



# Exploring Highly Antigenic Protein of *Campylobacter jejuni* for Designing Epitope Based Vaccine: Immunoinformatics Approach

Nayan Gupta<sup>1</sup> · Fariya Khan<sup>1</sup> · Ajay Kumar<sup>1</sup>

Accepted: 17 September 2018 / Published online: 22 September 2018  
© Springer Nature B.V. 2018

## Abstract

*Campylobacter jejuni* is a spiral-shaped Gram-negative and microaerophilic bacteria that causes bacterial diarrhea worldwide. Due to its complicated antigenic nature, lack of proper animal models, and a poor understanding of its pathogenicity there is no vaccine against *C. jejuni*. The most effective way of reducing the number of *Campylobacter* infections is by developing protective vaccines for humans. Due to the increasing resistance against antibiotics of bacterial strains epitopes-based vaccine could be the best approach to trigger an effective immune response. Immunoinformatics methods can be an alternative approach for developing epitope-based vaccine which is able to elucidate humoral and cell-mediated immune response inside the host body. Reverse vaccinology approach analyses the entire protein sequences of the pathogen using bioinformatics tools to select target proteins for their high-throughput expression and validation. Here in this research study, three epitopes MSNVYAYRF, LSDDINLNI and ATSTSTITL have been identified as the most potential epitopes binding with HLA-B\*58:01, HLA-A\*01:01 and HLA-B\*07:02 MHC class I allele, respectively. With the help of Autodock vina, these three epitopes have shown the lowest binding energies and form the most complex docked structure covering the maximum number of populations worldwide. Thus, we suggest these peptides may have a good potential if further presented for experimental analysis and can be proved to be helpful against *C. jejuni* infections causing diarrhea.

**Keywords** Immunoinformatics · Epitopes · Cytotoxic T cells (CTL) · MHC complex

## Introduction

*Campylobacter jejuni* is a Gram-negative spiral-shaped bacteria which is one of the leading causes of foodborne bacterial diseases worldwide. The isolation of *C. jejuni* was accomplished from human feces in 1968 (Dekeyser et al. 1972). Since its discovery, *C. jejuni* remains the most frequent cause of infectious diarrhea affecting over 450 million people every year throughout the world, attributing a large economic burden (Friedman et al. 2000). It is most commonly found in animal feces and therefore poultry is the major reservoir (Young et al. 2007). Diarrhea is the third leading cause of childhood mortality in India and it is

responsible for 13% of all death per year in children under 5 year of age killing in an estimated 300,000 children per year in India. *Campylobacter* infection affects more than 1.3 million people every year (CDC 2017). According to WHO report, diarrheal diseases are the most common infection resulting from unhygienic food causing 550 million people falling ill yearly out of which 220 million children below the age of 5 years and it is one of the key global causes of diarrheal diseases (WHO 2018). *C. jejuni* infects the organism by penetrating the gastrointestinal mucus, then adhering to the gut enterocytes and induced diarrhea by toxin release mainly enterotoxin and cytotoxin that correlate with the severity of enteritis (Gallardo et al. 1998; Wallis 1994). The limited data on the treatment of *C. jejuni* infection suggested that ciprofloxacin, erythromycin, and fluoroquinolones may shorten the duration of symptoms. However, treatment failure associated with the emergence of the quinolone-resistant strain of *C. jejuni* has been documented (Van et al. 2003). For the permanent solution with the infection of *C. jejuni*, there is need of a powerful vaccine that

---

Nayan Gupta and Fariya Khan contributed equally.

✉ Ajay Kumar  
ajaymtech@gmail.com

<sup>1</sup> Department of Biotechnology, Faculty of Engineering & Technology, Rama University, Kanpur, Uttar Pradesh 209217, India

could provide immunity against it but till now no vaccine is available against *Campylobacter*.

Reverse vaccinology is the best approach in terms of medical sciences that involves the use of genomic information with the use of computer for the preparation of vaccines without culturing microorganism (Kanampalliwar et al. 2013; Tang et al. 2012). This method helps in the development of vaccines that were previously difficult or impossible to make and therefore can lead to the discovery of the unique epitope-based vaccine that will improve existing vaccine technology (Sette and Rappuoli 2010). It allows the selection of antigenic regions from the pathogenic genomes and the most antigenic regions could be developed as potential vaccine candidates to trigger protective immune responses (Ada et al. 2018). Production of peptide or epitopes based vaccines are easy in comparison to the conventional vaccine and it is also more specific, cost-effective, less time consuming and safe (Kumar et al. 2015). An epitope-based vaccine recently drawn much attention in treating infectious diseases (Yang and Yu 2009). Immunization based on epitope-based vaccines is more powerful in stimulation of the cellular and humoral immune response (Bijker et al. 2007). These types of vaccines consist of highly immunogenic T and B cells epitopes, which provoke cytotoxic T cells (CTL), T<sub>H</sub> or B to specific epitopes (Baloria et al. 2012; Akhoon et al. 2011). B and T<sub>H</sub> cells play an important role in the induction of a protective immune response in many bacterial infections; thus, determination of peptides that induce T and B cells response is the crucial requirement for the design of effective epitope-based vaccines (Gupta et al. 2010, 2012). Due to the heterogeneous virulence nature of *C. jejuni* infection, there is a lack of proper animal models that can replicate the gastroenteritis caused by *C. jejuni* in humans. *C. jejuni* vaccine candidates that have been or are in clinical development include killed the whole cell, protein subunit, and capsule-conjugate products (Man 2011). A recombinant protein vaccine by the name ACE 393 has been tested in Phase II human challenge study but its efficacy failed to demonstrate protection (Riddle and Guerry 2016). To validate the analysis, two standard antigens were taken as control (CadF and FlpA) which helps in the colonization of the gut that is promoted by flagellum-mediated motility and binding to host tissue such as fibronectin (Flanagan et al. 2009; Jin et al. 2001).

In the present study, screening of the whole genome in search of probable antigens was performed by using the RV approach. Immuno-informatics or computational biology has added an unavoidable contribution to design epitope-based vaccines (Gupta et al. 2011). In this context, identification of potential epitopes from an antigen protein by in silico methods can be considered in such vaccines reducing the lengthy process for discovery of appropriate epitopes (Srivastava et al. 2011). Instead of selecting individual proteins,

the RV approach analyses the entire protein sequences of the pathogen using bioinformatics tools to select target proteins for their high-throughput expression and validation in animal models. The wide application of computational method provides an easy way to identify the epitopes of highly antigenic protein. Vaxign, a server that predicts the novel antigen which can be used as potential candidates for designing vaccine (Yongqun and Mobley 2010). Predicted epitopes will be further examined for different physiochemical and immunogenic properties as well as three-dimensional structural modeling and the docking studies will be performed to investigate the strong binding interaction with MHC I alleles. Nowadays, many online servers are available for predicting B and T cell epitopes. In this respect, the Immune Epitope Database (IEDB) server website (Vita et al. 2010) provides tools to predict both B and T cells epitopes. Other online servers such as ProPred (Singh and Raghava 2001), MHCpred (Guan et al. 2003) and SVMHC (Donnes and Elofsson 2002) have different tools for finding T cell epitopes. In order to examine the binding capability of these predicted epitopes with their corresponding alleles and population coverage analysis were performed.

