

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/345342684>

# Development of multi-epitope subunit vaccine for protection against the norovirus' infections based on computational vaccinology

Article in Journal of biomolecular Structure & Dynamics · November 2020

DOI: 10.1080/07391102.2020.1845799

CITATIONS

7

READS

406

15 authors, including:



**Syed Shujait Ali**

University of Swat

86 PUBLICATIONS 922 CITATIONS

[SEE PROFILE](#)



**Huma Farooque Hashmi**

Quaid-i-Azam University

5 PUBLICATIONS 33 CITATIONS

[SEE PROFILE](#)



**Shahzeb Khan**

Shanghai Jiao Tong University

10 PUBLICATIONS 121 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Network Pharmacology [View project](#)



Biological databases development [View project](#)



## Development of multi-epitope subunit vaccine for protection against the norovirus' infections based on computational vaccinology

Irfan Ahmad<sup>a#</sup>, Syed Shujait Ali<sup>a#</sup>, Bisma Zafar<sup>b</sup>, Huma Farooque Hashmi<sup>c</sup>, Ismail Shah<sup>a</sup>, Shahzeb Khan<sup>a</sup>, Muhammad Suleman<sup>a</sup>, Mazhar Khan<sup>d</sup>, Saif Ullah<sup>a</sup>, Shahid Ali<sup>a</sup>, Jafar Khan<sup>a</sup>, Mohammad Ali<sup>a</sup>, Abbas Khan<sup>e</sup> and Dong-Qing Wei<sup>e,f,g</sup>

**Q1** <sup>a</sup>Centre for Biotechnology and Microbiology, University of Swat, Swat, Pakistan; <sup>b</sup>Department of Biotechnology, University of Okara, Punjab, Pakistan; <sup>c</sup>School of Life Sciences, Shandong University, Shandong, P.R. China; <sup>d</sup>The CAS Key Laboratory of Innate Immunity and Chronic Diseases, Hefei National Laboratory for Physical Sciences at Microscale, School of Life Sciences, CAS Center for Excellence in Molecular Cell Science, University of Science and Technology of China (USTC), Collaborative Innovation Center of Genetics and Development, Hefei, P.R. China; <sup>e</sup>Department of Bioinformatics and Biological Statistics, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, P.R. China; <sup>f</sup>State Key Laboratory of Microbial Metabolism, Shanghai-Islamabad-Belgrade Joint Innovation Center on Antibacterial Resistances, Joint Laboratory of International Cooperation in Metabolic and Developmental Sciences, Ministry of Education and School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, P.R. China; <sup>g</sup>Peng Cheng Laboratory, Shenzhen, P.R. China

Communicated by Ramaswamy H. Sarma

**Q2**

### ABSTRACT

Human Norovirus belongs to a family *Calciviridae*, and was identified in the outbreak of gastroenteritis in Norwalk, due to its seasonal prevalence known as “winter vomiting disease.” Treatment of Norovirus infection is still mysterious because there is no effective antiviral drugs or vaccine developed to protect against the infection, to eradicate the infection an effective vaccine should be developed. In this study, capsid protein (A7YK10), small protein (A7YK11), and polyprotein (A7YK09) were utilized. These proteins were subjected to B and T cell epitopes prediction by using reliable immunoinformatics tools. The antigenic and non-allergenic epitopes were selected for the subunit vaccine, which can activate cellular and humoral immune responses. Linkers joined these epitopes together. The vaccine structure was modelled and validated by using Errat, ProSA, and rampage servers. The modelled vaccine was docked with TLR-7. The stability of the docked complex was evaluated by MD simulation. To apply the concept in a wet lab, the reverse translated vaccine sequence was cloned in pET28a (+). The vaccine developed in this study requires experimental validation to ensure its effectiveness against the disease.

### ARTICLE HISTORY

Received 8 April 2020  
Accepted 29 October 2020

### KEYWORDS




Norovirus; epitopes;  
docking; MD simulation;  
in-silico cloning

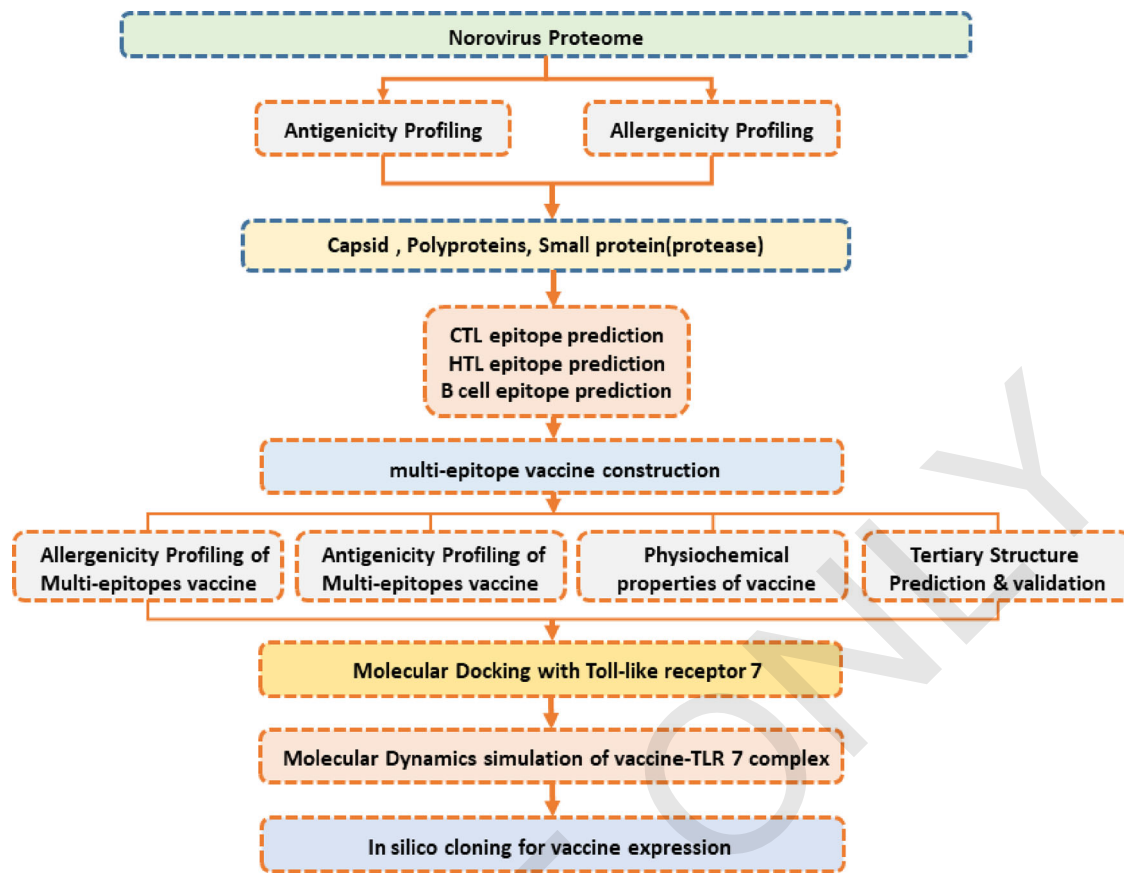
## Introduction

Norovirus (NVs), also known as “small round-structured virus,” and this RNA single-stranded virus is placed within family *Calciviridae*. NVs are responsible for the spread of non-bacterial human gastroenteritis (Lew et al., 1994). The Human norovirus (HuNV) was first identified in stool specimen, during gastroenteritis outbreak in Norwalk and was coined as Norwalk virus. More than a thousand strains of NVs are isolated, which are genetically and serologically different. The infected person has an abdominal cramp, stomach pain, diarrhea, vomiting, and nausea with mild pyrexia (Robilotti et al., 2015). The consumption of contaminated food and water is deemed essential for the development and spread of disease (Mead et al., 1999; White et al., 2016), and globally 20% of all diarrheal diseases are caused by HuNV, and nearly 21200 victims succumbed to death annually (Ahmed et al., 2014; Pires et al., 2015). HuNV interferes with interferon type I & III by influencing MHC-I expression and causing rapid infection. MHC-I plays a key role in providing immunity against viruses.

In this process, proteasome-mediated degraded peptides are presented to the CD8+ T-cells for evoking immune reactions (Mattison et al., 2018). The genome of HuNV is 7.5 kb, which consists of three open reading frames (ORF's), ORF1, ORF2, and ORF3. These ORF's (ORF1, ORF2, and ORF3) codes for a nonstructural protein, VP1 major capsid protein, and VP2 minor capsid proteins, respectively (McFadden et al., 2011).

In clinical samples, an electron microscope (EM) (Caul & Appleton, 1982) is used as a diagnostic tool for norovirus identification. ELISA (enzyme-linked immunosorbent assay) and molecular techniques are accessible for the diagnostic purpose of pathogens, including norovirus. EM is found in the well-equipped lab, and it is used to look for the pathogenic particles in the feces. RT-PCR (Ando et al., 1995) shows sensitivity for identification, and it also assists in the comprehension of these viruses molecular diversity (Fankhauser et al., 1998; Green et al., 1993; Maguire et al., 1999; Norcott et al., 1994; Vinje & Koopmans, 1996). Genomic characterization and molecular diversity are assessed by (HMA) heteroduplex mobility assay of various viruses which include;

**CONTACT** Abbas Khan  [abbaskhan@sjtu.edu.cn](mailto:abbaskhan@sjtu.edu.cn); Dong-Qing Wei  [dqwei@sjtu.edu.cn](mailto:dqwei@sjtu.edu.cn)  Department of Bioinformatics and Biological Statistics, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, 200240, P.R. China  
#contributed equally.



