSKDKPTIDVKM, E<sub>383-398</sub>: YIVVGRGEQQINHHWH, PrM<sub>56-71</sub>: LSAGNDPEDIDCWCTK, PrM<sub>93-108</sub>: SLTVQTHGESTLANKK, Capsid C<sub>2-17</sub>: SKKPGGPGKSRAVNML Capsid C<sub>17-32</sub>: LKRGMPrVLSLIGLKR), showed potential antigenicity, surface accessibility, non-toxicity, and non-allergenicity. These epitopes may induce production of WNV specific antibodies in human. Neutralizing antibodies (IgM and IgG) appear to terminate viremia by controlling hematogenous spread of the WNV virus to Central Nervous System [101]. It was reported that highly immunogenic E protein of WNV, specifically part of domain II and III, showed promising results to generate neutralizing antibodies in clinical trial [102]. Domain III of E protein consists of many neutralizing epitopes [89,103–107] such as amino acid residues of E<sub>302-309</sub>, E<sub>303-333</sub>, E<sub>333-365</sub>, and E<sub>389-391</sub> [108]. Even passively transferred E protein antisera eliminated viremia in mice and provided protection against WNV infection [109,110]. Passively transferred WNV immune sera thus protected B cell deficient mice from WNV infection [109]. One of our selected BCL E epitopes (E<sub>383-398</sub>: YIVVGRGEQQINHHWH) for vaccine construction contains experimentally proved amino acid residues capable of inducing neutralizing antibodies. Likewise E antigens, PrM antigens incorporated in ChimeriVax-WNV2 vaccine also induced neutralizing antibodies and until now are thought to be the most promising candidate as a human vaccine for WNV [111]. Although antibodies play crucial role in terminating viremia by controlling viral replication, T cell mediated responses are also vital and essential components in eradication of WNV from the body [101,112–117]. CTLs targets E, M, Capsid C, NS3, NS4A, and NS4B proteins of WNV [118]. In our study, we screened all structural and non-structural proteins of WNV to find out immunogenic CD8<sup>+</sup> T cell epitopes. MHC class I epitopes selected for vaccine construction (E<sub>480-488</sub>: SIALTFLAV, PrM<sub>153-161</sub>: VVFVLLLL, Capsid C<sub>51-59</sub>: LAFFRFTAI, NS1<sub>120-128</sub>: KSILFAPEL, NS2A<sub>9-17</sub>: FQLGLLVVF, NS2B<sub>13-21</sub>: LMFAIVGGL, NS3<sub>358-366</sub>: KTVWFVPSV, NS4A<sub>63-71</sub>: MTMGVFLL, NS5<sub>486-494</sub>: LEFEALGFL) were immunogenic, non-toxic, and non-allergenic. These were predicted to bind with human alleles and at least one mouse allele. From published studies, we found an epitope (E<sub>480-488</sub>: SIALTFLAV) of E protein overlapped with experimentally validated MHC class I binding epitope (E<sub>481-488</sub>: IALTFLAV) in mouse model [95]. MHC class I molecule or T cell deficient mice have shown increased mortality while surviving mice have exhibited sustained and higher WNV burdens in both CNS and spleen [101,114]. Several studies reported that patients suffering from acquired deficiencies, genetic diseases with impaired T cell function and hematologic malignancies, have increased risk of developing West Nile Neuro-invasive Disease [119,120]. CTLs recognizing a WNV infected cell, proliferate and release different pro-inflammatory cytokines to directly lyse the infected cell [115,121,122]. Secretion of Interferon- $\gamma$  by  $\alpha\beta$  and  $\gamma\delta$  T cells, a pro-inflammatory cytokine, showed reduced WNV load in mice brains. Adoptive transfer of these cells also protected mice by reducing the susceptibility to lethal WNV infection [119–121]. Hence, considering these aspects, selected CTL epitopes may trigger potential T cell immune responses *in vivo*.

CD4<sup>+</sup> T cell (HTL) is another key immune cell type that acquires Th1 or Th2 phenotypes and stimulates immune responses [123,124]. CD8<sup>+</sup> T cell, natural killer cell, and macrophages are activated by Th1 response. Whereas Th2 is responsible for the activation of B cell, isotype switching of B cells, affinity maturation, and antibody production which eliminates extracellular pathogens [115,116,125,126]. In our study, to induce HTL responses through vaccination, screening of potential HTL epitopes were performed. Shortlisted MHC class II epitopes from all structural and non-structural proteins of WNV were antigenic, non-toxic, and non-allergenic that can induce strong HTL immune responses. Importance of CD4<sup>+</sup> T cell responses in WNV were experimentally evaluated. Experiments in mouse models with CD4<sup>+</sup> T cell and MHC class II deficiency aided in longer WNV persistence in brain and spinal cord. Levels of IgM and IgG were significantly decreased in MHC class II and HTL deficient mouse after a certain period and prolonged the viral persistence [127]. Only IFN- $\gamma$  inducing immunogenic HTL epitopes (Capsid C<sub>52-66</sub>: AFFRFTAIAPTRAVL, PrM<sub>129-143</sub>: LRNPGYALVAAVIGW, E<sub>356-370</sub>: VTVNPFVSVATANAK, NS2A<sub>85-99</sub>: IQPVFMVASFLKARW, NS2B<sub>103-117</sub>: VCLAI-SAYTPWAILP, NS3<sub>544-558</sub>: PVWLAYKVAAGVSY, NS4A<sub>68-82</sub>: FFLMQRKGIGKIGL, NS4B<sub>34-48</sub>: LDLRPATAWSLYAVT, NS5<sub>475-489</sub>: AIWFMWLGARFLEFE) were used in vaccine construct since IFN- $\gamma$  induces MHC class II expression and activation of macrophages [128]. It is well known that  $\gamma\delta$  T cells also play an important role in WNV infection by releasing IFN- $\gamma$  molecule in a MHC independent restriction manner [118].