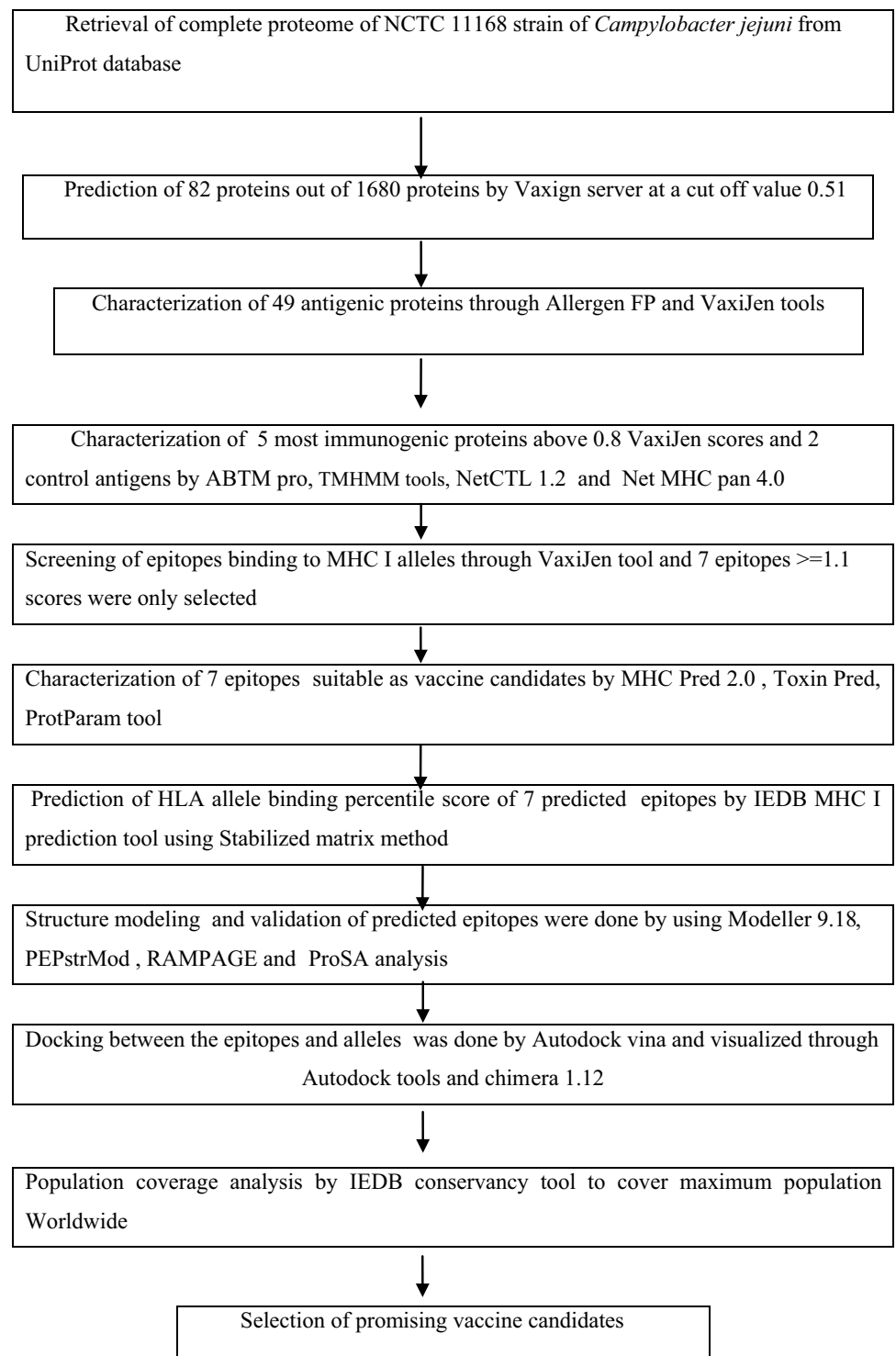
## Methodology

The *C. jejuni* strain NCTC 11,168 has a circular chromosome of 1,641,481 base pairs (30.6% G+C) that encodes 1680 proteins (Parkhill et al. 2000). The most striking feature of this strain was the presence of hypervariable sequences and lack of repetitive DNA sequences within this genome. The steps involved in the screening of the complete genome have been portrayed in Fig. 1.

## Retrieval of Proteome Data and Identification of Antigens

The complete proteome of diarrheagenic *Campylobacter jejuni* strain NCTC 11,168 was isolated from UniProt database having proteome id UP000000799 (<http://www.uniprot.org/proteomes/>) showing the total number of 1680 protein coding sequences with their complete protein information. Adhesin Probability  $\geq 0.51$  of the filtered antigens are screened from the complete proteome of *Campylobacter jejuni* through Vaxign Server (Xiang and He 2009) (<http://www.violinet.org/vaxign/index.php>). Vaxign is a web-based software that works on the principle of reverse vaccinology and predicts the vaccine targets by computing the complete genome using default parameters. This server uses the default cutoff value of 0.51 to predict adhesion probability of the vaccine candidate. The adhering capability of the pathogenic strain is an important characteristic in causing

**Fig. 1** Flowchart portraying the screening of potential vaccine candidates from the *Campylobacter jejuni* genome



infection and therefore considered to be the major target in the development of vaccine.

### Prediction of Antigenic and Allergic Properties

Selected proteins were further subjected to VaxiJen tool for the characterization of the antigenic nature of protein

(Doytchinova et al. 2007). VaxiJen is the first server to predict the protective antigens mainly based on their physico-chemical properties of proteins irrespective of the sequence alignment. In order to achieve high accuracy of prediction, only those proteins that show above 0.51 cut off scores were selected and therefore 49 proteins were calculated for analysis. These antigenic proteins sequences were also tested for

their allergenicity prediction by using Allergen FP v.1.0 tool available at <http://www.ddg-pharmfac.net/AllergenFP/index.html>. It is a novel fingerprint based approach that differentiates between the allergens and non-allergens (Dimitrov et al. 2014).

### Prediction of T Cell Epitopes

The binding capability of the T cell epitopes with HLA class I alleles were checked by using NetCTL 1.2 tool (Larsen et al. 2007) and Net MHC pan 4.0 (Vanessa et al. 2017). NetCTL 1.2 server predicts CTL epitopes in any given protein sequences, proteasomal C terminal cleavage and TAP transport efficiency. This server allows for the predictions of CTL epitopes restricted to 12 MHC class I supertype. It uses the different algorithm for prediction of MHC class I binding and proteasomal cleavage by using artificial neural networks and TAP transport efficiency prediction through weight matrix method. To confirm the predicted peptides analyze through NetCTL 1.2 server it was further validated by using Net MHC pan 4.0. This server predicts binding of peptides to an MHC molecule of the antigenic sequence using artificial neural networks (ANNs).

To confirm the most probable epitopes with the higher confidence level, the final peptides were analyzed again through VaxiJen v2.0 web server and peptides that have shown high scores were selected for the analysis. High scorer peptides showing score  $\geq 1.1$  were shortlisted for the toxicity prediction. On the basis of these scores, best vaccine candidates were selected for further analysis.

### Characterization of Target Proteins as a Vaccine Candidate

SPAAN is a non-homology based software program that is based on the properties of the sequence presented. It identifies top-scoring novel adhesins with the high level of confidence, therefore, the adhesion property of the antigens was checked by SPAAN (Sachdeva et al. 2005). Localization of these proteins was predicted by PSORT b 3.0 (Yu et al. 2010). PSORTb has remained the most precise bacterial protein, subcellular localization (SCL) predictor which analyzed the subcellular localization by using multiple analytical modules. Prediction of the subcellular localization of proteins through computational approach is a valuable tool for genome analysis and annotation and it can provide information regarding its function in an organism. Transmembrane regions were also checked by TMHMM method (Krogh et al. 2001). TMHMM is a membrane protein topology prediction method and works on the principle of hidden Markov model. It predicted the number of transmembrane helices within the protein sequences and discriminate between soluble and membrane proteins with the high level of accuracy.