**Figure 1.** Schematic representation of the steps involved in epitope-based vaccine designing. A multi-step approach was used to construct the final vaccine candidate. Finally, validation through MD simulation and in silico expression was achieved.

Norovirus (Mattick et al., 2000), measles virus (Kreis & Whistler, 1997), polioviruses (Chezzi & Schoub, 1996), hepatitis C virus (Weiner et al., 1998) and polioviruses (Chezzi & Schoub, 1996). Cross challenge studies and IEM (immune electron microscopy) studies (Kapikian et al., 1972) was previously utilized for NV antigenic diversity assessment, before the development of rNV (recombinant-NV) capsid protein.

Non-bacterial gastroenteritis is still a great challenge, and there is no effective licensed vaccine available for its treatment (Mattison et al., 2018). Researchers are trying their best to launch an effective vaccine against norovirus. However, their investigations are either in clinical trials or in pre-clinical stages (Cortes-Penfield et al., 2017). These investigations may result in two norovirus vaccines in the future, bivalent GI.1/GII.4 intramuscular VLP vaccines (in phase II b clinical trial), and monovalent GI.1 oral pill recombinant adenovirus vaccine (phase I trials) (Mattison et al., 2018). NoV infection is equally pervasive in developing as well as developed countries. Children and elders are severely affected by these infections. Effective vaccines and drug designing would be instrumental in controlling the higher mortality rates caused by these infections (Pires et al., 2015). Vaccines evoke innate and cellular immune responses to develop antibodies and memory cells, which may provide long-lasting protection from specific serotypes (Pulendran & Ahmed, 2011). In-silico study based on immunoinformatics approaches was applied to pinpoint effective epitopes or hits as a potential candidate for vaccine or drug designing (Adhikari et al., 2018; Dash et al., 2017; Kalliamurthi et al., 2018; Khan, Junaid, et al.,

2018; Khan, Khan, Rehman, et al., 2019). Computational approaches are beneficial to predict the antigen without culturing the pathogenic strain experimentally (Agrawal & Raghava, 2018; Ali et al., 2019; Khan, Khan, Ali, et al., 2019; Rappuoli, 2000; Sette & Fikes, 2003).

In this study, a computational analysis was performed to predict the epitopes based effective vaccine against the NoV. T-cell, B-cell, and HTL epitopes were predicted and analyzed using defined criteria for selecting the potential epitopes for the final vaccine construct. Further implementation of molecular docking with TLR receptors, molecular dynamics simulation, and codon optimization for expression confirmed the potential of the final vaccine construct. Thus, this study provides a way towards the development of a potential vaccine candidate against the Norovirus (Figure 1).

## Results

### Protein collection

The amino acid sequences of capsid protein (UniProt id: A7YK10), polyprotein (UniProt id: A7YK09), and small protein (A7YK11) of *Norovirus* as well as Beta-defensin 3 (Q5U7J2), an adjuvant was retrieved from UniProtKB in FASTA format. These proteins were antigenic in nature based on antigenicity scores of 0.48, 0.52, and 0.49, respectively, as calculated by the VaxiJen server and were selected for the designing of a multi-epitope vaccine by immunoinformatics approach.

**Table 1.** Selected CTL epitopes. All the epitopes are based on their scores.

Proteins	Position	Peptide	Affinity	affinity rescale	Cleavage	TAP	Combine score	MHC binding	Epitope
A7YK10	469–477	QSDALLIRY	0.7635	3.2418	0.9510	2.8490	3.5269	Yes	
	91–100	YLAHLSAMY	0.5251	2.2294	0.8881	2.7960	2.5024	Yes	
A7YK09	1044–1053	TSSGDFLKY	0.5906	2.5076	0.9717	2.9870	2.8027	Yes	
	1513–1522	MQESEFSFY	0.5161	2.1913	0.5334	2.9500	2.4188	Yes	
A7YK11	122–131	AVDWSGTRY	0.5730	2.4330	0.9756	3.1630	2.7375	Yes	
	140–149	FSGGFTPSY	0.3257	1.3828	0.9690	2.6770	1.6620	Yes	

**Table 2.** HTL selected epitopes. The epitopes are based on percentile rank.

Proteins	Allele	Start	End	Peptide	Method	Percentile rank
Capsid protein	HLA-DRB1*01:02	145	159	PHVMCDVRALEPIQL	sturniolo	0.02
	HLA-DRB1*01:03	373	387	KVYASLAAAAPLDLV	NetMHCIIpan	0.20
	HLA-DRB1*01:02	117	131	FTAGKVVVALVPPYF	sturniolo	0.44
Polyprotein	HLA-DRB1*01:02	1	15	MRMATPSSASSVRNT	sturniolo	0.12
	HLA-DRB1*01:01	1631	1645	GEKYRTVASRSVSKE	Consensus (comb.lib./simm/nn)	0.19
	HLA-DRB1*01:03	816	830	ITSILQAAGTAFSIY	NetMHCIIpan	0.25
	HLA-DRB1*01:02	617	631	CRRIDFLVYAESPVV	sturniolo	0.32
	HLA-DRB1*01:03	1089	1103	LALAVRMGSQAARI	NetMHCIIpan	0.51
	HLA-DRB1*01:02	193	207	PLDPAELRKCVGMTV	sturniolo	0.59

**Table 3.** Selected HTL epitopes analysis for interferon production.

Serial No	Sequence	Method	Result	Score
1	PHVMCDVRALEPIQL	MERCI	NEGATIVE	1
2	GEKYRTVASRSVSKE	MERCI	NEGATIVE	1
3	ITSILQAAGTAFSIY	MERCI	NEGATIVE	1
4	CRRIDFLVYAESPVV	MERCI	NEGATIVE	1
5	KVYASLAAAAPLDLV	SVM	POSITIVE	0.32853096
6	FTAGKVVVALVPPYF	SVM	NEGATIVE	−0.37555421
7	MRMATPSSASSVRNT	SVM	POSITIVE	0.98957435
8	LALAVRMGSQAARI	SVM	POSITIVE	0.087092709
9	PLDPAELRKCVGMTV	SVM	NEGATIVE	−0.024602308

**Table 4.** Predicted Linear B cell epitopes with their respective scores.

Rank	Sequence	Score	Start position
1	GPGGEKYYRTVASRSVSKEGP	<b>0.91</b>	222
2	VMCDVRALEPIQLGPGPGHV	<b>0.89</b>	127
3	GNSISTGPGPGGGGLMGIIG	<b>0.87</b>	334
4	LEPIQLPLLDGPGPMRMAT	<b>0.86</b>	190
5	PGPGPGVMCDVRALEPIQLP	<b>0.86</b>	159
6	PGPGITSILQAAGTAFSIYG	<b>0.84</b>	241

### B-cell epitope prediction for norovirus

B-cell epitope prediction was performed on ABCpreds server, and six epitopes of 20-mer in length with a score higher than 0.83 were selected for further analysis. Conformational B-cell epitopes were identified on Discotope 2.0, and out of 359 residues, 50 B-cell residues were predicted. The predicted B-cell Linear and discontinuous epitopes are given in Tables 4 and 5.

### Final multi-epitope vaccine construct

The final multi-epitope vaccine construct was composed of 6 CTL and B-cell epitopes while 9 HTL epitopes were selected based on high binding affinity scores. The topological representation of the final vaccine candidate is given in Figure 2. AAY linkers were used to combine CTL epitopes, and GPGPG linkers joined HTL epitopes. In contrast, EAAAK linker was used for attachment of adjuvant to the N-terminal of vaccine, which amplifies its function.

### Prediction and validation of tertiary structure

The 3D structure of the vaccine was predicted by the Robetta server. Five models were generated, and after evaluation, model three Figure 3(A) was selected for further analysis. The epitopes used to construct the final structure of the multi-epitopes vaccine is given in Figure 3(B). Each epitope such as CTL, HTL and B-cell epitopes are coloured differently and shown as ribbon. For better representation, a surface shade has been applied.

### CTL (cytotoxic T lymphocytes) epitopes prediction

NetCTL1.2 server predicted a total of 51 CTL epitopes of 9-mer in length. In these epitopes, six non-allergenic epitopes (Table 1) were selected for vaccine designing based on a high binding affinity score. Based on the predicted scores, two epitopes were selected from A7YK09, A7YK10, and A7YK11 each.

### HTL (helper T lymphocytes) prediction

MHC-II prediction module of IEDB was used for Helper T Lymphocytes (HTL) epitopes prediction for HLA-DRB1\*01:01, HLA-DRB1\*01:02, HLA-DRB1\*01:03, HLA-DRB1\*01:04, and HLA-DRB1\*01:05 Human alleles. 9 HTL epitopes with the highest binding affinity were selected. The selected epitopes are situated at position 145–159, 373–387, 117–131 (capsid protein), 1–15, 1631–1645, 816–830, 617–631, 1089–1103 and 193–207 (polyproteins). (Table 2 represent selected HTL epitopes).