Conservancy analysis of epitopes in different WNV serovars suggested universal candidacy of the designed vaccine. Also, population coverage calculated for conserved MHC class I and II epitopes showed >99.99% worldwide coverage in a combined manner. EAAAK linker used in the constructed vaccine would ensure *in vivo* separation of individual epitopes effectively in natural environment [67,80]. As GGGS linker is superior to AAY linker [129] in epitope based vaccine, we used this linker to separate CTL epitopes from PADRE sequence as well as individual CTL epitopes. HLA alleles are highly polymorphic in different populations, PADRE sequence was incorporated in the vaccine construct to overcome the problem. PADRE sequence binds to a wide variety of DR alleles [130] and ensure maximum efficacy and potency of subunit vaccines and also provide better CTL responses than PADRE deficient vaccines [131,132]. The adjuvant used in our study (Beta-defensin-3) demonstrated efficiency as an immune-stimulator in wet lab experiments [133–136]. The invasin molecule used in the vaccine construct can enhance the efficiency of human adenovirus based vaccines [137]. Finally, several predicted physiochemical parameters ensured immunogenicity, stability of vaccine construct and solubility in *E. coli* heterologous expression system [32]. The constructed multi-subunit vaccine showed higher binding affinity to human TLR-4 than TLR-3, TLR-7 and TLR-8. Molecular Dynamics simulation proved the stability of TLR-4 bound vaccine-receptor complex in cellular environment. Moreover, reversely transcribed optimized nucleotides of Adenovirus based subunit vaccine were found ideal for better heterologous expression. Kozak sequence incorporated at the upstream of optimized cDNA facilitate *in vivo* expression of stable mRNA [138] and may aid its recognition by eukaryotic ribosomes for efficient translation [139]. In recent years, new platforms of vaccine development such as virus like particles (VLPs), DNA, and mRNA-based vaccines are gaining attention. Vaccines are produced in both non-mammalian expression systems (bacterial, yeast, plant, and insect) and mammalian expression systems (including human cell lines) [140]. Mammalian expression system promote proper folding and post-translational modification (PTM) of designed vaccine, generally the preferred platform [140,141]. VLPs, however, do not favor mammalian expression systems for subunit vaccines due to its high production costs and low production yield [142]. On the contrary, adenoviral vector based DNA vaccines have been widely studied and extensively

evaluated [143,144]. Adenoviral vaccines can be easily generated with higher tier at relatively low production cost [145].

Adenovirus based vaccines showed promising results in clinical trials against several infectious diseases such as HIV, Malaria, Hepatitis C virus, Ebola virus, Tuberculosis, Rotavirus and also cancers such as Prostate cancer, Lymphoma, melanoma, and others [146]. It is noteworthy that several Adenovirus based COVID-19 vaccines have already been approved for human use such as ChAdOx1 nCoV-19 or AZD1222 by University of Oxford and AstraZeneca, AD26.COV.2.S or JNJ-78436735 by Johnson & Johnson, Sputnik V by Gamaleya and Ad5-nCoV by CanSino Biologics Inc [147]. Johnson & Johnson have developed Adenovirus based Ebola Vaccine (Ad26.ZEBOV) which is currently in use with another vaccine as second dose [148]. They are also conducting clinical trial on Adenovirus based vaccine for Zika virus. Hence, Adenovirus based vaccine would be a potential candidate for eukaryotic expression system and many pharmaceutical companies have great interest for in Adenovirus based commercial vaccine development. Although our proposed multi-subunit vaccine showed promising results, it is based only on computational and reverse vaccinomics approaches. Therefore, to find out the most efficient candidate against WNV infection, we recommend further laboratory validation of the proposed vaccine construct and also comparison of the effectivity between protein vaccine and Adenovirus based vaccine.

## Conclusion

West Nile virus (WNV) is continuously spreading across Europe and other continents, i.e. North and South America, Asia, and many other regions of the world. Using immunoinformatics approach, we designed a multi epitope-based subunit vaccine against WNV. Human (HLA) and Mice (H-2) allele specific immunogenic epitopes including Cytotoxic T-lymphocytes (CTL), Helper T-lymphocytes (HTL), and B cell lymphomas (BCL) were screened and selected for vaccine design. Final vaccine constructs consist of B and T-cell epitopes, Universal pan-HLA DR or PADRE (13 mer) sequence,  $\beta$ -defensin-3 adjuvant (45 mer), and Invasin molecule at N-terminal of the construct with fusion protein linkers. Whereas, Invasin molecule (16 mer) was added at the C-terminal end to elicit strong immune responses. Conservancy and population coverage analysis of the promiscuous epitopes revealed the robust immune response against multiple serovars of the WNV and various ethnicities. Molecular docking revealed stronger binding affinity with human TLR-4; later confirmed by molecular dynamics simulation. Our computational analysis suggests that the constructed vaccine was highly immunogenic, safe, non-toxic, and stable in nature. Eventually, *in silico* cloning approaches of the vaccine construct ensure high expression within microbial system.

## Supporting information

**S1 Fig. Discontinuous B cell epitope (residual position of 359–386 highlighted in purple color) in the 3D model of Vaccine protein.** The Highest predicted score was obtained 0.735. (TIFF)

**S2 Fig. *In silico* restriction cloning of the gene sequence of final vaccine construct into pAdV-Track-CMV vector.** a) Gene of Interest was inserted between *Bgl II* and *EcoRV* restriction sites (red marked). Recombinant plasmid consists of 10688 base pairs. b) Predicted secondary structure of mRNA for vaccine expressed through pAdVTrack-CMV. The 5' end of predicted mRNA structure does not contain any pseudoknot or hairpin. (TIF)

**S1 Table. Quality assessment of vaccine models (3D structures) and refined structure by PROCHECK server.**

(DOCX)

**S1 File. Linear B Lymphocytes (BCL) epitope screening from structural proteins (Envelope, PrM, Capsid C) of WNV.**

(XLSX)

**S2 File. Cytotoxic T Lymphocytes (CTL) epitope screening from structural and non-structural proteins of WNV.**

(XLSX)

**S3 File. Helper T Lymphocytes (HTL) epitope screening from structural and non-structural proteins of WNV.**

(XLSX)

**S4 File. Dataset of the sequences used in the study.**

(DOCX)

**S5 File. Sequence composition of the vaccine construct (Amino acid sequences and Optimized codons).**

(DOCX)

**S6 File. Discontinuous epitopes predicted in the final vaccine structure.**

(XLSX)

**S1 Graphical abstract.**

(TIF)