### Prediction of Physicochemical Properties of Proteins

Identification of physical and chemical properties have been performed by ProtParam tool (Gasteiger et al. 2005). These parameters include molecular weight, theoretical pI, extinction coefficient, estimated half-life, instability index, aliphatic index and grand average of hydropathicity. ABT-Mpro is a server that characterizes whether a given protein sequence is a transmembrane protein or not. It further identifies the probabilities of the transmembrane protein predicted by ABTMpro being an alpha-helical transmembrane protein or a Beta Barrel transmembrane protein. It consists of a Support Vector Machine, which utilizes features such as amino acid composition and properties, reduced alphabet composition, predicted secondary structure, evolutionary information etc. ABTM pro is available at <http://scratch.proteomics.ics.uci.edu/>.

### IEDB Percentile Score Calculation of Predicted Epitopes

Binding affinities of the peptides were calculated by IEDB tool by applying Stabilized Matrix Method against the MHC restricted HLA alleles (Peters and Sette 2005). Toxicity of all the selected peptides were examined by Toxin Pred tool (Gupta et al. 2013). Toxin Pred is an in silico tool that predicts and design the toxic as well as non-toxic peptides. Here, in our study, we require non-toxic peptides, therefore, computational methods are useful in predicting the toxicity of peptides but also facilitate the designing of better useful peptides with lower toxicity level (Khan et al. 2017). This tool not only discriminates against toxic peptides from non-toxic peptides but also save time and money.

### Molecular Modeling of HLA Alleles and Epitopes

The three-dimensional structure of the epitopes was designed by using PEPstrMOD (Kaur et al. 2007), a server that predicts the 3D structures of the small peptide as well as peptides having non-natural residues, terminal modifications, cyclization, and post-translational modifications, etc. The sequence of HLA alleles was downloaded from IPD-IMGT/HLA Database (Robinson et al. 2015). The 3-D structures of these HLA alleles were modeled by Modeller 9.18 (Sali et al. 1995). MODELLER is used for homology or comparative modeling of the three-dimensional structures of the proteins and provides an alignment of a protein sequence to be modeled with their known structural template. It automatically calculates five models and the best of all model was selected on the basis of their lowest DOPE score. The most acceptable model was further validated by RAMPAGE

and ProSA analysis (Wiederstein and Sippl 2007). This tool is widely used to check 3D models of protein structures for potential errors.

### Molecular Docking Analysis of the Peptide and HLA Allele Complex

After structural modeling, docking of the selected peptides with their favored HLA alleles was performed by using AutoDock Vina (Trott and Olson 2010). The tool generates 9 docked conformations that were ranked on the basis of binding energy, torsion energy, geometry and electrostatic energy. The best output was finalized on the basis of highest binding energy score. AutoDock Vina did considerably better in terms of speed and accuracy. Preparation of ligand and protein was done by using Autodock 4.2 of MGL tool 1.5.4. The chimera tool was used for the visualization of the 3-D structure with their Binding position, H-bonding between the HLA alleles and peptides.

### Worldwide Population Coverage Analysis

The frequency of distribution of human HLA alleles binding to MHC molecules was predicted by MHC Pred (Guan et al. 2003) and their conservancy analysis among the predicted epitopes was carried out by IEDB population coverage tool (Bui et al. 2006). All the parameters were set to default values that have a dataset containing the total number of 3245 alleles for 16 geographical areas, 21 ethnicities, and 115 countries. The predicted antigenic epitopes with their corresponding HLA alleles were submitted to this database and it determines the percentage of the number of population interacting to a given set of epitopes.

## Results and Discussion

### Identification of Protective Antigens and Antigenicity Prediction

In the present study, the screening of complete proteome of *C. jejuni* in order to identify protective antigens with higher affinity was done by Vaxign server. Identification of a set of 82 proteins was shortlisted on the basis of adhesion value above 0.51 which indicates the adhesion properties of the filtered proteins. Subsequently, the next step is the filtering of predicted proteins via VaxiJen to select only those sequences having the high probability of antigenic epitopes. Out of 82 proteins, 49 proteins were predicted as potential antigens and furthermore can be characterized by their allergic properties with AllergenFP tool to exclude the allergens proteins. The non-allergens with high VaxiJen scores above

**Table 1** Screening of complete proteome through Vaxign server by keeping the cut off value  $\geq 0.51$  and further filtered by using VaxiJen and AllergenFP

S. no.	Uniprot id	VaxiJen	Allergen FP
1	Q0PCA5	0.3867	Allergen
2	Q0PC76	0.5793	Allergen
3	Q0PC48	0.371	Non-allergen
4	Q0PC46	0.5572	Non-allergen
5	Q0PC44	0.7782	Allergen
6	Q0PC26	0.5479	Non-allergen
7	Q0PC23	0.6809	Non-allergen
8	Q9PIZ3	0.3871	Non-allergen
9	Q0PBY5	0.4596	Non-allergen
10	Q0PBW1	0.5736	Non-allergen
11	Q0PBV8	0.6633	Non-allergen
12	Q0PBT9	0.2171	Allergen
13	Q0PBT8	0.3796	Non-allergen
14	Q0PBD5	0.6475	Non-allergen
15	Q0PBB5	0.4647	Allergen
16	Q0PBA4	0.5608	Non-allergen
17	Q0PB95	0.5547	Allergen
<b>18</b>	<b>Q0PB90</b>	<b>0.8359</b>	<b>Non-allergen</b>
<b>19</b>	<b>Q0PB82</b>	<b>0.8586</b>	<b>Non-allergen</b>
20	Q0PB28	0.3158	Non-allergen
21	Q9PHW6	0.7146	Non-allergen
22	Q0PAV5	0.5272	Allergen
23	Q0PAV3	0.2984	Non-allergen
24	Q0PAT7	0.3391	Allergen
25	Q0PAS3	0.7902	Non-allergen
26	Q0PAS1	0.4407	Non-allergen
27	Q0PAN9	0.7387	Non-allergen
28	Q9PHM1	0.4427	Non-allergen
29	Q0PAM0	0.5649	Allergen
30	Q9PPM0	0.6974	Non-allergen
31	P96747	0.5544	Non-allergen
32	Q0PAE4	0.5087	Non-allergen
33	Q0PAE2	0.4144	Non-allergen
34	Q0PAE1	0.6852	Allergen
35	Q0PAE0	0.7900	Allergen
36	Q0PAB4	0.5206	Non-allergen
37	Q0PA94	0.2177	Non-allergen
38	Q0PA76	0.5452	Allergen
39	Q0PA73	0.3329	Non-allergen
<b>40</b>	<b>Q0PA31</b>	<b>0.9818</b>	<b>Non-allergen</b>
41	Q0PA24	0.3475	Allergen
42	Q0PA18	0.4531	Non-allergen
43	Q0PA11	0.6976	Non-allergen
44	Q0P9Z2	0.4270	Non-allergen
45	Q0P9Z1	0.5586	Allergen
46	Q46121	0.5656	Non-allergen
47	Q0P9Y3	0.2010	Non-allergen
48	Q0P9W3	0.7357	Non-allergen