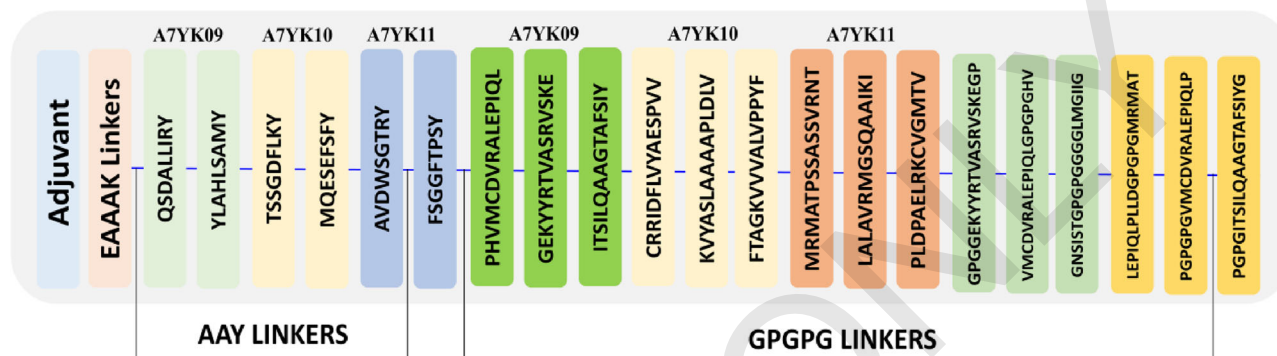
### Interferon production

Online server IFN-epitope was used to identify the HTL epitope with the potential to induce the release of IFN-gamma from T (CD4+) cell. This analysis resulted in three epitopes (out of nine HTL epitopes) with the ability to produce T cells for interferon production (Table 3).



**Table 5.** Predicted discontinuous B-cell epitopes residues with their contact number and Discotope score.

S.NO	Residues	Contact number	Discotope score
1	LYS, TYR, PRO, GLN, ASN, GLY, PRO, GLY, GLY, GLU	3, 6, 4, 0, 4, 1, 0, 12, 2, 4	-2.885, -2.029, -3.589 -3.886, -1.879, 1.667 2.652, 3.693, 2.327, 0.024
2	LYS, PRO, TYR, GLY, PRO, GLY, PRO, GLY, HIS, GLY	8, 9, 11, 13, 3, 13, 5, 0, 0, 6	-0.356, -2.670, -2.082, 1.008, 2.997, 4.536, 5.611, 5.290, 5.382, 5.370
3	GLU, LYS, TYR, TYR, ARG, THR, ALA, LYS, PRO, GLY	13, 9, 18, 32, 6, 3, 7, 0, 8, 4	5.690, 5.180, 4.948, 2.551, 1.623, 0.693, -2.447, -3.572, -2.696, -2.886
4	GLY, ILE, ILE, GLY, ASN, SER, ILE, GLY, PRO, GLY	4, 35, 7, 30, 3, 16, 3, 4, 13, 36	-1.302, 0.705, 2.215, 3.458, 5.151, 4.587, 2.787, 2.029, 0.101, 0.667
5	PRO, GLY, PRO, GLY, PRO, GLY, ILE, ASN, GLY, GLY	11, 28, 7, 0, 7, 12, 6, 0, 16, 15	-0.292, 0.722, 1.126, 1.693, 1.776, 1.315, -0.535, -2.624, 0.504, -2.062



**Figure 2.** Final vaccine construct. Six different CTL epitopes, while nine HTL epitopes from three different proteins and six B-cell epitopes were combined to construct the multi-epitope subunit vaccine using linkers. An adjuvant to the N-terminal has also been added. The blue line shows the linkers AAY and GPGPG.

The selected model was validated by ProSA-web, (Figure 4(A)), RAMPAGE (Figure 4(B)) and ERRAT. Modeled protein Ramachandran plot analysis revealed that most favored regions, additionally allowed regions, generously allowed region and disallowed region, contain 70.8%, 22.3%, 6.8%, and 0.4% of residues, respectively. ProSA-web and ERRAT were used for the evaluation of potential and quality errors in the 3D crude model. Modeled protein overall quality factor was found 70.88% utilizing ERRAT. ProSA-web is being used for prediction of Z-score prediction, which is reported to be  $-4.85$ .

### Antigenicity, allergenicity and physiochemical parameter prediction of final vaccine construct

The antigenicity of the final vaccine construct was predicted on Vaxijen and ANTIGENpro servers by selecting the bacteria model at 0.4% threshold. The antigenicity scores predicted by Vaxijen and ANTIGENpro are 0.94% and 0.7124%, respectively, which indicates the antigenic nature of the final vaccine. For allergenicity, AlgPred was used and  $-0.537$  score indicating the non-allergen nature of the vaccine using the default threshold of  $-0.40$ . The physiochemical parameters, including molecular weight (MW) and theoretical isoelectric point value (PI) of vaccine, were found 50.64 kDa and 8.98, respectively, as predicted by ProtParam. The PI value (8.98) suggesting the basic nature of the vaccine. Moreover, the half-life in mammalian reticulocytes, yeast, and *E. Coli* was found 30 h (in vitro); 20 h (in vivo), and 10 h (in vivo), respectively. The instability index score of 31.05 refers to the stable nature of the protein. The GRAVY (Grand average of

hydropathicity) and the aliphatic index were found 73.44 and  $-0.060$ , respectively.

## Prediction of secondary structure

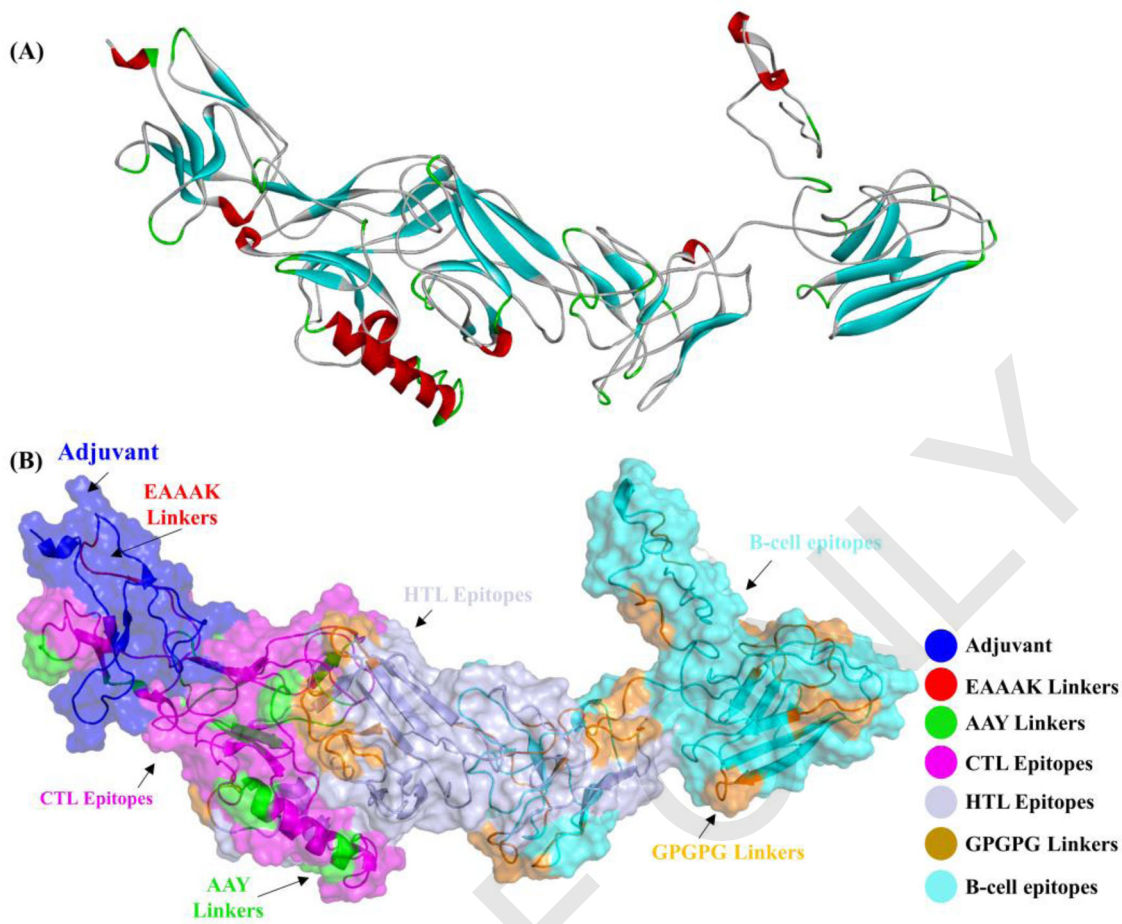
Secondary structure for vaccine predicted by PSIPRED program suggests the presence of 20.83% alpha-helix, 22.59% beta-strand, and 56.58% coil as shown in [Figure 5](#).