## Author Contributions

**Conceptualization:** Md Tahsin Khan, Arzuba Akter, Shakhinur Islam Mondal.

**Data curation:** Md Tahsin Khan, Rahatul Islam, Araf Mahmud, Ahasanul Kobir, Shakhinur Islam Mondal.

**Formal analysis:** Md Tahsin Khan, Rahatul Islam, Tarhima Jahan Jerin, Araf Mahmud, Sahara Khatun, Ahasanul Kobir.

**Investigation:** Arzuba Akter.

**Methodology:** Md Tahsin Khan, Rahatul Islam, Tarhima Jahan Jerin, Araf Mahmud, Arzuba Akter, Shakhinur Islam Mondal.

**Project administration:** Arzuba Akter, Shakhinur Islam Mondal.

**Resources:** Arzuba Akter, Shakhinur Islam Mondal.

**Supervision:** Shakhinur Islam Mondal.

**Validation:** Md Tahsin Khan, Rahatul Islam, Md Nahidul Islam, Shakhinur Islam Mondal.

**Visualization:** Shakhinur Islam Mondal.

**Writing – original draft:** Md Tahsin Khan, Rahatul Islam, Arzuba Akter, Shakhinur Islam Mondal.

**Writing – review & editing:** Arzuba Akter, Shakhinur Islam Mondal.

## References

1. Final Cumulative Maps and Data: Centers for Disease Control and Prevention; 2019. <https://www.cdc.gov/westnile/statsmaps/cumMapsData.html>.
2. Kramer LD, Styer LM, Ebel GD. A global perspective on the epidemiology of West Nile virus. *Annu Rev Entomol.* 2008; 53:61–81. <https://doi.org/10.1146/annurev.ento.53.103106.093258> PMID: 17645411
3. Murray KO, Garcia MN, Rahbar MH, Martinez D, Khuwaja SA, Arafat RR, et al. Survival analysis, long-term outcomes, and percentage of recovery up to 8 years post-infection among the Houston West Nile virus cohort. *PloS one.* 2014; 9(7).
4. Weatherhead JE, Miller VE, Garcia MN, Hasbun R, Salazar L, Dimachkie MM, et al. Long-term neurological outcomes in West Nile Virus–Infected patients: an observational study. *The American journal of tropical medicine and hygiene.* 2015; 92(5):1006–12. <https://doi.org/10.4269/ajtmh.14-0616> PMID: 25802426
5. Nolan MS, Podoll AS, Hause AM, Akers KM, Finkel KW, Murray KO. Prevalence of chronic kidney disease and progression of disease over time among patients enrolled in the Houston West Nile virus cohort. *PloS one.* 2012; 7(7). <https://doi.org/10.1371/journal.pone.0040374> PMID: 22792293
6. Zeller HG, Schuffenecker I. West Nile virus: an overview of its spread in Europe and the Mediterranean basin in contrast to its spread in the Americas. *European Journal of Clinical Microbiology and Infectious Diseases.* 2004; 23(3):147–56. <https://doi.org/10.1007/s10096-003-1085-1> PMID: 14986160
7. Waller FM, Reche PA, Flower DR. West Nile Virus Vaccine Design by T Cell Epitope Selection: In Silico Analysis of Conservation, Functional Cross-Reactivity with the Human Genome, and Population Coverage. *Journal of Immunology Research.* 2020; 2020.
8. Ronca SE, Murray KO, Nolan MS. Cumulative incidence of West Nile virus infection, continental United States, 1999–2016. *Emerging infectious diseases.* 2019; 25(2):325. <https://doi.org/10.3201/eid2502.180765> PMID: 30666940
9. Rapaport DC, Rapaport DCR. *The art of molecular dynamics simulation: Cambridge university press;* 2004.
10. Napp S, Petrić D, Busquets NJP. West Nile virus and other mosquito-borne viruses present in Eastern Europe. *health g.* 2018; 112(5):233–48. <https://doi.org/10.1080/20477724.2018.1483567> PMID: 29979950
11. Khan SA, Dutta P, Khan AM, Chowdhury P, Borah J, Doloi P, et al. West Nile virus infection, Assam, India. 2011; 17(5):947.
12. Anukumar B, Sapkal GN, Tandale BV, Balasubramanian R, Gangale DJJoCV. West Nile encephalitis outbreak in Kerala, India, 2011. 2014; 61(1):152–5. <https://doi.org/10.1016/j.jcv.2014.06.003> PMID: 24985196
13. Khatun T, Chatterjee SJToTRSoTM. Emergence of West Nile virus in West Bengal, India: a new report. *Hygiene.* 2017; 111(4):178–84. <https://doi.org/10.1093/trstmh/trx033> PMID: 28673021
14. Shukla J, Saxena D, Rathinam S, Lalitha P, Joseph CR, Sharma S, et al. Molecular detection and characterization of West Nile virus associated with multifocal retinitis in patients from southern India. 2012; 16(1):e53–e9.
15. Tandel K, Sharma S, Dash PK, Shukla J, Parida MJJomv. Emergence of human West Nile Virus infection among pediatric population in Madhya Pradesh, India. 2019; 91(3):493–7.
16. Rodrigues F, Bright Singh P, Dandawate C, Soman R, Guttikar S, Kaul HJJJP. Isolation of Japanese encephalitis and West Nile viruses from mosquitoes collected in Andhra Pradesh. 1980; 4:149–53.
17. Pavri K, Singh KJlJomr. Isolation of West Nile Virus from *Culex fatigans* Mosquitoes from Western India. 1965; 53(6):501–5.
18. Barrett AD. Economic burden of West Nile virus in the United States. *The American journal of tropical medicine and hygiene.* 2014; 90(3):389. <https://doi.org/10.4269/ajtmh.14-0009> PMID: 24515938
19. Staples JE, Shankar MB, Sejvar JJ, Meltzer MI, Fischer M. Initial and long-term costs of patients hospitalized with West Nile virus disease. *The American journal of tropical medicine and hygiene.* 2014; 90(3):402–9. <https://doi.org/10.4269/ajtmh.13-0206> PMID: 24515937
20. Brinton MA. Replication cycle and molecular biology of the West Nile virus. *Viruses.* 2014; 6(1):13–53.
21. Xu Z, Anderson R, Hobman TC. The capsid-binding nucleolar helicase DDX56 is important for infectivity of West Nile virus. *Journal of virology.* 2011; 85(11):5571–80. <https://doi.org/10.1128/JVI.01933-10> PMID: 21411523
22. Kanai R, Kar K, Anthony K, Gould LH, Ledizet M, Fikrig E, et al. Crystal structure of West Nile virus envelope glycoprotein reveals viral surface epitopes. *Journal of virology.* 2006; 80(22):11000–8. <https://doi.org/10.1128/JVI.01735-06> PMID: 16943291