**Table 1** (continued)

S. no.	Uniprot id	VaxiJen	Allergen FP
49	Q9PNZ0	0.3095	Non-allergen
50	Q0P9T3	0.5119	Allergen
51	Q0P9T1	0.2535	Non-allergen
52	P45492	0.6514	Non-allergen
53	Q0P9N2	0.6710	Non-allergen
54	P0C633	0.5108	Non-allergen
55	Q0P9B4	0.4059	Non-allergen
56	Q0P921	0.4279	Allergen
57	Q0P920	0.3031	Non-allergen
<b>58</b>	<b>Q0P914</b>	<b>0.8535</b>	<b>Non-allergen</b>
59	Q0P904	0.5283	Non-allergen
60	Q0P8Z8	0.3073	non-allergen
61	P80672	0.6142	Allergen
62	Q0P8Y7	0.6484	Allergen
63	Q0P8W8	0.4104	Non-allergen
64	P56964	0.7211	Non-allergen
65	P56963	0.7571	Non-allergen
66	Q0P8Q8	0.5676	Allergen
67	Q0P8Q1	0.6954	Non-allergen
68	Q0P8M8	0.6550	Non-allergen
69	Q9PMJ8	0.7886	Allergen
70	Q9PMJ7	0.5131	Non-allergen
<b>71</b>	<b>Q0P8F1</b>	<b>0.8564</b>	<b>Non-allergen</b>
72	Q0P8E9	0.6043	Non-allergen
73	Q0P893	0.4293	Non-allergen
74	Q0P893	0.4283	Non-allergen
75	Q0P807	0.4955	Allergen
76	Q0P7x9	0.4177	Allergen
77	Q0P7x5	0.7783	Non-allergen
78	Q0P7W4	0.4375	Non-allergen
79	Q0P7W1	0.3830	Non-allergen
80	Q0P7V2	0.7041	Allergen
81	Q0P7R9	0.2355	Non-allergen
82	Q0P7Q2	0.7888	Non-allergen

Proteins highlighted in bold fonts are selected above 0.8 VaxiJen score

0.8 were selected for the further analysis- Q0PB90, Q0PB82, Q0PA31, Q0P914, Q0P8F1 (Table 1).

### Characterization of Target Antigens as Potential Vaccine Candidates

To characterize predicted antigens as good vaccine candidates, we have applied various computational methods to distinguish their immunological properties (Khan et al. 2018). The SPAAN program (Sachdeva et al. 2005) is applied to calculate the adhesion property of the target antigens, all the 5 proteins were predicted as adhesions.

Outer membrane surface proteins were considered to be a good choice for vaccine development and so the prediction of localization of these proteins was done by PSORT b 3.0. Usually, proteins having more than two transmembrane region are not considered a good vaccine candidate due to the difficulty in cloning expressing and purifying. Here we had used the TMHMM method which clearly depicted that all the 5 selected proteins were having less than 1 transmembrane region. Since it is difficult to clone, express and purify protein antigens with 1 or more TM region and is a major obstacle in the potential vaccine candidates selection. Furthermore, ABTM pro was applied to predict alpha helix and beta barrel transmembrane proteins with the overall accuracy of 97% (Singh et al. 2017) (Table 2).

### Prediction and Analysis of MHC Restricted T Cell Epitopes

Epitopes that are capable of inducing T-cell immune response are considered to be a good vaccine candidate. In the recent years, all the efforts have been devoted to the identification of MHC restricted T cell epitopes (Singh et al. 2010). For the identification of these epitopes, amino acid sequences of all the predicted proteins were subjected to Net CTL 1.2 for T- cell epitope prediction keeping the combined epitope prediction score of 0.75 (Tables 3 and 4). Thus, 5 predicted antigens sequences along with two control proteins were screened for their promiscuous HLA class I-restricted supertype binding epitopes using the NetCTL algorithm. The control antigen proteins were included in the study to compare the epitope density/immunogenicity level of the predicted antigens. Cytotoxic T cells are considered to be important in the immune system's response to disease as they have the capability to recognize defective cells by binding to peptides presented on the cell surface by MHC (Major Histocompatibility Complex) class I molecules (Huber et al. 2014). To validate the peptides predicted through the NetCTL algorithm another latest more precise server Net MHC pan 4.0 were applied and results were compared by comparing the peptides scores. The threshold for the strong binders was kept at its default value 0.5 and therefore peptides showing below this values were considered to be the weak binder (Tables 5 and 6). In addition, the selected epitopes were further subjected to VaxiJen server in order to predict their antigenicity to elicit an effective immune response. As the epitopes displaying the top scores  $\geq 1.1$  were selected and low score binders were eliminated as shown in Table 7.

We have presented all the selected 7 epitopes to MHC Pred and IEDB to predict the binding energies to their corresponding HLA alleles. Peptides binding to MHC class I

**Table 2** Adhesion probability of target proteins, localization, and transmembrane helices prediction

Antigen no.	Uniprot id	Localization	Adhesin (probability)	Predicted TM protein using ABTM pro			Number of predicted transmembrane regions
				$\alpha$ helix	$\beta$ barrel	Total	
1	Q0PB90	Putative periplasmic protein	0.566	0.0657	0.7674	0.1667 Non TM	0
2	Q0PB82	Extracellular	0.537	0.0355	0.03494	0.9294 Non TM	0
3	Q0PA31	Unknown	0.547	0.1324	0.0184	0.8490 Non TM	0
4	Q0P914	Unknown	0.636	0.0848	0.01739	0.8977 Non TM	0
5	Q0P8F1	Unknown	0.647	0.1036	0.0158	0.8805 Non TM	0
CadF	Q0P8D9	Outer membrane	0.481	0.0147	0.9815	0.0037 Non TM	0
FlpA	P45492	Outer membrane	0.624	0.0920	0.1161	0.7917 Non TM	0

**Table 3** Prediction of antigenic epitopes through Net CTL 1.2 for A supertypes

Protein Id	A1	A2	A3	A24	A26
Q0PB90	LSDDINLNI (175: 2.0521)	VLLSSLVAV (4: 1.3484)	LLSTGLFAK (14: 1.3731)	KYPDMTFTM (98: 1.6860)	NVKGNFKDY (44: 1.4057)
Q0PB82	QVDKQIEDY (22: 2.9105)	LLNDLLNAI (119: 1.2336)	QMSNVYAYR (76: 1.0439)	NYAQMSNVY (73: 1.0942)	NSNSNLSNY (66: 1.1849)
Q0PA31	NSLNSLSQY (5: 1.2310)	SLNSLSQYV (6: 1.2814)	–	–	KIVELTAKM (97: 1.2555)
Q0P914	–	AIFSTTTKV (12: 1.1968)	RSYAWQKMR (68: 1.1824)	QYKKNEETL (79: 1.1841)	AVREDFRSY (62: 1.2367)
Q0P8F1	MINPIQQSY (1: 1.3776)	–	MINPIQQSY (1: 1.3149)	–	DTKATAAAI (52: 0.8306)
Q0P8D9	ITPTLNINY (25: 2.3761)	FLCLGLASV (5: 1.1864)	ISFGFGGKK (165:1.4998)	GYHFDDFWL (51: 1.3995)	SVANELEKY (271: 2.2164)
P45492	SSDSYTLSF (164:3.1484)	KLSSDSYTL (162:1.2548)	KANASISIK (105:1.4030)	KFYNAKLNLF (190:1.1063)	NLKEDLLNY (205:1.6233)