### Subunit vaccine molecular docking with an immune receptor (TLR-7)

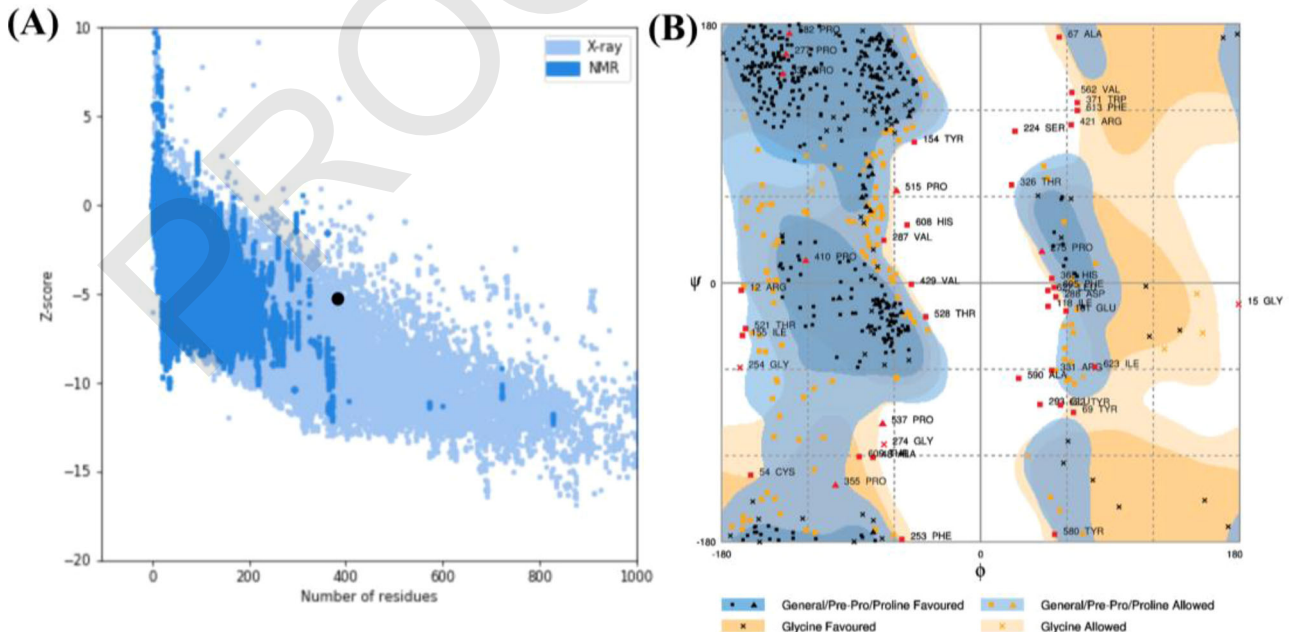
For docking of TLR-7 was docked with the multi-epitopes vaccine, using an online server ZDOCK. Overall, ten complexes were generated, and the most suitable vaccine-TLR complex was selected based on correct conformation and binding ([Figure 6](#)). The PDBsum server reported 79 interface residues with one salt bridge, and eleven hydrogen bonds were reported between the vaccine and TLR-7.

### MD simulation of immune receptor-vaccine complex

Molecular dynamics (MD) simulation was performed to check the stability and fluctuation of vaccine construct and TLR-7 complex. The computed RMSD and RMSF for the vaccine (protein) and its side chain, as well as their graph, are shown in [Figure 7](#). The RMSD and RMSF of protein and side-chain residues respectively were checked at 100 ns simulation time to estimate the stability of the system. Overall fluctuation (RMSD) rate for the simulated system was found 4–5 Å (TLR-7), and RMSF (residual fluctuation) for maximum residues



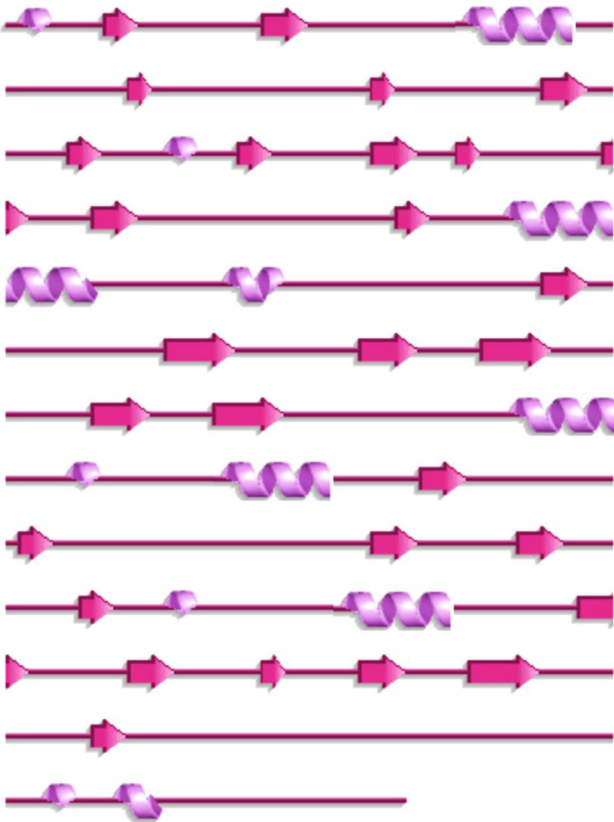
**Figure 3.** Predicted structure of the vaccine construct includes CTL, HTL and B-cell epitopes. (A) Showing the Ribbon representation of the final vaccine structure (B) Showing the final structure of multi-epitopes vaccine construct showing Adjuvant, linkers, CTL, HTL and B-cell epitopes coloured differently.



**Figure 4.** Validation of the 3D final vaccine model. (A) PROSA showing Z-score ( $-4.85$ ) for 3D structure validation (B) In Ramachandran analysis residues were allocated; the most favored region 70.8%, 22.3%, 6.8%, and 0.4%.

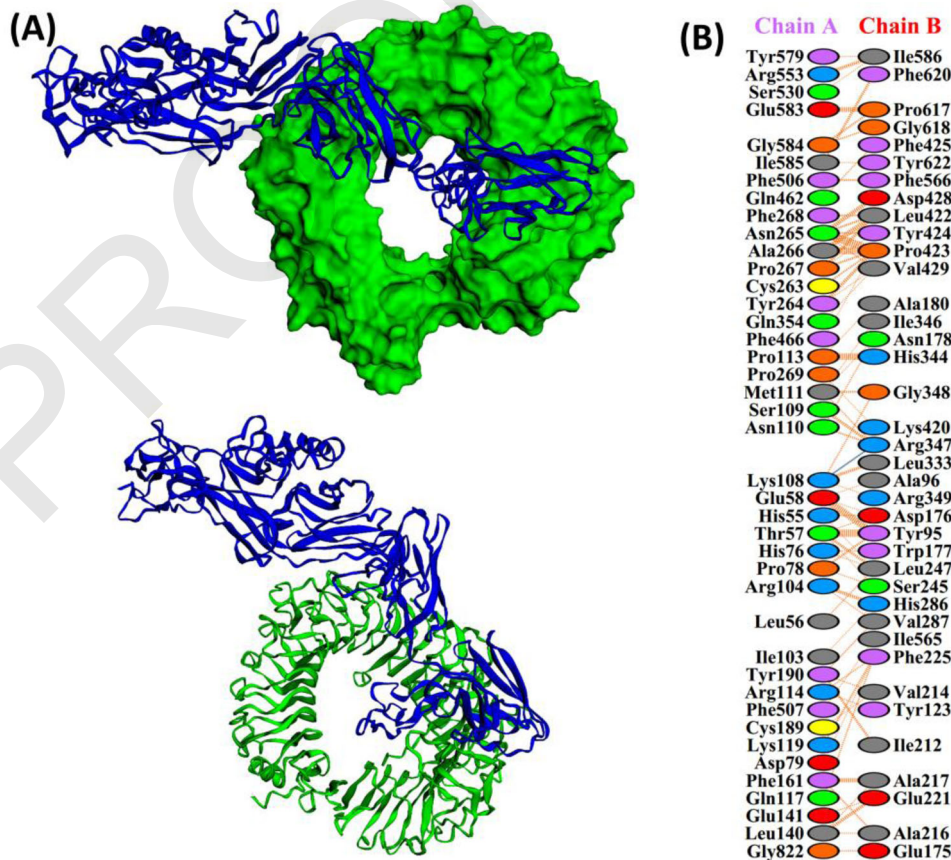


COLOR  
Online /  
B&W in  
Print



**Figure 5.** Represent the secondary structure of the final vaccine construct in which alpha helix, beta-strands, and coils were identified (20.83% alpha-helix, 22.59% beta-strand, and 56.58% coil).

COLOR  
Online /  
B&W in  
Print



**Figure 6.** (A) TLR-7 (PDB ID: 5GMF) and vaccine docked complex. Magenta colour represents the vaccine while the green and blue (surface) colour represents the receptor. (B) Showing the interaction pattern of the TLR-7 and vaccine construct.

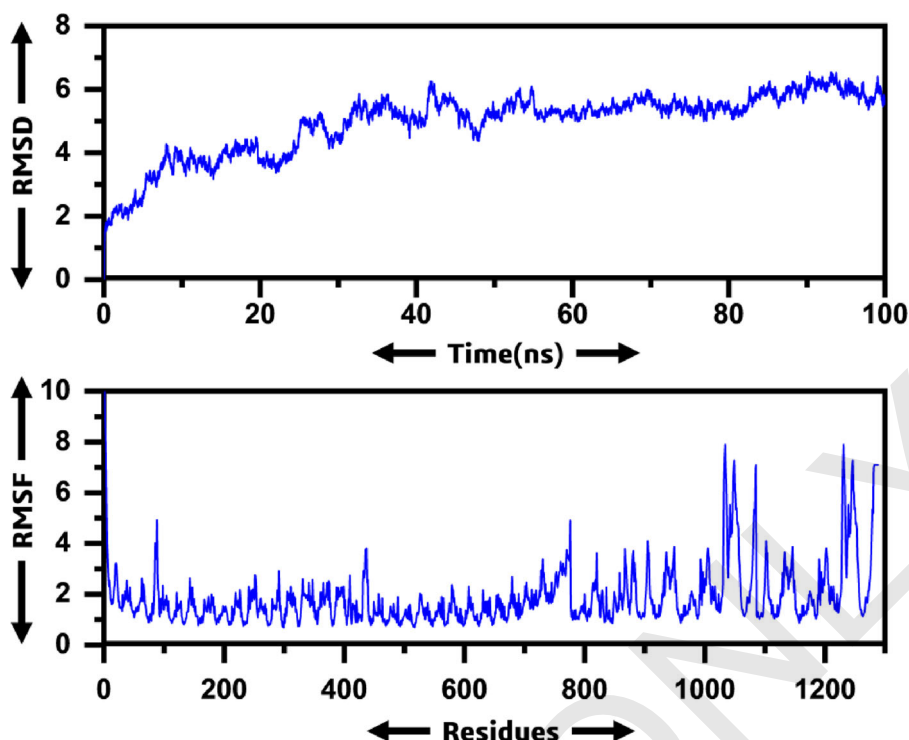
were found in acceptable range while some residues exhibit higher fluctuation.