23. Yang J-S, Ramanathan MP, Muthumani K, Choo AY, Jin S-H, Yu Q-C, et al. Induction of inflammation by West Nile virus capsid through the caspase-9 apoptotic pathway. *Emerging infectious diseases*. 2002; 8(12):1379. <https://doi.org/10.3201/eid0812.020224> PMID: 12498651
24. Dayan GH, Pugachev K, Bevilacqua J, Lang J, Monath TP. Preclinical and clinical development of a YFV 17 D-based chimeric vaccine against West Nile virus. *Viruses*. 2013; 5(12):3048–70. <https://doi.org/10.3390/v5123048> PMID: 24351795
25. Durbin AP, Wright PF, Cox A, Kagucia W, Elwood D, Henderson S, et al. The live attenuated chimeric vaccine rWN/DEN4Δ30 is well-tolerated and immunogenic in healthy flavivirus-naïve adult volunteers. *Vaccine*. 2013; 31(48):5772–7. <https://doi.org/10.1016/j.vaccine.2013.07.064> PMID: 23968769
26. Martin JE, Pierson TC, Hubka S, Rucker S, Gordon IJ, Enama ME, et al. A West Nile virus DNA vaccine induces neutralizing antibody in healthy adults during a phase 1 clinical trial. *Journal of Infectious Diseases*. 2007; 196(12):1732–40.
27. Coller B-A, Pai V, Weeks-Levy C, Ogata SA. Recombinant subunit West Nile virus vaccine for protection of human subjects. Google Patents; 2012.
28. Parvizpour S, Pourseif MM, Razmara J, Rafi MA, Omid Y. Epitope-based vaccine design: a comprehensive overview of bioinformatics approaches. *Drug Discovery Today*. 2020. <https://doi.org/10.1016/j.drudis.2020.03.006> PMID: 32205198
29. Kazi A, Chuah C, Majeed ABA, Leow CH, Lim BH, Leow CY. Current progress of immunoinformatics approach harnessed for cellular-and antibody-dependent vaccine design. *Pathogens and global health*. 2018; 112(3):123–31. <https://doi.org/10.1080/20477724.2018.1446773> PMID: 29528265
30. Bahrami AA, Payandeh Z, Khalili S, Zakeri A, Bandehpour M. Immunoinformatics: In Silico Approaches and Computational Design of a Multi-epitope, Immunogenic Protein. *International reviews of immunology*. 2019; 38(6):307–22. <https://doi.org/10.1080/08830185.2019.1657426> PMID: 31478759
31. Zhang L. Multi-epitope vaccines: a promising strategy against tumors and viral infections. *Cellular & molecular immunology*. 2018; 15(2):182–4. <https://doi.org/10.1038/cmi.2017.92> PMID: 28890542
32. Ali M, Pandey RK, Khatoon N, Narula A, Mishra A, Prajapati VK. Exploring dengue genome to construct a multi-epitope based subunit vaccine by utilizing immunoinformatics approach to battle against dengue infection. *Scientific reports*. 2017; 7(1):1–13.
33. Kalyanaraman N. In silico prediction of potential vaccine candidates on capsid protein of human bocavirus 1. *Molecular immunology*. 2018; 93:193–205. <https://doi.org/10.1016/j.molimm.2017.11.024> PMID: 29207326
34. Khan M, Khan S, Ali A, Akbar H, Sayaf AM, Khan A, et al. Immunoinformatics approaches to explore *Helicobacter Pylori* proteome (Virulence Factors) to design B and T cell multi-epitope subunit vaccine. *Scientific reports*. 2019; 9(1):1–13.
35. Chauhan V, Rungta T, Goyal K, Singh MP. Designing a multi-epitope based vaccine to combat Kaposi Sarcoma utilizing immunoinformatics approach. *Scientific reports*. 2019; 9(1):1–15.
36. Michel-Todó L, Reche PA, Bigey P, Pinazo M-J, Gascón J, Alonso-Padilla J. In silico design of an epitope-based vaccine ensemble for Chagas disease. *Frontiers in Immunology*. 2019; 10:2698. <https://doi.org/10.3389/fimmu.2019.02698> PMID: 31824493
37. Majid M, Andleeb S. Designing a multi-epitopic vaccine against the enterotoxigenic *Bacteroides fragilis* based on immunoinformatics approach. *Scientific Reports*. 2019; 9(1):1–15.
38. Ojha R, Pareek A, Pandey RK, Prusty D, Prajapati VK. Strategic development of a next-generation multi-epitope vaccine to prevent Nipah virus zoonotic infection. *ACS omega*. 2019; 4(8):13069–79. <https://doi.org/10.1021/acsomega.9b00944> PMID: 31460434
39. Islam R, Parvez MSA, Anwar S, Hosen MJJIIMU. Delineating blueprint of an epitope-based peptide vaccine against the multiple serovars of dengue virus: A hierarchical reverse vaccinology approach. 2020; 20:100430.
40. Hoque H, Islam R, Ghosh S, Rahaman MM, Jewel NA, Miah MAJH. Implementation of in silico methods to predict common epitopes for vaccine development against Chikungunya and Mayaro viruses. 2021; 7(3):e06396.
41. Khan MT, Islam MJ, Parihar A, Islam R, Jerin TJ, Dhote R, et al. Immunoinformatics and molecular modeling approach to design universal multi-epitope vaccine for SARS-CoV-2. 2021; 24:100578.
42. Jiang P, Cai Y, Chen J, Ye X, Mao S, Zhu S, et al. Evaluation of tandem *Chlamydia trachomatis* MOMP multi-epitopes vaccine in BALB/c mice model. *Vaccine*. 2017; 35(23):3096–103. <https://doi.org/10.1016/j.vaccine.2017.04.031> PMID: 28456528
43. Zhu S, Feng Y, Rao P, Xue X, Chen S, Li W, et al. Hepatitis B virus surface antigen as delivery vector can enhance *Chlamydia trachomatis* MOMP multi-epitope immune response in mice. *Applied*