**Table 4** Prediction of antigenic epitopes through Net CTL 1.2 for B supertypes

	B7	B8	B27	B39	B44	B58	B62
Q0PB90	ATSTSTITL (167: 0.9280)	MKKVLSSL (1: 1.3003)	KRSDFKAT (160: 0.7754)	GKEKIGFSL (147: 1.0677)	–	KIKRSDFKF (158: 1.2032)	KIKRSDFKF (158: 0.9472)
Q0PB82	RFRQNEGL (84: 0.8202)	–	YRTISSYFY (5: 1.6122)	EQSKNNTL (112: 1.3676)	NEQSKNNL (111: 1.1593)	MSNVYAYF (77: 1.8605)	AQMSNVYAY (75: 1.4850)
Q0PA31	APAQALNQ (61: 1.1685)	EIQAKIVEL (93: 0.9351)	–	SKNQEQNL (25: 1.0289)	–	–	SQIATLNAQ (118: 0.8641)
Q0P914	AVREDFRY (62: 0.7932)	QYKKNEEL (79: 0.9944)	FRSYAWQM (67: 1.0970)	SADEFQAL (26: 1.4030)	NEETLLNL (83: 1.2517)	–	AVREDFRSY (62: 1.3854)
Q0P8F1	–	QSYVANTAL (7: 0.9343)	–	QSYVANTL (7: 1.5950)	–	KASKIAEQI (36: 1.0795)	MINPIQQSY (1: 1.3496)
Q0P8D9	TPTLNYNF (26: 1.2191)	KIKEIAKVL (230: 1.3936)	RRVDAKFIL (310: 1.7497)	FRLSDSLAL (134: 2.0143)	GEKFYFYL (97: 2.1167)	ISLEGHFGF (210: 1.9197)	WVSTLGISF (159: 1.3660)
P45492	LPFIMFSML (358: 1.2713)	NIKIRSNEI (93: 1.4201)	YKLNLDLKF (316: 1.0957)	FLSIGIAVL (6: 1.6274)	YENQLNQV (30: 1.4166)	KSIQSNLVF (124: 2.0329)	SMLMGGASF (364: 1.4943)

**Table 5** Prediction of antigenic epitopes through Net MHC pan 4.0 for A supertypes

Protein id	A1	A2	A3	A24	A26
Q0PB90	LSDDINLNI (175: 0.729)	VLLSSLVAV (4: 0.924)	LLSTGLFAK (14: 0.7055)	KYPDMTFTM (98: 0.9765)	NVKGNFKDY (44: 0.44087)
Q0PB82	QVDKQIEDY (22: 0.9558)	LLNDLLNAI (119: 0.8682)	–	NYAQMSNVY (73: 0.4090)	NSNSNLSNY (66: 0.4364)
Q0PA31	NSLNSLSQY (5: 0.5518)	SLNSLSQYV (6: 0.8445)	SLNSLSQYV (6: 0.5884)	–	NSLNSLSQY (5: 0.3310)
Q0P914	–	TLLNKLFTT (86: 0.6868)	AIFSTTTKK (12: 0.8881)	QYKKNEETL (79: 0.5182)	AVREDFRSY (62: 0.5847)
Q0P8F1	MINPIQQSY (1: 0.8238)	–	ALNTNRIDK (14: 0.6516)	–	MINPIQQSY (1: 0.7634)
Q0P8D9	HTDNIGSRAY (251: 0.9953)	VLFADNNV (13: 0.8909)	TINPTFQEK (222: 0.7813)	TYLSAIKGI (86: 0.8464)	DVGEKFYFY (95: 0.7927)
P45492	YLDNLDIKFY (183: 0.9781)	KLSSDSYTL (162: 0.8464)	KIRSNEIYK (95: 0.8745)	NYLDNLDIKF (182: 0.9496)	NLKEDLLNY (205: 0.5140)

**Table 6** Prediction of antigenic epitopes through Net MHC pan 4.0 for B supertypes

	B7	B8	B27	B39	B44	B58	B62
Q0PB90	ATSTSTITL (167:0.0914)	NGKIKRSDF (1: 0.5402)	–	AHTDVGFI (29: 0.6864)	AEFKLDVI (63: 0.8640)	KIKRSDFF (158:0.4837)	LQISNVKNF (40: 0.4848)
Q0PB82	–	SMRAQSASV (93: 0.5828)	YRTISSYEV (5: 0.8454)	EQSKNNTL (112: 0.3531)	NEQSKNTL (111:0.7600)	MSNVYAF (77: 0.8339)	AQMSNVYAY (75: 0.9628)
Q0PA31	APAQALAQ (61: 0.1260)	EIQAKIVEL (93: 0.7279)	–	SKNQEQNL (25: 0.5084)	AEIQAKIEL (92: 0.9841)	KSIESQITL (114:0.5260)	–
Q0P914	–	KVKEKNTSA (19: 0.4188)	FRSYAWQM (67: 0.4681)	SADEFQAL (26: 0.5405)	NEETLLNL (83: 0.8009)	–	AVREDFRSY (62: 0.9394)
Q0P8F1	–	QIKNGTYKI (43: 0.1682)	–	–	TENDKASKI (32: 0.4912)	KASKIAEQI (36: 0.8176)	MINPIQQSY (1: 0.9279)
Q0P8D9	RAKSVANL (268: 0.5712)	KIKEIAKVL (230: 0.4036)	NRYAPGIRL (42: 0.9795)	FRLSDSLAL (134: 0.9592)	GEKFYFYL (97: 0.7079)	KTTDITRTY (79: 0.9595)	KTTDITRTY (79: 0.8374)
P45492	LPFIMFSML (358: 0.3616)	NIKIRSNEI (93: 0.8720)	–	QQDPKISSF (151: 0.3784)	YENQLNQV (30: 0.9662)	KSIQSNLVF (124: 0.9483)	SMLMGGASF (364: 0.9092)

**Table 7** Total number of epitopes selected on the basis of high VaxiJen score above 1.1

S. no.	Epitopes	HLA allele	VaxiJen
1	<b>LSDDINLNI</b>	HLA-A*01:01	2.11
2	KYPDMTFTM	HLA-A*24:03	1.14
3	<b>ATSTSTITL</b>	HLA-B*07:02	1.28
4	GKEKIGFSL	HLA-B*39:01	1.13
5	KIKRSDFKF	HLA-B*58:01	1.77
6	<b>MSNVYAYRF</b>	HLA-B*58:01	1.90
7	APAQALNAQ	HLA-B*07:02	1.29
8	ITPTLNINY(+)	HLA-A*01:06	1.45
9	KANASISIK(+)	HLA-A*30:01	1.57

alleles ( $IC_{50} \leq 500$  nM) were selected for the further analysis. The solubility of all the selected epitopes was predicted

by the ProtParam tool and molecular weight, as well as other properties, were mentioned in Table 8.

The result from toxicity prediction clearly shows that all the selected antigenic epitopes are non-toxic in nature (Table 9).