### Codon optimization and in silico cloning

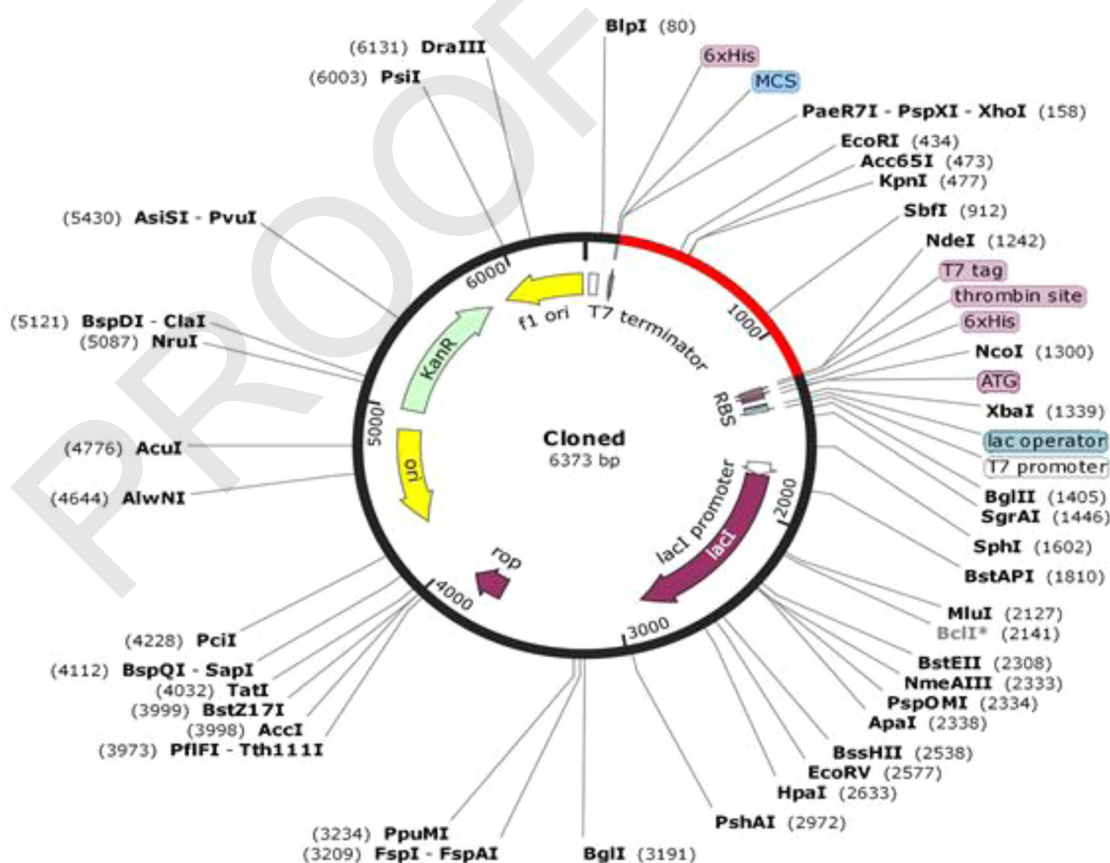
To assure the maximal expression of the protein, the vaccine codon was optimized in *E. coli* (strain ATCC 27325/DSM5911/W3110/K12) using Java Codon Adoption Tool. The optimized length of the codon sequence is 1077 nucleotides. The average GC content was found 58.54% (optimum range 30%–70%), and CAI (codon adaption index) was 0.9688, which indicates possibilities of good expression in host *E. coli*. Finally, using restriction enzymes restriction clone was formed, and adapted codon sequence was inserted in pET28a (+) plasmid. The designed construct is shown in Figure 8.

### Discussion

Norovirus proteins such as capsid, polyprotein, and protease were found antigenic and are vital for infection and replication within the host. Therefore, they are considered very important for subunit vaccine analysis. Immunization or vaccination is a universally recognized method to eradicate or control the infection. The advancement in computational approaches and their applications to biological research ushered a new era of subunit vaccination designing in which



**Figure 7.** Molecular dynamics simulation of the receptor-vaccine complex. The first graph is showing the RMSD of the complex (X-axis = Time in ns and Y-axis = RMSD) while the second graph is showing the RMSF of the complex (X-axis = Time in residue and Y-axis = RMSF).



**Figure 8.** In silico restriction cloning of the final vaccine construct into pET28a (+) expression vector where Red part representing the vaccine insert and black circle showing the vector.



the most accurate or exact antigenic part is identified and used as an immunization tool instead of the whole pathogen. The use of computational techniques to design small molecule inhibitors and vaccine candidate is the most frequently and time-saving practice (Khan, Ali, Khan, et al., 2020; Khan, Ali, Wang, et al., 2020; Khan, Junaid, et al., 2018; Khan, Saleem, et al., 2018). The fast accumulation of vast genomic and proteomic data of pathogens, including norovirus, is helpful in the development of epitope-based effective vaccines to control and eradicate the diseases caused by various pathogens. Computational approaches predicted epitopes (CTL & HTL) based on norovirus proteins and validation scores of these epitopes suggest their use for subunit vaccine construction. MHC (major histocompatibility complex) is of different types; MHC-I carries a peptide of 9-mer to the surface of the cell and acts like an impulsive signal for cytotoxic T cell, which leads to cell destruction by activating a complementary immune cascade. MHC-II molecules present peptide of 15-mer to Helper T lymphocytes. Our final subunit vaccine is composed of high-affinity CTL and HTL epitopes to elicit immunity. The allergenicity and antigenicity values of the vaccine were calculated, and these values indicate the non-allergen nature of the vaccine and are capable of provoking immune response due to antigenic nature.

In addition to these epitopes, B cell linear epitopes were also predicted, which help in B-cell maturation, in order to produce antibodies. Physicochemical properties of vaccines such as molecular weight, theoretical PI, aliphatic array, and thermal stability, were calculated. The molecular weight of the vaccine was 50.64 kDa, which is a suitable range for the subunit vaccine; the theoretical PI score is 8.98, which indicates that the vaccine is basic in nature. The aliphatic array suggests that vaccines have aliphatic side chains, and instability index endorses the thermally stable nature of the vaccine. For prediction and analysis of the secondary structure of the vaccine, PSIPRED V3.3 was used, which indicates the presence of 20.83% alpha-helix, 22.59% beta-strand, and 56.58% coil. Besides, the 3D structure obtained by homology modeling work comprises of adequate information on the spatial arrangement of such essential protein residues as well as useful guidance in the study of protein normal function, dynamics, and interaction with ligand as well as other proteins. To pinpoint the error in the final 3D structure of vaccine, different structural validation tools were used to detect errors. From the main Ramachandran plot, it was found that the overall model is satisfactory because most residues were found in the most-favour region while few were present in the disordered region.

Furthermore, the vaccine was docked with TLR-7 to understand the immune response towards the final vaccine structure. Energy minimization was conducted to minimize the potential energy of the entire system for the overall conformational stability of the docked vaccine protein-TLR-7. Energy minimization repairs the structure's unnecessary topology by ditching certain protein atoms and thus forms a more relatively stable structure with adequate stereochemistry is thus formed.

To obtain maximum expression (transcription and translation) of vaccine protein in the host (*E. coli* strain k-12), codon optimization was accomplished by CAI (codon adaptation index). The solubility of overexpressed recombinant protein in the host (*E. coli*) is one of the crucial requirements of many biochemical processes. Our vaccine protein shows a suitable proportion of solubility in the host. The essential goals of many mechanical and biomedical applications are to strengthen the protein. The newly designed vaccine was very effective and showing good immunogenic response in animals. However, when these vaccines were applied to humans, the response was not similar due to the complexity of human immunopathology. In this study, modern immunoinformatics approaches were incorporated to develop new thermostable, cheap and effective subunit vaccines. These vaccines are safe and immunogenic to exploit the immune system to provide protection from norovirus infection.

Azim et al. (2019) also published their data using immunoinformatic approaches for designing multi-epitope vaccine against human norovirus. In their analysis they selected two proteins for multi-epitope vaccine designing whereas in this study whole proteome was used. The antigenicity score of their final vaccine construct was 0.60 as estimated by Vaxijen server and in this study the antigenicity score was 0.94 (Vaxijen). The comparison of both scores suggest that the vaccine construct used in this study is more immunogenic. The instability index value 31.05 indicates the stable nature of vaccine. Azim et al. reported 21.33% alpha helix, 7.7% sheet and 70.9% coils is secondary structure, whereas in this study 20.83% alpha-helix, 22.59% beta-strand, and 56.58% coil were present. Higher %age of coil suggest that it would be denatured easily. The half life is very important for expression and purification of proteins, in Azim et al the reported half-life is 1 h (in vitro) within mammalian reticulocytes, whereas the half-life in this study was 30 h (in vitro). We performed MD simulation up to 100 ns to check the stability of the complex. In this study we used TLR 7 (Azim et al used TLR 8). The CAI value in this study is 0.96 higher than reported in Azim et al. This probably suggest the vaccine construct is robust compared the construct used in Azim et al.