- microbiology and biotechnology. 2014; 98(9):4107–17. <https://doi.org/10.1007/s00253-014-5517-x> PMID: 24458565
44. Yusufu M, Shalitanati A, Yu H, Moming A, Li Y, Deng F, et al. Immune responses in mice induced by multi-epitope DNA vaccine and protein vaccine of Crimean-Congo Hemorrhagic Fever Virus. *BioRxiv*. 2019:719724.
  45. Lin X, Chen S, Xue X, Lu L, Zhu S, Li W, et al. Chimerically fused antigen rich of overlapped epitopes from latent membrane protein 2 (LMP2) of Epstein–Barr virus as a potential vaccine and diagnostic agent. *Cellular & molecular immunology*. 2016; 13(4):492–501. <https://doi.org/10.1038/cmi.2015.29> PMID: 25864917
  46. Wilkins MR, Gasteiger E, Bairoch A, Sanchez JC, Williams KL, Appel RD, et al. Protein identification and analysis tools in the ExPASy server. *Methods in molecular biology (Clifton, NJ)*. 1999; 112:531–52. <https://doi.org/10.1385/1-59259-584-7-531> PMID: 10027275.
  47. Saha S, Raghava GPS. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. *Proteins: Structure, Function, and Bioinformatics*. 2006; 65(1):40–8.
  48. Larsen JEP, Lund O, Nielsen M. Improved method for predicting linear B-cell epitopes. *Immunome research*. 2006; 2(1):2. <https://doi.org/10.1186/1745-7580-2-2> PMID: 16635264
  49. Emimi EA, Hughes JV, Perlow D, Boger J. Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. *Journal of virology*. 1985; 55(3):836–9. <https://doi.org/10.1128/JVI.55.3.836-839.1985> PMID: 2991600
  50. Karplus P, Schulz G. Prediction of chain flexibility in proteins. *Naturwissenschaften*. 1985; 72(4):212–3.
  51. Parker J, Guo D, Hodges R. New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites. *Biochemistry*. 1986; 25(19):5425–32. <https://doi.org/10.1021/bi00367a013> PMID: 2430611
  52. Larsen MV, Lundegaard C, Lamberth K, Buus S, Lund O, Nielsen M. Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. *BMC bioinformatics*. 2007; 8(1):424. <https://doi.org/10.1186/1471-2105-8-424> PMID: 17973982
  53. Doytchinova IA, Flower DR. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC bioinformatics*. 2007; 8(1):4. <https://doi.org/10.1186/1471-2105-8-4> PMID: 17207271
  54. Dimitrov I, Flower DR, Doytchinova I, editors. AllerTOP—a server for in silico prediction of allergens. *BMC bioinformatics*; 2013: BioMed Central. <https://doi.org/10.1186/1471-2105-14-S6-S4> PMID: 23735058
  55. Gupta S, Kapoor P, Chaudhary K, Gautam A, Kumar R, Raghava GP, et al. In silico approach for predicting toxicity of peptides and proteins. *PloS one*. 2013; 8(9). <https://doi.org/10.1371/journal.pone.0073957> PMID: 24058508
  56. Calis JJ, Maybeno M, Greenbaum JA, Weiskopf D, De Silva AD, Sette A, et al. Properties of MHC class I presented peptides that enhance immunogenicity. *PLoS computational biology*. 2013; 9(10). <https://doi.org/10.1371/journal.pcbi.1003266> PMID: 24204222
  57. Krogh A, Larsson B, Von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of molecular biology*. 2001; 305(3):567–80. <https://doi.org/10.1006/jmbi.2000.4315> PMID: 11152613
  58. Andreatta M, Nielsen M. Gapped sequence alignment using artificial neural networks: application to the MHC class I system. *Bioinformatics*. 2016; 32(4):511–7. <https://doi.org/10.1093/bioinformatics/btv639> PMID: 26515819
  59. Nielsen M, Lundegaard C, Worning P, Lauemøller SL, Lamberth K, Buus S, et al. Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. *Protein Science*. 2003; 12(5):1007–17. <https://doi.org/10.1110/ps.0239403> PMID: 12717023
  60. Vita R, Mahajan S, Overton JA, Dhanda SK, Martini S, Cantrell JR, et al. The immune epitope database (IEDB): 2018 update. *Nucleic acids research*. 2019; 47(D1):D339–D43. <https://doi.org/10.1093/nar/gky1006> PMID: 30357391
  61. Dhanda SK, Vir P, Raghava GP. Designing of interferon-gamma inducing MHC class-II binders. *Biology direct*. 2013; 8(1):30. <https://doi.org/10.1186/1745-6150-8-30> PMID: 24304645
  62. Bui H-H, Sidney J, Li W, Fusseder N, Sette A. Development of an epitope conservancy analysis tool to facilitate the design of epitope-based diagnostics and vaccines. *BMC bioinformatics*. 2007; 8(1):361. <https://doi.org/10.1186/1471-2105-8-361> PMID: 17897458
  63. Bui H-H, Sidney J, Dinh K, Southwood S, Newman MJ, Sette A. Predicting population coverage of T-cell epitope-based diagnostics and vaccines. *BMC bioinformatics*. 2006; 7(1):153. <https://doi.org/10.1186/1471-2105-7-153> PMID: 16545123

64. Reginald K, Chan Y, Plebanski M, Poh CL. Development of peptide vaccines in dengue. *Current pharmaceutical design*. 2018; 24(11):1157–73. <https://doi.org/10.2174/1381612823666170913163904> PMID: 28914200
65. Rana A, Akhter Y. A multi-subunit based, thermodynamically stable model vaccine using combined immunoinformatics and protein structure based approach. *Immunobiology*. 2016; 221(4):544–57. <https://doi.org/10.1016/j.imbio.2015.12.004> PMID: 26707618
66. Chen X, Zaro JL, Shen W-C. Fusion protein linkers: property, design and functionality. *Advanced drug delivery reviews*. 2013; 65(10):1357–69. <https://doi.org/10.1016/j.addr.2012.09.039> PMID: 23026637
67. Pandey RK, Sundar S, Prajapati VK. Differential expression of miRNA regulates T cell differentiation and plasticity during visceral leishmaniasis infection. *Frontiers in microbiology*. 2016; 7:206. <https://doi.org/10.3389/fmicb.2016.00206> PMID: 26941729
68. Hebditch M, Carballo-Amador MA, Charonis S, Curtis R, Warwicker J. Protein–Sol: a web tool for predicting protein solubility from sequence. *Bioinformatics*. 2017; 33(19):3098–100. <https://doi.org/10.1093/bioinformatics/btx345> PMID: 28575391
69. Saha S, Raghava G. AlgPred: prediction of allergenic proteins and mapping of IgE epitopes. *Nucleic acids research*. 2006; 34(suppl\_2):W202–W9. <https://doi.org/10.1093/nar/gkl343> PMID: 16844994
70. Jones DT. Protein secondary structure prediction based on position-specific scoring matrices. *Journal of molecular biology*. 1999; 292(2):195–202. <https://doi.org/10.1006/jmbi.1999.3091> PMID: 10493868
71. Xu D, Zhang Y. Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization. *Biophysical journal*. 2011; 101(10):2525–34. <https://doi.org/10.1016/j.bpj.2011.10.024> PMID: 22098752
72. Laskowski RA, Rullmann JAC, MacArthur MW, Kaptein R, Thornton JM. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *Journal of biomolecular NMR*. 1996; 8(4):477–86. <https://doi.org/10.1007/BF00228148> PMID: 9008363
73. Dubitzky W, Wolkenhauer O, Yokota H, Cho K-H. *Encyclopedia of systems biology*: Springer Publishing Company, Incorporated; 2013.
74. Sivalingam GN, Shepherd AJ. An analysis of B-cell epitope discontinuity. *Molecular immunology*. 2012; 51(3–4):304–9. <https://doi.org/10.1016/j.molimm.2012.03.030> PMID: 22520973
75. Ponomarenko J, Bui H-H, Li W, Fusseder N, Bourne PE, Sette A, et al. ElliPro: a new structure-based tool for the prediction of antibody epitopes. *BMC bioinformatics*. 2008; 9(1):514. <https://doi.org/10.1186/1471-2105-9-514> PMID: 19055730
76. Solanki V, Tiwari V. Subtractive proteomics to identify novel drug targets and reverse vaccinology for the development of chimeric vaccine against *Acinetobacter baumannii*. *Scientific reports*. 2018; 8(1):1–19.
77. Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ. PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic acids research*. 2005; 33(suppl\_2):W363–W7. <https://doi.org/10.1093/nar/gki481> PMID: 15980490
78. Allen MP. Introduction to molecular dynamics simulation. *Computational soft matter: from synthetic polymers to proteins*. 2004; 23:1–28.
79. Voter AF. A method for accelerating the molecular dynamics simulation of infrequent events. *The Journal of chemical physics*. 1997; 106(11):4665–77.
80. Hajjigharamani N, Nezafat N, Eslami M, Negahdaripour M, Rahmatabadi SS, Ghasemi Y. Immunoinformatics analysis and in silico designing of a novel multi-epitope peptide vaccine against *Staphylococcus aureus*. *Infection, Genetics and Evolution*. 2017; 48:83–94. <https://doi.org/10.1016/j.meegid.2016.12.010> PMID: 27989662
81. Nezafat N, Karimi Z, Eslami M, Mohkam M, Zandian S, Ghasemi Y. Designing an efficient multi-epitope peptide vaccine against *Vibrio cholerae* via combined immunoinformatics and protein interaction based approaches. *Computational biology and chemistry*. 2016; 62:82–95. <https://doi.org/10.1016/j.compbiolchem.2016.04.006> PMID: 27107181
82. Narula A, Pandey RK, Khatoon N, Mishra A, Prajapati VK. Excavating chikungunya genome to design B and T cell multi-epitope subunit vaccine using comprehensive immunoinformatics approach to control chikungunya infection. *Infection, Genetics and Evolution*. 2018; 61:4–15. <https://doi.org/10.1016/j.meegid.2018.03.007> PMID: 29535024
83. Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJ. GROMACS: fast, flexible, and free. *Journal of computational chemistry*. 2005; 26(16):1701–18. <https://doi.org/10.1002/jcc.20291> PMID: 16211538
84. Scott WR, Hünenberger PH, Tironi IG, Mark AE, Billeter SR, Fennen J, et al. The GROMOS biomolecular simulation program package. *The Journal of Physical Chemistry A*. 1999; 103(19):3596–607.