### Modeling and Refinement of the Identified Peptides and MHC Alleles

The three-dimensional structural information of proteins plays a significant role in providing the insight information on their molecular functions and identification of their binding sites. Mapping of the 3-D structure of the MHC I HLA alleles—HLA-A\*01:01, HLA-A\*24:03, HLA-B\*07:02, HLA B\*39:01, and HLA-\*58:01 with respect to its highest scoring epitopes were generated by Modeller 9.18 which used the template base model that was retrieved from Protein Data Bank. In order to refine the quality of structures



**Table 8** Prediction of vaccine candidate parameters through Prot Param tool

Epitope	Molecular weight	Instability index	pI	Hydropathicity
LSDDINLNI	1016.12	57.71	3.56	0.200
KYPDMTFTM	1133.34	24.73	5.83	−0.567
ATSTSTITL	893.99	8.89	5.57	0.633
GKEKIGFSL	978.16	−9.98	8.59	−0.200
KIKRSDFKF	1168.40	49.99	10.29	−1.156
MSNVYAYRF	1150.32	56.86	8.34	−0.078
APAQALNAQ	882.97	51.69	5.57	−0.122
ITPTLNINY(+)	1098.22	8.89	5.52	−0.478
KANASISIK (+)	931.10	−0.54	10.00	−0.033

**Table 9** Selection of vaccine candidates using MHC Pred and IEDB tools to check their binding affinity to the maximum number of HLA alleles and toxicity prediction through Toxin Pred tool

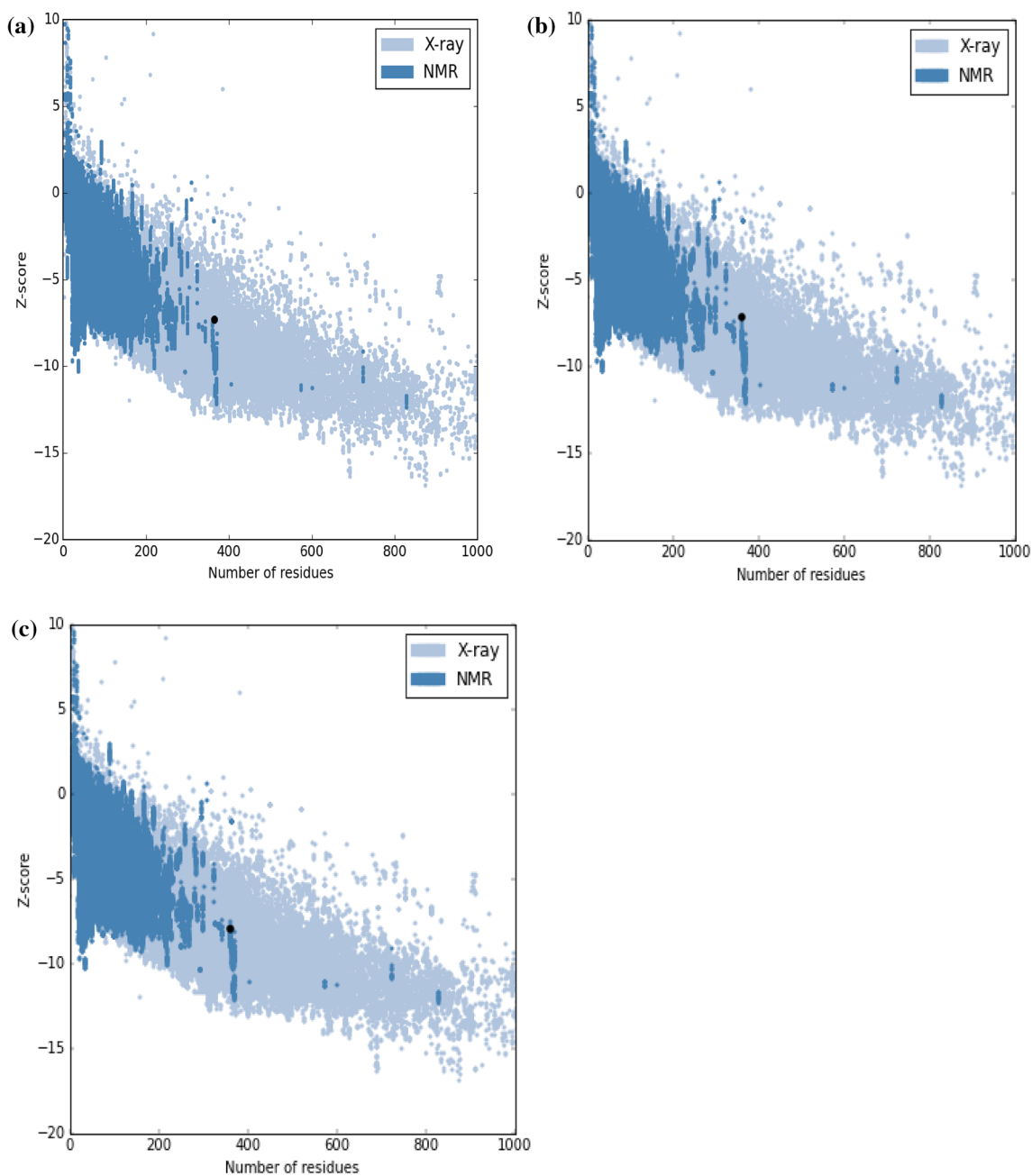
Epitope	No. of HLA binders	HLA-allele	Toxicity score	Toxicity prediction
LSDDINLNI	7	HLA-A*01:01,HLA-A*02:02,HLA-A*02:06,HLA-A*11:01,HLA-A*31:01,HLA-A*68:01, HLA-A*02:50	−1.07	Non-toxin
KYPDMTFTM	10	HLA-A*24:03, HLA-A*24:02, HLA-A*02:50, HLA-A*02:01,HLA-A*02:02, HLA-A*02:03,HLA-A*02:06,HLA-A*11:01,HLA-A*68:01,HLA-B*35:01	−0.52	Non-toxin
ATSTSTITL	10	HLA-A*02:50, HLA-B*15:02,HLA-A*01:01,HLA-A*02:02,HLA-A*02:03,HLA-A*02:06,HLA-A*03:01,HLA-A*11:01,HLA-A*68:01	−1.19	Non-toxin
GKEKIGFSL	9	HLA-A*02:50, HLA-B*15:02, HLA-A*02:02,HLA-A*02:03,HLA-A*02:06, HLA-A*03:01, HLA-A*11:01,HLA-B*35:01	−1.04	Non-toxin
KIKRSDFKF	10	HLA-A*02:50, HLA-B*15:02, HLA-A*32:01,HLA-A*02:02,HLA-A*02:06,HLA-A*03:01,HLA-A*11:01, HLA-A*68:01,HLA-B*35:01	−0.73	Non-toxin
MSNVYAYRF	12	HLA-A*32:01, HLA-A*02:50, HLA-B*15:02, HLA-B*15:01, HLA-A*26:02,HLA-A*02:03, HLA-A*03:01,HLA-A*11:01,HLA-A*31:01,HLA-A*68:01,HLA-A*68:02,HLA-B*35:01	−1.24	Non-toxin
APAQALNAQ	7	HLA-A*02:50, HLA-B*07:02, HLA-A*02:02, HLA-A*02:03,HLA-A*02:06,HLA-A*11:01,HLA-A*68:02	−0.95	Non-toxin

generated through MODELLER, these were further subjected to RAMPAGE tool. The analysis indicated that all the modeled structure's residues were > 90% in their respective favored region, and therefore approved the quality of the predicted models. Furthermore, the three selected protein models were validated through ProSA analysis (Fig. 2). Now it came to light that only segments of antigenic protein or epitopes were enough to induce the desired immune response in comparison to the whole protein (Huber et al. 2014).

### Binding Energy Calculation of the HLA Allele-Epitope Complex

Using Autodock vina, the docking of the predicted set of epitopes and HLA allele was performed. Different ten conformations were generated by Autodock vina showing the binding interaction and the best conformation was selected

on the basis of binding energy score. The lower binding energy is stronger the bonding interaction between HLA allele and epitope. The protein and the ligand preparation was done by using Autodock 4.2 tool and PDB files were converted into the PDBQT files required for docking through Autodock vina. The interactions between MSNVYAYRF binding with HLA-B\*58:01, LSDDINLNI with HLA-A\*01:01 and ATSTSTITL with HLA-B\*07:02 are shown in Table 10. The binding energies are −9.2, −7.6 and −7.4 kcal/mol, respectively. Along with these epitopes, two best peptides from the control antigens were also docked to compare the docking result, ITPTLNINY(+) shown −7.7 score and KANASISIK(+) shown −7.5 score respectively. Three epitopes with the best energies were selected and were analyzed by Chimera 1.2 tool and docked complex was generated and visualized through Chimera as shown in Figs. 3, 4 and 5.