## Conclusion

In this study, the main focus was to apply, *in-silico* approaches to design an effective multi-epitope vaccine against norovirus based on three proteins due to their antigenic nature. To design a vaccine, B and T cell epitopes were predicted, which is due to the presentation of pathogen epitope by MHC-I and II. Suitable linkers were used to fuse these epitopes. Vaccine tertiary structure was predicted and validated to ensure the functionality of the vaccine. The physicochemical properties such as antigenicity, allergenicity, and stability were computed. The vaccine was docked with TLR-7 to check the vaccine affinity towards the receptor.

Finally, for codon optimization, the protein was reversely translated which ensure maximum expression of the vaccine in the host (*E. coli*). A wet lab experimental validation is

needed to assure the activity of constructed vaccine. This study can help in controlling norovirus infection.

## Methodology

### Collection of norovirus proteins for vaccine preparation

The three proteins, capsid, polyproteins, and small basic proteins, were collected from UniProtKB (<https://www.uniprot.org>) (Consortium, 2018). These proteins were found antigenic through the VaxiJen server (<http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html>), and they were considered for vaccine designing (Doytchinova & Flower, 2007a, 2007b, 2008).

### Prediction of the (CTL) epitope

NetCTL 1.2 is an online server (<http://www.cbs.dtu.dk/services/NetCTL/>) in which CTL epitopes were predicted against these proteins (Larsen et al., 2007). Prediction of these CTL epitopes is based upon three essential parameters, including peptide attached to MHC-proteasomal C-terminal degradation activity, TAP (Transporter Associated with Antigen Processing) delivery accuracy and I. Artificial neural network was used to predict the attachment of peptide to MHC-I and proteasomal C-terminal degradation while TAP deliverance score was predicted by the weight matrix. For CTL epitope identification threshold was set as 0.75.

### Prediction of helper T-cell epitopes

IEDB (Immune Epitope Database) is an online server (<http://tools.iedb.org/mhcii/>), which is employed for prediction of Helper T cells lymphocytes of 15-mer length for five human alleles (HLA-DRB1\*01:01, HLA-DRB1\*01:02, HLA-DRB1\*01:03, HLA-DRB1\*01:04, HLA-DRB1\*01:05) (Nielsen & Lund, 2009; Wang et al., 2008, 2010). This server speculates epitopes based on receptor affinity, which is figured out from IC50 value (binding score) given to each epitope. Epitopes with higher binding affinity mostly have IC50 score <50 nM. The IC50 value determines epitopes affinity; value such is <500 nM represent moderate affinity while <5000 nM instantly shows epitopes having low affinity. The affinity of epitopes is inversely related to the value of percentile rank.

### Prediction of B cell lymphocytes

ABCpred is an online server (<http://crdd.osdd.net/raghava/abcpred/>) employed for the prediction of linear B-cell epitopes. The ABCpred predicts B-cell epitopes (linear) with a precision of 75% (0.49 Sensitivity and 0.75 specificity). Conformational B-cell epitopes prediction is carried out by DiscoTope 2.0 an online server (<http://www.cbs.dtu.dk/services/DiscoTope/>) (Kringelum et al., 2012). The process of computation in this method is based on log-odds ratios (ratio among amino acid composition in epitopes and residues of non-epitopes) and surface accessibility. The sensitivity and

specificity are found 0.47 and 0.75, respectively, on the default threshold (3.7).

### Multi-epitopes vaccine sequence construction

High scoring epitopes, including other parameters, filtered out the final epitopes from CTL and HTL. These selected epitopes were used to construct the final multi-epitopes subunit vaccine. Linker such as AAY and GPGPG interconnect these epitopes and also enhanced effective separation and presentation of epitopes. These linkers have two essential roles; first, they assist binding of HLA-II epitopes and also immune processing, and second, they restrain epitopes numbers by specifying cleavage point (Bergmann et al., 1996; Dorosti et al., 2019; Livingston et al., 2002; Saadi et al., 2017). It has also been reported that these linkers specifically the GPGPG perform antigen presentation. Furthermore, the GPGPG space construct provoke the HTL response which direct the DNA or polypeptide immunization and thus prevent the infection. Further, to boost the immune responses, an adjuvant (TLR-7 agonist Beta-defensin 3 (Q5U7J2)) was affixed to the vaccine to construct N-terminus by EAAAK linker (Arai et al., 2001). The EAAAK linkers helps in efficient separation and eventually assist the reduction in the domain-domain interaction of different vaccine.

### Subunit vaccine and interferon

Online server IFNepitope (<http://crdd.osdd.net/raghava/ifnepitope/predict.php>) assists in designing and predicting the amino acid sequence from proteins, which have the potential to release IFN-gamma from CD4+ T cells. This server helps in the designing of effective and better subunit vaccine by identifying peptide which binds to MHC II and releases IFN-gamma (Dhanda et al., 2013)

### Prediction of vaccine allergenicity

For allergenicity prediction of multi-epitope subunit vaccine AlgPred (<http://crdd.osdd.net/raghava/algpred/>) web tool was used, in which algorithm like (IgEpitope + SVMc + MAST + ARPs BLAST) was applied (Saha & Raghava, 2006). This hybrid prediction veracity is about 85% at 0.4 thresholds. The server employed six distant routes to act. These approaches abet allergic protein prediction with perfection.

### Vaccine antigenicity prediction

Online server ANTIGENpro (<http://scratch.proteomics.ics.uci.edu/>) was used for antigenicity prediction (Cheng et al., 2005). Server to predict protein antigenicity used two approaches; primary protein sequence is presented multiple times, and five different machine-learning algorithms were applied to produce the result based on protein microarray data analysis.

### Physiochemical parameters and identification of domain

ProtParam (<http://web.expasy.org/protparam/>), an online server was used for the prediction of many physiochemical parameters such as theoretical PI, instability index, amino acid composition, in vitro and in vivo half-life, molecular weight, aliphatic index and also the (GRAVY) which is a grand average of hydropathicity (Gasteiger et al., 2005).

### Prediction of secondary structure

The primary amino-acid sequence was used to predict the secondary structure of Protein by PDBsum. Sequences that show homology to vaccine protein were selected for structure prediction. Position position-specific iterated (PSI-BLAST) identified homologous residues.

### Prediction of the tertiary structure

For prediction and analysis of protein tertiary structure, an online server Robetta (<http://rosetta.bakerlab.org>) was utilized, which is an automated tool (Kim et al., 2004). Protein sequences were submitted in FASTA format to predict 3D structure. Models were generated after parsing the structure into respective domains. This model is based either on comparative modelling or de novo structure; for comparative modelling homologs sequence was identified by BLAST, 3D-Jury or FFAS03, and PSI-BLAST, which is then used as a template. If homologs were not found, de novo structure was generated using the Rosetta fragment insertion method.

### Validation of tertiary structure

To validate the predicted model is an essential step, ProSA-web an online server (<https://prosa.services.came.sbg.ac.at/prosa.php>) was used for validation of tertiary structure. For input structure, overall quality scoring was computed. The predicted protein will have an error when the scoring characteristic was found arbitrary then native proteins. For statistical analysis of non-bonded interaction, ERRAT server was used (<https://servicesn.mbi.ucla.edu/ERRAT/>) (Colovos & Yeates, 1993), and Ramachandran plot analysis was performed on RAMPAGE server (<http://mordred.bioc.cam.ac.uk/~rapper/rampage.php>) (Colovos & Yeates, 1993). This server utilized PROCHECK principle to validate the structure of protein through Ramachandran plot and separate plots for Proline and Glycine residues.

### Vaccine and TLR-7 docking

ZDOCK server (<http://zdock.umassmed.edu/>) was used for the docking of the final vaccine with TLR-7 (Enosi Tuipulotu et al., 2018). The ZDOCK webserver produces quick and accurate complexes. Refinement and post-processing of the complexes were subjected to the Firedock server (Pierce et al., 2014).

### MD (Molecular dynamic simulation) of receptor-vaccine complex

For selected complexes, amber 14 (Pearlman et al., 1995) was used to conduct MD simulation. The system was neutralized and solvated with TIP3P water box. Two stages of energy minimization, followed by gentle heating and equilibration, were performed (Salomon-Ferrer et al., 2013). After equilibration, 100 ns simulation was conducted. For post-simulation trajectories analysis (RMSD and RMSF), 2.0 ps time scale was used and for trajectory, sampling using CPPTRAJ and PTRAJ (Roe & Cheatham, 2013) implemented in AMBER 14.

### Codon optimization and in silico cloning of vaccine

For effective expression of the vaccine in a host (*E. coli* strain K12) was performed. Reverse translation and codon optimization were performed using JCAT (Java Codon Adaption tool). Norovirus genome expression is distinct from the expression of vector genome. Codon optimization ensures maximum expression of vaccine with in the vector. In order to get the desired result, three additional options were selected, such as prokaryote ribosome binding site, restriction enzymes cleavage and rho-independent transcription termination. JCat output includes a codon adaptation index (CAI) and to ensure the high-level protein expression percentage GC content is used (Khan, Khan, Rehman, et al., 2019). To clone the desired gene in *E. coli* pET-28a (+) two restriction sites NdeI and XhoI were added to the C and N terminal of the sequence, respectively. Finally, the adapted sequence having the restriction site was inserted to pET-28a (+) to maximize the vaccine expression.