85. Huang J, MacKerell AD Jr. CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *Journal of computational chemistry*. 2013; 34(25):2135–45. <https://doi.org/10.1002/jcc.23354> PMID: 23832629
86. He T-C, Zhou S, Da Costa LT, Yu J, Kinzler KW, Vogelstein B. A simplified system for generating recombinant adenoviruses. *Proceedings of the National Academy of Sciences*. 1998; 95(5):2509–14. <https://doi.org/10.1073/pnas.95.5.2509> PMID: 9482916
87. SnapGene. Insightful Science. [snapgene.com](http://snapgene.com).
88. Gruber AR, Lorenz R, Bernhart SH, Neuböck R, Hofacker IL. The vienna RNA websuite. *Nucleic acids research*. 2008; 36(suppl\_2):W70–W4. <https://doi.org/10.1093/nar/gkn188> PMID: 18424795
89. Davis BS, Chang G-JJ, Cropp B, Roehrig JT, Martin DA, Mitchell CJ, et al. West Nile virus recombinant DNA vaccine protects mouse and horse from virus challenge and expresses in vitro a noninfectious recombinant antigen that can be used in enzyme-linked immunosorbent assays. *Journal of virology*. 2001; 75(9):4040–7. <https://doi.org/10.1128/JVI.75.9.4040-4047.2001> PMID: 11287553
90. Johnson AJ, Martin DA, Karabatsos N, Roehrig JT. Detection of anti-arboviral immunoglobulin G by using a monoclonal antibody-based capture enzyme-linked immunosorbent assay. *Journal of clinical microbiology*. 2000; 38(5):1827–31. <https://doi.org/10.1128/JCM.38.5.1827-1831.2000> PMID: 10790108
91. Nash D, Mostashari F, Fine A, Miller J, O'Leary D, Murray K, et al. The outbreak of West Nile virus infection in the New York City area in 1999. *New England Journal of Medicine*. 2001; 344(24):1807–14.
92. Pathogens NEID. National Institute of Allergy and Infectious Diseases [June 03, 2020]. [https://www.niaid.nih.gov/research/emerging-infectious-diseases-pathogens?fbclid=IwAR39n4fn1t1byJb6bJWFH-UdRxxhNNO445D\\_mQQvzUdVsVz8b-6-XWHMsmA](https://www.niaid.nih.gov/research/emerging-infectious-diseases-pathogens?fbclid=IwAR39n4fn1t1byJb6bJWFH-UdRxxhNNO445D_mQQvzUdVsVz8b-6-XWHMsmA).
93. Iyer AV, Kousoulas KG. A review of vaccine approaches for West Nile virus. *International journal of environmental research and public health*. 2013; 10(9):4200–23. <https://doi.org/10.3390/ijerph10094200> PMID: 24025396
94. Monath TP. Prospects for development of a vaccine against the West Nile virus. *Annals of the New York Academy of Sciences*. 2001; 951(1):1–12. <https://doi.org/10.1111/j.1749-6632.2001.tb02680.x> PMID: 11797767
95. Purtha WE, Myers N, Mitaksov V, Sitati E, Connolly J, Fremont DH, et al. Antigen-specific cytotoxic T lymphocytes protect against lethal West Nile virus encephalitis. *European journal of immunology*. 2007; 37(7):1845–54. <https://doi.org/10.1002/eji.200737192> PMID: 17559174
96. Larsen MV, Lelic A, Parsons R, Nielsen M, Hoof I, Lamberth K, et al. Identification of CD8+ T cell epitopes in the West Nile virus polyprotein by reverse-immunology using NetCTL. *PLoS one*. 2010; 5(9). <https://doi.org/10.1371/journal.pone.0012697> PMID: 20856867
97. Diez-Rivero CM, Reche PA. CD8 T cell epitope distribution in viruses reveals patterns of protein biosynthesis. *PLoS one*. 2012; 7(8). <https://doi.org/10.1371/journal.pone.0043674> PMID: 22952734
98. Cherryholmes GA, Stanton SE, Disis ML. Current methods of epitope identification for cancer vaccine design. *Vaccine*. 2015; 33(51):7408–14. <https://doi.org/10.1016/j.vaccine.2015.06.116> PMID: 26238725
99. Liu F, Wang X, Zheng M, Xiong F, Liu X, Zhou L, et al. Immunization with DNA prime-subunit protein boost strategy based on influenza H9N2 virus conserved matrix protein M1 and its epitope screening. *Scientific reports*. 2020; 10(1):1–12.
100. Pumchan A, Krobthong S, Roytrakul S, Sawatdichaikul O, Kondo H, Hirono I, et al. Novel Chimeric Multiepitope Vaccine for Streptococcosis Disease in Nile Tilapia (*Oreochromis niloticus* Linn.). *Scientific Reports*. 2020; 10(1):1–13.
101. Shrestha B, Diamond MS. Role of CD8+ T cells in control of West Nile virus infection. *Journal of virology*. 2004; 78(15):8312–21. <https://doi.org/10.1128/JVI.78.15.8312-8321.2004> PMID: 15254203
102. Ulbert S, Magnusson SE. Technologies for the development of West Nile virus vaccines. *Future microbiology*. 2014; 9(10):1221–32. <https://doi.org/10.2217/fmb.14.67> PMID: 25405890
103. Gordon S, Lawson L, Rabinowitz S, Crocker P, Morris L, Perry V. Antigen markers of macrophage differentiation in murine tissues. *Macrophage biology and activation*: Springer; 1992. p. 1–37.
104. Goverdhan M, Kulkarni A, Gupta A, Tupe C, Rodrigues J. Two-way cross-protection between West Nile and Japanese encephalitis viruses in bonnet macaques. *Acta virologica*. 1992; 36(3):277–83. PMID: 1360756
105. Halevy M, Akov Y, Ben-Nathan D, Kobiler D, Lachmi B, Lustig S. Loss of active neuroinvasiveness in attenuated strains of West Nile virus: pathogenicity in immunocompetent and SCID mice. *Archives of virology*. 1994; 137(3–4):355–70. <https://doi.org/10.1007/BF01309481> PMID: 7944955
106. Hall RA, Burgess GW, Kay BH, Clancy P. Monoclonal antibodies to Kunjin and Kokobero viruses. *Immunology and cell biology*. 1991; 69(1):47–9. <https://doi.org/10.1038/icb.1991.7> PMID: 1651286