**Fig. 2** ProSA analysis all selected protein chains in PDB determined by X-ray crystallography (light blue) or NMR spectroscopy (dark blue) with respect to their length. **a** HLA-A\*01:01 allele representing

–7.27 Z score **b** HLA-B\*07:02 allele representing –7.17 Z score **c** HLA-B\*58:01 allele representing –7.92 Z score

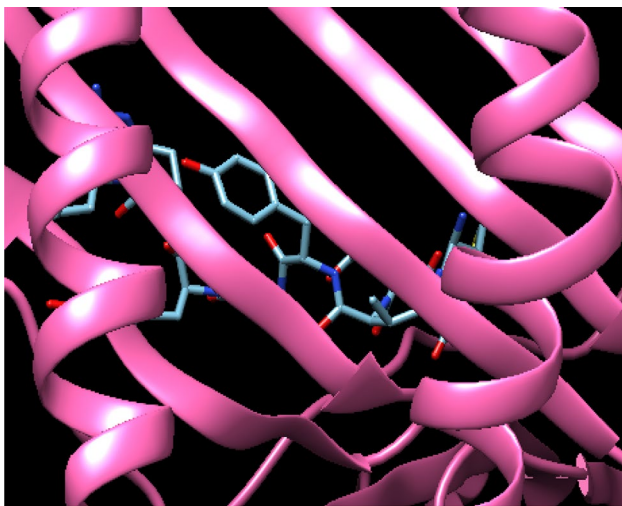
### Worldwide Population Coverage Analysis

Population coverage analysis was performed for selected epitopes along with their corresponding MHC-I alleles predicted through IEDB tool. The three predicted epitopes MSNVYAYRF, LSDDINLNI and ATSTSTITL have shown the immune response elicitation of the 56.35%, 42.41%, and 53.11% total world population. Maximum population

coverage was for observed for epitope MSNVYAYRF to be 70.91% (Fig. 6), epitope LSDDINLNI to be 56.39% (Fig. 7) and epitope ATSTSTITL to be 66.85% (Fig. 8) were found in the population of Northeast Asia. These epitopes have shown no population coverage in the region of South Africa. The results from population coverage analysis analyzed through IEDB tool clearly indicate the distribution of epitopes in the different region.

**Table 10** Binding energy calculation of the best-identified epitopes interacting with HLA alleles by using Autodock vina

Epitope	HLA allele	Binding energies (Kcal/mol)
<b>LSDDINLNI</b>	HLA-A*01:01	<b>-7.6</b>
KYPDMTFTM	HLA-A*24:03	-6.5
<b>ATSTSTITL</b>	HLA-B*07:02	<b>-7.4</b>
GKEKIGFSL	HLA-B*39:01	-6.5
KIKRSDFKF	HLA-B*58:01	-7.4
<b>MSNVYAYRF</b>	HLA-B*58:01	<b>-9.2</b>
APAQALNAQ	HLA-B*07:02	-7.2
ITPTLNyny(+)	HLA-A*01:06	-7.7
KANASISIK(+)	HLA-A*30:01	-7.5

**Fig. 3** The docked complex of epitope MSNVYAYRF with HLA allele HLA-B\*58:01 visualized through Chimera 1.12

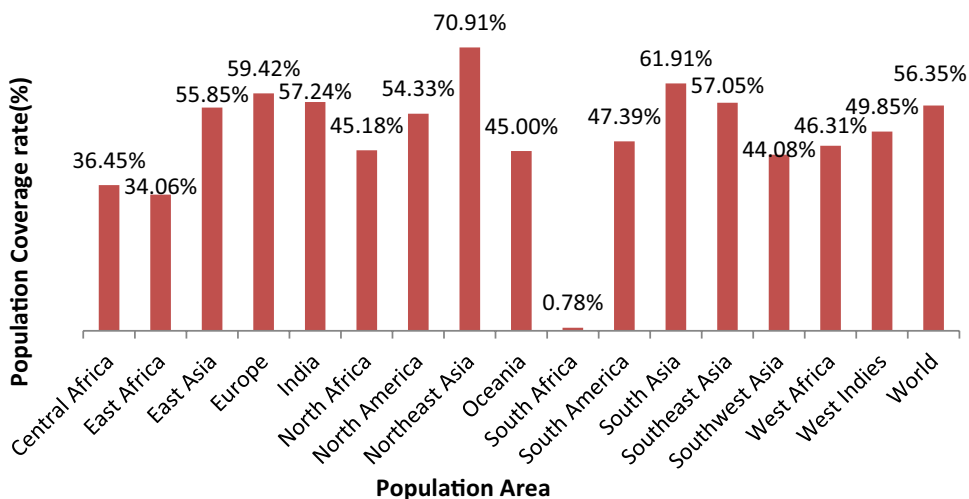
## Conclusion

The computational approach can be the helpful method in predicting the best epitopes on the basis of their antigenic nature, non-toxic score, interacting with the maximum number of HLA alleles and with higher population coverage in vaccine development. Considering all the above parameter, three epitopes MSNVYAYRF, LSDDINLNI and ATSTSTITL predicted to have the most considerable binding with HLA-B\*58:01, HLA-A\*01:01 and HLA-B\*07:02 MHC class I allele and lowest binding energy values providing stability to the peptide and MHC complex. Therefore, these peptide candidates can be further suggested for experimental laboratory analysis for the development of an effective vaccine against diarrhea. Screenings through experimental methods are time-consuming

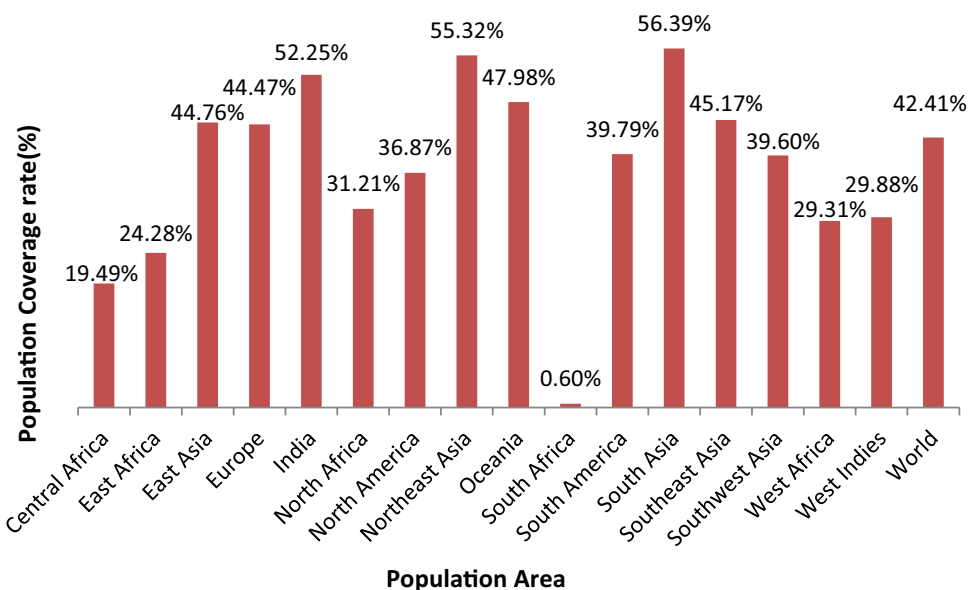
**Fig. 4** The docked complex of epitope LSDDINLNI with HLA allele HLA-A\*01:01 visualized through Chimera 1.12**Fig. 5** The docked complex of epitope ATSTSTITL with HLA allele B\*07:02 visualized through Chimera 1.12

and laborious and therefore can be replaced by reverse vaccinology based immunoinformatics approach. Therefore, the current study used in silico methods to reduce the time-consuming laboratory experiments and to avoid a hit and trial method. Thus, our findings can be helpful in developing vaccines against *C. jejuni* infection after experimentally tested in the laboratory.