### Disclosure statement

No potential conflict of interest was reported by the author(s).

### Funding

Dong-Qing Wei is supported by the grants from the Key Research Area Grant 2016YFA0501703 of the Ministry of Science and Technology of China, the National Natural Science Foundation of China (Contract no. 61832019, 61503244), the Science and Technology Commission of Shanghai Municipality (Grant: 19430750600), the Natural Science Foundation of Henan Province (162300410060) and Joint Research Funds for Medical and Engineering and Scientific Research at Shanghai Jiao Tong University (YG2017ZD14). The computations were partially performed at the Pengcheng Lab and the Center for High-Performance Computing, Shanghai Jiao Tong University.

### References

- Adhikari, U. K., Tayebi, M., & Rahman, M. M. (2018). Immunoinformatics approach for epitope-based peptide vaccine design and active site prediction against polyprotein of emerging oropouche virus. *Journal of Immunology Research*, 2018, 6718083. doi: 10.1155/2018/6718083
- Agrawal, P., & Raghava, G. P. S. (2018). Prediction of antimicrobial potential of a chemically modified peptide from its tertiary structure.



- Frontiers in Microbiology*, 9, 2551. <https://doi.org/10.3389/fmicb.2018.02551>
- Ahmed, S. M., Hall, A. J., Robinson, A. E., Verhoef, L., Premkumar, P., Parashar, U. D., Koopmans, M., & Lopman, B. A. (2014). Global prevalence of norovirus in cases of gastroenteritis: A systematic review and meta-analysis. *The Lancet. Infectious Diseases*, 14(8), 725–730. doi: 10.1016/S1473-3099(14)70767-4
- Ali, A., Khan, A., Kaushik, A. C., Wang, Y., Ali, S. S., Junaid, M., Saleem, S., Cho, W. C. S., Mao, X., & Wei, D.-Q. (2019). Immunoinformatic and systems biology approaches to predict and validate peptide vaccines against Epstein-Barr virus (EBV). *Scientific Reports*, 9(1), 1–12.
- Ando, T., Monroe, S. S., Gentsch, J. R., Jin, Q., Lewis, D. C., & Glass, R. I. (1995). Detection and differentiation of antigenically distinct small round-structured viruses (Norwalk-like viruses) by reverse transcription-PCR and southern hybridization. *Journal of Clinical Microbiology*, 33(1), 64–71. <https://doi.org/10.1128/JCM.33.1.64-71.1995>
- Arai, R., Ueda, H., Kitayama, A., Kamiya, N., & Nagamune, T. (2001). Design of the linkers which effectively separate domains of a bifunctional fusion protein. *Protein Engineering*, 14(8), 529–532. <https://doi.org/10.1093/protein/14.8.529>
- Bergmann, C. C., Yao, Q., Ho, C. K., & Buckwold, S. L. (1996). Flanking residues alter antigenicity and immunogenicity of multi-unit CTL epitopes. *Journal of Immunology (Baltimore, MD.: 1950)*, 157(8), 3242–3249.
- Caul, E., & Appleton, H. (1982). The electron microscopical and physical characteristics of small round human fecal viruses: An interim scheme for classification. *Journal of Medical Virology*, 9(4), 257–265. <https://doi.org/10.1002/jmv.189090403>
- Cheng, J., Randall, A. Z., Sweredoski, M. J., & Baldi, P. (2005). SCRATCH: A protein structure and structural feature prediction server. *Nucleic Acids Research*, 33(Web Server issue), W72–W76. <https://doi.org/10.1093/nar/gki396>
- Chezzi, C., & Schoub, B. D. (1996). Differentiation between vaccine-related and wild-type polioviruses using a heteroduplex mobility assay. *Journal of Virological Methods*, 62(2), 93–102.
- Colovos, C., & Yeates, T. O. (1993). Verification of protein structures: Patterns of nonbonded atomic interactions. *Protein Science: A Publication of the Protein Society*, 2(9), 1511–1519. <https://doi.org/10.1002/pro.5560020916>
- Consortium, T. U. (2018). UniProt: A worldwide hub of protein knowledge. *Nucleic Acids Research*, 47(D1), D506–D515. <https://doi.org/10.1093/nar/gky1049>
- Cortes-Penfield, N. W., Trautner, B. W., & Jump, R. L. P. (2017). Urinary tract infection and asymptomatic bacteriuria in older adults. *Infectious Disease Clinics of North America*, 31(4), 673–688. <https://doi.org/10.1016/j.idc.2017.07.002>
- Dash, R., Das, R., Junaid, M., Akash, M. F. C., Islam, A., & Hosen, S. Z. (2017). In silico-based vaccine design against Ebola virus glycoprotein. *Advances and Applications in Bioinformatics and Chemistry: AABC*, 10, 11–28. <https://doi.org/10.2147/AABC.S115859>
- Dhanda, S. K., Vir, P., & Raghava, G. P. S. (2013). Designing of interferon-gamma inducing MHC class-II binders. *Biology Direct*, 8(1), 30. <https://doi.org/10.1186/1745-6150-8-30>
- Dorosti, H., Eslami, M., Negahdaripour, M., Ghoshoon, M. B., Gholami, A., Heidari, R., Dehshahri, A., Erfani, N., Nezafat, N., & Ghasemi, Y. (2019). Vaccinomics approach for developing multi-epitope peptide pneumococcal vaccine. *Journal of Biomolecular Structure and Dynamics*, 37(13), 3524–3535. <https://doi.org/10.1080/07391102.2018.1519460>
- Doytchinova, I. A., & Flower, D. R. (2007a). Identifying candidate subunit vaccines using an alignment-independent method based on principal amino acid properties. *Vaccine*, 25(5), 856–866. <https://doi.org/10.1016/j.vaccine.2006.09.032>
- Doytchinova, I. A., & Flower, D. R. (2007b). VaxiJen: A server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics*, 8(1), 4.
- Doytchinova, I. A., & Flower, D. R. (2008). Bioinformatic approach for identifying parasite and fungal candidate subunit vaccines. *Open Vaccine J*, 1(1), 4.
- Enosi Tuipulotu, D., Netzler, N. E., Lun, J. H., Mackenzie, J. M., & White, P. A. (2018). TLR7 agonists display potent antiviral effects against norovirus infection via innate stimulation. *Antimicrobial Agents and Chemotherapy*, 62(5), 1–15. <https://doi.org/10.1128/aac.02417-17>
- Fankhauser, R. L., Noel, J. S., Monroe, S. S., Ando, T., & Glass, R. I. (1998). Molecular epidemiology of “Norwalk-like viruses” in outbreaks of gastroenteritis in the United States. *The Journal of Infectious Diseases*, 178(6), 1571–1578. <https://doi.org/10.1086/314525>
- Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S. e., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). Protein identification and analysis tools on the ExPASy server. In J. M. Walker (Ed.), *The proteomics protocols handbook* (pp. 571–607). Humana Press.
- Green, J., Norcott, J., Lewis, D., Arnold, C., & Brown, D. (1993). Norwalk-like viruses: Demonstration of genomic diversity by polymerase chain reaction. *Journal of Clinical Microbiology*, 31(11), 3007–3012. <https://doi.org/10.1128/JCM.31.11.3007-3012.1993>
- Kaliyathur, S., Selvaraj, G., Junaid, M., Khan, A., Gu, K., & Wei, D.-Q. (2018). Cancer immunoinformatics: A promising era in the development of peptide vaccines for human papillomavirus-induced cervical cancer. *Current Pharmaceutical Design*, 24(32), 3791–3817. <https://doi.org/10.2174/1381612824666181106094133>
- Kapikian, A. Z., Wyatt, R. G., Dolin, R., Thornhill, T. S., Kalica, A. R., & Chanock, R. M. (1972). Visualization by immune electron microscopy of a 27-nm particle associated with acute infectious nonbacterial gastroenteritis. *Journal of Virology*, 10(5), 1075–1081. <https://doi.org/10.1128/JVI.10.5.1075-1081.1972>
- Khan, A., Ali, S. S., Khan, M. T., Saleem, S., Ali, A., Suleman, M., Babar, Z., Shafiq, A., Khan, M., & Wei, D.-Q. (2020). Combined drug repurposing and virtual screening strategies with molecular dynamics simulation identified potent inhibitors for SARS-CoV-2 main protease (3CLpro). *Journal of Biomolecular Structure and Dynamics*, 1–12.
- Khan, A., Junaid, M., Kaushik, A. C., Ali, A., Ali, S. S., Mehmood, A., & Wei, D.-Q. (2018). Computational identification, characterization and validation of potential antigenic peptide vaccines from hrHPVs E6 proteins using immunoinformatics and computational systems biology approaches. *PloS One*, 13(5), e0196484.
- Khan, A., Saleem, S., Idrees, M., Ali, S. S., Junaid, M., Kaushik, A. C., & Wei, D.-Q. (2018). Allosteric ligands for the pharmacologically important Flavivirus target (NS5) from ZINC database based on pharmacophoric points, free energy calculations and dynamics correlation. *Journal of Molecular Graphics & Modelling*, 82, 37–47. <https://doi.org/10.1016/j.jmgm.2018.03.004>
- Khan, M., Khan, S., Ali, A., Akbar, H., Sayaf, A. M., Khan, A., & Wei, D.-Q. (2019). Immunoinformatics approaches to explore Helicobacter Pylori proteome (Virulence Factors) to design B and T cell multi-epitope subunit vaccine. *Scientific Reports*, 9(1), 1–13.
- Khan, M. T., Ali, A., Wang, Q., Irfan, M., Khan, A., Zeb, M. T., Zhang, Y. J., Chinnasamy, S., & Wei, D.-Q. (2020). Marine natural compounds as potent inhibitors against the main protease of SARS-CoV-2. A molecular dynamic study. *Journal of Biomolecular Structure and Dynamics*, 1–14.
- Khan, S., Khan, A., Rehman, A. U., Ahmad, I., Ullah, S., Khan, A. A., Ali, S. S., Afridi, S. G., & Wei, D.-Q. (2019). Immunoinformatics and structural vaccinology driven prediction of multi-epitope vaccine against Mayaro virus and validation through in-silico expression. *Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases*, 73, 390–400. <https://doi.org/10.1016/j.meegid.2019.06.006>
- Kim, D. E., Chivian, D., & Baker, D. (2004). Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Research*, 32(Web Server issue), W526–W531. <https://doi.org/10.1093/nar/gkh468>
- Kreis, S., & Whistler, T. (1997). Rapid identification of measles virus strains by the heteroduplex mobility assay. *Virus Research*, 47(2), 197–203.
- Kringelum, J. V., Lundegaard, C., Lund, O., & Nielsen, M. (2012). Reliable B cell epitope predictions: Impacts of method development and improved benchmarking. *PLoS Computational Biology*, 8(12), e1002829. <https://doi.org/10.1371/journal.pcbi.1002829>
- Larsen, M. V., Lundegaard, C., Lamberth, K., Buus, S., Lund, O., & Nielsen, M. (2007). Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. *BMC Bioinformatics*, 8, 424. <https://doi.org/10.1186/1471-2105-8-424>