107. Hase T, Summers P, Dubois D. Ultrastructural changes of mouse brain neurons infected with Japanese encephalitis virus. *International journal of experimental pathology*. 1990; 71(4):493. PMID: [2169298](#)
108. Oliphant T, Engle M, Nybakken GE, Doane C, Johnson S, Huang L, et al. Development of a humanized monoclonal antibody with therapeutic potential against West Nile virus. *Nature medicine*. 2005; 11(5):522–30. <https://doi.org/10.1038/nm1240> PMID: [15852016](#)
109. Diamond MS, Shrestha B, Marri A, Mahan D, Engle M. B cells and antibody play critical roles in the immediate defense of disseminated infection by West Nile encephalitis virus. *Journal of virology*. 2003; 77(4):2578–86. <https://doi.org/10.1128/jvi.77.4.2578-2586.2003> PMID: [12551996](#)
110. Wang T, Anderson JF, Magnarelli LA, Wong SJ, Koski RA, Fikrig E. Immunization of mice against West Nile virus with recombinant envelope protein. *The Journal of Immunology*. 2001; 167(9):5273–7. <https://doi.org/10.4049/jimmunol.167.9.5273> PMID: [11673542](#)
111. Tyler KL. Current developments in understanding of West Nile virus central nervous system disease. *Current opinion in neurology*. 2014; 27(3):342. <https://doi.org/10.1097/WCO.000000000000088> PMID: [24722324](#)
112. Shrestha B, mmiSamuel MA, Diamond MS. CD8+ T cells require perforin to clear West Nile virus from infected neurons. *Journal of virology*. 2006; 80(1):119–29. <https://doi.org/10.1128/JVI.80.1.119-129.2006> PMID: [16352536](#)
113. Camenga DL, Nathanson N, Cole GA. Cyclophosphamide-potentiated West Nile viral encephalitis: relative influence of cellular and humoral factors. *Journal of Infectious Diseases*. 1974; 130(6):634–41. <https://doi.org/10.1093/infdis/130.6.634> PMID: [4372273](#)
114. Wang Y, Lobigs M, Lee E, Müllbacher A. CD8+ T cells mediate recovery and immunopathology in West Nile virus encephalitis. *Journal of virology*. 2003; 77(24):13323–34. <https://doi.org/10.1128/jvi.77.24.13323-13334.2003> PMID: [14645588](#)
115. Kesson AM, Blanden RV, Mullbacher A. The primary in vivo murine cytotoxic T cell response to the flavivirus, West Nile. *Journal of general virology*. 1987; 68(7):2001–6. <https://doi.org/10.1099/0022-1317-68-7-2001> PMID: [3496425](#)
116. Liu Y, Blanden RV, Müllbacher A. Identification of cytolytic lymphocytes in West Nile virus-infected murine central nervous system. *Journal of general virology*. 1989; 70(3):565–73. <https://doi.org/10.1099/0022-1317-70-3-565> PMID: [2543751](#)
117. Nathanson N, Cole GA. Immunosuppression and experimental virus infection of the nervous system. *Advances in virus research*. 16: Elsevier; 1970. p. 397–448. [https://doi.org/10.1016/s0065-3527\(08\)60028-7](https://doi.org/10.1016/s0065-3527(08)60028-7) PMID: [4993221](#)
118. Lanteri MC, Heitman JW, Owen RE, Busch T, Geffer N, Kiely N, et al. Comprehensive analysis of West Nile virus-specific T cell responses in humans. *The Journal of infectious diseases*. 2008; 197(9):1296–306. <https://doi.org/10.1086/586898> PMID: [18422442](#)
119. Murray K, Baraniuk S, Resnick M, Arafat R, Kilborn C, Cain K, et al. Risk factors for encephalitis and death from West Nile virus infection. *Epidemiology & Infection*. 2006; 134(6):1325–32. <https://doi.org/10.1017/S0950268806006339> PMID: [16672108](#)
120. Pruitt AA. Central nervous system infections in cancer patients. *Neurologic clinics*. 1991; 9(4):867–88. PMID: [1758429](#)
121. Douglas M, Kesson A, King N. CTL recognition of west Nile virus-infected fibroblasts is cell cycle dependent and is associated with virus-induced increases in class I MHC antigen expression. *Immunology*. 1994; 82(4):561. PMID: [7835918](#)
122. Kulkarni A, Mullbacher A, Blanden R. In vitro T-cell proliferative response to the flavivirus, West Nile. *Viral immunology*. 1991; 4(2):73–82. <https://doi.org/10.1089/vim.1991.4.73> PMID: [1722099](#)
123. Mosmann TR, Coffman R. TH1 and TH2 cells: different patterns of lymphokine secretion lead to different functional properties. *Annual review of immunology*. 1989; 7(1):145–73. <https://doi.org/10.1146/annurev.iy.07.040189.001045> PMID: [2523712](#)
124. Reinherz EL, Schlossman SF. The differentiation and function of human T lymphocytes. *Cell*. 1980; 19(4):821–7. [https://doi.org/10.1016/0092-8674\(80\)90072-0](https://doi.org/10.1016/0092-8674(80)90072-0) PMID: [6991122](#)
125. Lafuente EM, Reche PA. Prediction of MHC-peptide binding: a systematic and comprehensive overview. *Current pharmaceutical design*. 2009; 15(28):3209–20. <https://doi.org/10.2174/138161209789105162> PMID: [19860671](#)
126. Mason P, Dalrymple J, Gentry M, McCown J, Hoke C, Burke DS, et al. Molecular characterization of a neutralizing domain of the Japanese encephalitis virus structural glycoprotein. *Journal of general virology*. 1989; 70(8):2037–49. <https://doi.org/10.1099/0022-1317-70-8-2037> PMID: [2549181](#)
127. Sitati EM, Diamond MS. CD4+ T-cell responses are required for clearance of West Nile virus from the central nervous system. *Journal of virology*. 2006; 80(24):12060–9. <https://doi.org/10.1128/JVI.01650-06> PMID: [17035323](#)

128. Kashyap M, Jaiswal V, Farooq U. Prediction and analysis of promiscuous T cell-epitopes derived from the vaccine candidate antigens of *Leishmania donovani* binding to MHC class-II alleles using in silico approach. *Infection, Genetics and Evolution*. 2017; 53:107–15. <https://doi.org/10.1016/j.meegid.2017.05.022> PMID: 28549876
129. Yang Y, Sun W, Guo J, Zhao G, Sun S, Yu H, et al. In silico design of a DNA-based HIV-1 multi-epitope vaccine for Chinese populations. *Human vaccines & immunotherapeutics*. 2015; 11(3):795–805. <https://doi.org/10.1080/21645515.2015.1012017> PMID: 25839222
130. Li X, Yang X, Jiang Y, Liu J. A novel HBV DNA vaccine based on T cell epitopes and its potential therapeutic effect in HBV transgenic mice. *International immunology*. 2005; 17(10):1293–302. <https://doi.org/10.1093/intimm/dxh305> PMID: 16113237
131. Ghaffari-Nazari H, Tavakkol-Afshari J, Jaafari MR, Tahaghoghi-Hajghorbani S, Masoumi E, Jalali SA. Improving multi-epitope long peptide vaccine potency by using a strategy that enhances CD4+ T help in BALB/c mice. *PloS one*. 2015; 10(11).
132. Wu C-Y, Monie A, Pang X, Hung C-F, Wu TC. Improving therapeutic HPV peptide-based vaccine potency by enhancing CD4+ T help and dendritic cell activation. *Journal of biomedical science*. 2010; 17(1):88. <https://doi.org/10.1186/1423-0127-17-88> PMID: 21092195
133. Mohan T, Verma P, Rao DN. Novel adjuvants & delivery vehicles for vaccines development: a road ahead. *The Indian journal of medical research*. 2013; 138(5):779. PMID: 24434331
134. Mohan T, Sharma C, Bhat AA, Rao D. Modulation of HIV peptide antigen specific cellular immune response by synthetic  $\alpha$ - and  $\beta$ -defensin peptides. *Vaccine*. 2013; 31(13):1707–16. <https://doi.org/10.1016/j.vaccine.2013.01.041> PMID: 23384751
135. Yang D, Biragyn A, Kwak LW, Oppenheim JJ. Mammalian defensins in immunity: more than just microbicidal. *Trends in immunology*. 2002; 23(6):291–6. [https://doi.org/10.1016/s1471-4906\(02\)02246-9](https://doi.org/10.1016/s1471-4906(02)02246-9) PMID: 12072367
136. Mohan T, Mitra D, Rao D. Nasal delivery of PLG microparticle encapsulated defensin peptides adjuvanted gp41 antigen confers strong and long-lasting immunoprotective response against HIV-1. *Immunologic research*. 2014; 58(1):139–53. <https://doi.org/10.1007/s12026-013-8428-5> PMID: 23666811
137. Li H, Ning P, Lin Z, Liang W, Kang K, He L, et al. Co-expression of the C-terminal domain of *Yersinia enterocolitica* invasin enhances the efficacy of classical swine-fever-vectored vaccine based on human adenovirus. *Journal of biosciences*. 2015; 40(1):79–90. <https://doi.org/10.1007/s12038-014-9495-z> PMID: 25740144
138. Ólafsdóttir G, Svansson V, Ingvarsson S, Marti E, Torsteinsdóttir S. In vitro analysis of expression vectors for DNA vaccination of horses: the effect of a Kozak sequence. *Acta Veterinaria Scandinavica*. 2008; 50(1):44. <https://doi.org/10.1186/1751-0147-50-44> PMID: 18983656
139. Kozak M. Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. *Nucleic acids research*. 1984; 12(2):857–72. <https://doi.org/10.1093/nar/12.2.857> PMID: 6694911
140. Ghaderi D, Zhang M, Hurtado-Ziola N, Varki AJB, Reviews GE. Production platforms for biotherapeutic glycoproteins. Occurrence, impact, and challenges of non-human sialylation. 2012; 28(1):147–76. <https://doi.org/10.5661/bger-28-147> PMID: 22616486
141. Durocher Y, Butler MJCoib. Expression systems for therapeutic glycoprotein production. 2009; 20(6):700–7.
142. Dai S, Wang H, Deng FJJoiS. Advances and challenges in enveloped virus-like particle (VLP)-based vaccines. 2018; 2(2).
143. Xiang K, Ying G, Yan Z, Shanshan Y, Lei Z, Hongjun L, et al. Progress on adenovirus-vectored universal influenza vaccines. 2015; 11(5):1209–22.
144. Kallel H, Kamen AAJBj. Large-scale adenovirus and poxvirus-vectored vaccine manufacturing to enable clinical trials. 2015; 10(5):741–7.
145. Farnós O, Gelaye E, Trabelsi K, Bernier A, Subramani K, Kallel H, et al. Establishing a Robust Manufacturing Platform for Recombinant Veterinary Vaccines: An Adenovirus-Vector Vaccine to Control Newcastle Disease Virus Infections of Poultry in Sub-Saharan Africa. 2020; 8(2):338.
146. Cai X, Bai H, Zhang X. Vaccines And Advanced Vaccines:-A landscape for advanced vaccine technology against infectious disease, COVID-19 and tumor. 2020.
147. Guleria V, Jaiswal VJG. Comparative transcriptome analysis of different stages of *Plasmodium falciparum* to explore vaccine and drug candidates. 2020; 112(1):796–804.
148. Milligan I, Gibani M, Sewell R, Clutterbuck E, Campbell D, Plested E, et al. Immunogenicity and reactogenicity of novel adenovirus type 26 and modified vaccinia Ankara-vectored Ebola vaccines: A randomized clinical trial. 2016; 315(15).