- Lew, J. F., Kapikian, A. Z., Valdesuso, J., & Green, K. Y. (1994). Molecular characterization of Hawaii virus and other Norwalk-like viruses: Evidence for genetic polymorphism among human caliciviruses. *Journal of Infectious Diseases*, 170(3), 535–542. <https://doi.org/10.1093/infdis/170.3.535>
- Livingston, B., Crimi, C., Newman, M., Higashimoto, Y., Appella, E., Sidney, J., & Sette, A. (2002). A rational strategy to design multiepitope immunogens based on multiple Th lymphocyte epitopes. *Journal of Immunology (Baltimore, MD.: 1950)*, 168(11), 5499–5506. <https://doi.org/10.4049/jimmunol.168.11.5499>
- Maguire, A. J., Green, J., Brown, D. W., Desselberger, U., & Gray, J. J. (1999). Molecular epidemiology of outbreaks of gastroenteritis associated with small round-structured viruses in East Anglia, United Kingdom, during the 1996–1997 season. *Journal of Clinical Microbiology*, 37(1), 81–89. <https://doi.org/10.1128/JCM.37.1.81-89.1999>
- Mattick, K. L., Green, J., Punia, P., Belda, F. J., Gallimore, C. I., & Brown, D. W. (2000). The heteroduplex mobility assay (HMA) as a pre-sequencing screen for Norwalk-like viruses. *Journal of Virological Methods*, 87(1–2), 161–169.
- Mattison, C. P., Cardemil, C. V., & Hall, A. J. (2018). Progress on norovirus vaccine research: Public health considerations and future directions. *Expert Review of Vaccines*, 17(9), 773–784. <https://doi.org/10.1080/14760584.2018.1510327>
- McFadden, N., Bailey, D., Carrara, G., Benson, A., Chaudhry, Y., Shortland, A., Heeney, J., Yarovsky, F., Simmonds, P., Macdonald, A., & Goodfellow, I. (2011). Norovirus regulation of the innate immune response and apoptosis occurs via the product of the alternative open reading frame 4. *PLoS Pathogens*, 7(12), e1002413. <https://doi.org/10.1371/journal.ppat.1002413>
- Mead, P. S., Slutsker, L., Dietz, V., McCaig, L. F., Bresee, J. S., Shapiro, C., Griffin, P. M., & Tauxe, R. V. (1999). Food-related illness and death in the United States. *Emerging Infectious Diseases*, 5(5), 607–625. <https://doi.org/10.3201/eid0505.990502>
- Nielsen, M., & Lund, O. (2009). NN-align. An artificial neural network-based alignment algorithm for MHC class II peptide binding prediction. *BMC Bioinformatics*, 10, 296. <https://doi.org/10.1186/1471-2105-10-296>
- Norcott, J., Green, J., Lewis, D., Estes, M., Barlow, K., & Brown, D. (1994). Genomic diversity of small round structured viruses in the United Kingdom. *Journal of Medical Virology*, 44(3), 280–286. <https://doi.org/10.1002/jmv.1890440312>
- Pearlman, D. A., Case, D. A., Caldwell, J. W., Ross, W. S., Cheatham, T. E., DeBolt, S., Ferguson, D., Seibel, G., & Kollman, P. (1995). AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Computer Physics Communications*, 91(1–3), 1–41. [https://doi.org/10.1016/0010-4655\(95\)00041-D](https://doi.org/10.1016/0010-4655(95)00041-D)
- Pierce, B. G., Wiehe, K., Hwang, H., Kim, B.-H., Vreven, T., & Weng, Z. (2014). ZDOCK server: Interactive docking prediction of protein-protein complexes and symmetric multimers. *Bioinformatics (Oxford, England)*, 30(12), 1771–1773. <https://doi.org/10.1093/bioinformatics/btu097>
- Pires, S. M., Fischer-Walker, C. L., Lanata, C. F., Devleeschauwer, B., Hall, A. J., Kirk, M. D., Duarte, A. S. R., Black, R. E., & Angulo, F. J. (2015). Aetiology-specific estimates of the global and regional incidence and mortality of diarrhoeal diseases commonly transmitted through food. *PLoS One*, 10(12), e0142927. <https://doi.org/10.1371/journal.pone.0142927>
- Pulendran, B., & Ahmed, R. (2011). Immunological mechanisms of vaccination. *Nature Immunology*, 12(6), 509–517. <https://doi.org/10.1038/ni.2039>
- Rappuoli, R. (2000). Reverse vaccinology. *Current Opinion in Microbiology*, 3(5), 445–450. [https://doi.org/10.1016/S1369-5274\(00\)00119-3](https://doi.org/10.1016/S1369-5274(00)00119-3)
- Robilotti, E., Deresinski, S., & Pinsky, B. A. (2015). Norovirus. *Clinical Microbiology Reviews*, 28(1), 134–164. doi: 10.1128/CMR.00075-14
- Roe, D. R., & Cheatham, T. E. (2013). PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data. *Journal of Chemical Theory and Computation*, 9(7), 3084–3095. <https://doi.org/10.1021/ct400341p>
- Saadi, M., Karkhah, A., & Nouri, H. R. (2017). Development of a multi-epitope peptide vaccine inducing robust T cell responses against brucellosis using immunoinformatics based approaches. *Infection, Genetics and Evolution*, 51, 227–234. <https://doi.org/10.1016/j.meegid.2017.04.009>
- Saha, S., & Raghava, G. P. S. (2006). AlgPred: Prediction of allergenic proteins and mapping of IgE epitopes. *Nucleic Acids Research*, 34(Web Server issue), W202–W209. <https://doi.org/10.1093/nar/gkl343>
- Salomon-Ferrer, R., Case, D. A., & Walker, R. C. (2013). An overview of the Amber biomolecular simulation package. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 3(2), 198–210. doi: 10.1002/wcms.1121
- Sette, A., & Fikes, J. (2003). Epitope-based vaccines: An update on epitope identification, vaccine design and delivery. *Current Opinion in Immunology*, 15(4), 461–470.
- Vinje, J., & Koopmans, M. P. (1996). Molecular detection and epidemiology of small round-structured viruses in outbreaks of gastroenteritis in the Netherlands. *The Journal of Infectious Diseases*, 174(3), 610–615. <https://doi.org/10.1093/infdis/174.3.610>
- Wang, P., Sidney, J., Dow, C., Mothe, B., Sette, A., & Peters, B. (2008). A systematic assessment of MHC class II peptide binding predictions and evaluation of a consensus approach. *PLoS Computational Biology*, 4(4), e1000048. <https://doi.org/10.1371/journal.pcbi.1000048>
- Wang, P., Sidney, J., Kim, Y., Sette, A., Lund, O., Nielsen, M., & Peters, B. (2010). Peptide binding predictions for HLA DR, DP and DQ molecules. *BMC Bioinformatics*, 11, 568. <https://doi.org/10.1186/1471-2105-11-568>
- Weiner, A., Houghton, M., Polito, A., Massaro, A., Bonino, F., Brunetto, M., Danielle, F., Calabrese, G., Guaschino, R., Dinello, R., & Quan, S. (1998). Hepatitis C virus heteroduplex tracking assay for genotype determination reveals diverging genotype 2 isolates in Italian hemodialysis patients.
- White, M. B., Rajagopalan, S., & Yoshikawa, T. T. (2016). Infectious Diarrhea: Norovirus and Clostridium difficile in older adults. *Clinics in Geriatric Medicine*, 32(3), 509–522. <https://doi.org/10.1016/j.cger.2016.02.008>