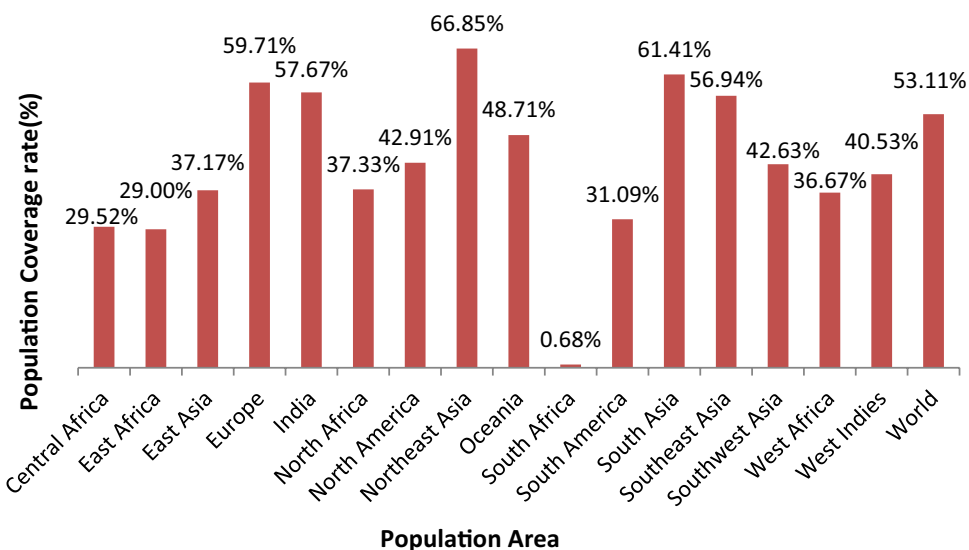
**Fig. 6** Worldwide population conservancy analysis for epitope MSNVYAYRF through IEDB server



**Fig. 7** Worldwide population conservancy analysis for epitope LSDDINLNI through IEDB server



**Fig. 8** Worldwide population conservancy analysis for epitope ATSTSTITL through IEDB server



**Acknowledgements** The authors wish to acknowledge the efforts of Dean, Faculty of Engineering and Technology, Rama University Uttar Pradesh, Kanpur for his experimental guidance during the course of experimental work. We are also grateful to supporting staff of Faculty of Engineering and Technology, Rama University Uttar Pradesh, Kanpur for providing necessary facilities.

## Compliance with Ethical Standards

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

- Ada K, Candy C, Abu B, Abdul M, Chiu H, Boon HL, Chiu YL (2018) Current progress of immunoinformatics approach harnessed for cellular- and antibody-dependent vaccine design. *Pathogens Global Health* 112:3
- Akhood BA, Slatia PS, Sharma P, Gupta SK, Verma V (2011) In silico identification of novel protective VSG antigens expressed by *Trypanosoma brucei* and an effort for designing a highly immunogenic DNA vaccine using IL-12 as adjuvant. *Microb Pathog* 51(1–2):77–87. <https://doi.org/10.1016/j.micpath.2011.01.011>
- Baloria U, Akhood BA, Gupta SK, Sharma S, Verma V (2012) In silico proteomic characterization of human epidermal growth factor receptor 2 (HER-2) for the mapping of high affinity antigenic determinants against breast cancer. *Amino Acids* 42(4):1349–1360. <https://doi.org/10.1007/s00726-00010-00830-x>
- Bijker Melief CJ, Offringa R, van der Burg SH (2007) Design and development of synthetic peptide vaccines: past, present and future. *Expert Rev Vac* 6(4):591–603. <https://doi.org/10.1586/14760584.6.4.591>
- Bui H, Sidney J, Dinh K, Southwood S, Newman MJ, Sette A (2006) Predicting population coverage of T-cell epitope-based diagnostics and vaccines. *BMC Bioinform* 7:153
- Dekeyser P, Gossuin-Detrain M, Butzler JP, Sternon J (1972) Acute enteritis due to related vibrio: first positive stool cultures. *J Infect Dis* 125:390–392
- Dimitrov I, Naneva L, Doytchinova I, Bangov I (2014) AllergenFP: allergenicity prediction by descriptor fingerprints. *Bioinformatic* 30:846–851
- Donnes P, Elofsson A (2002) Prediction of MHC class I binding peptides, using SVMHC. *BMC Bioinform* 3:25
- Doytchinova IA, Darren R, Flower (2007) VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinform* 8:4
- Flanagan RC, Neal-McKinney JM, Dhillon AS, Miller WG, Konkel ME (2009) Examination of *Campylobacter jejuni* putative adhesins leads to the identification of a new protein, designated FlpA, required for chicken colonization. *Infect Immun* 77(6):2399–2407
- Friedman CR, Neimann J, Wegener HC, Tauxe RV (2000) Epidemiology of *Campylobacter jejuni* infections in the United States and other industrialized nations. In: *Campylobacter*. ASM International, Washington, pp 121–138
- Gallardo F, Gascón J, Ruiz J, Corachan M, Jimenez de Anta M, Vila J (1998) *Campylobacter jejuni* as a cause of traveler's diarrhea: clinical features and antimicrobial susceptibility. *J Travel Med* 5(1):23–26
- Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A (2005) Protein identification and analysis tools on the Expasy server. In: Walker JM (ed) *The proteomics protocols handbook*. Humana Press, pp 571–607
- Guan P, Doytchinova IA, Zygouri C, Flower DR (2003) MHCpred: a server for quantitative prediction of peptide—MHC binding. *Nucleic Acids Res* 31:3621–3624
- Gupta SK, Smita S, Sarangi AN, Srivastava M, Akhood BA, Rahman Q, Gupta SK (2010) In silico CD4? T-cell epitope prediction and HLA distribution analysis for the potential proteins of *Neisseria meningitidis* serogroup B: a clue for vaccine development. *Vaccine* 28(43):7092–7097. <https://doi.org/10.1016/j.vaccine.2010.08.005>
- Gupta SK, Srivastava M, Akhood BA, Smita S, Schmitz U, Wolkenhauer O, Vera J, Gupta SK (2011) Identification of immunogenic consensus T-cell epitopes in globally distributed influenza-A H1N1 neuraminidase. *Infect Genet Evol* 11(2):308–319. <https://doi.org/10.1016/j.meegid.2010.10.013>
- Gupta SK, Srivastava M, Akhood BA, Gupta SK, Grabe N (2012) In silico accelerated identification of structurally conserved CD8? and CD4? T-cell epitopes in high-risk HPV types. *Infect Genet Evol* 12(7):1513–1518. <https://doi.org/10.1016/j.meegid.2012.02.022>
- Gupta S, Kapoor P, Chaudhary K, Gautam A, Kumar R (2013) In silico approach for predicting toxicity of peptides and proteins. *PLoS ONE* 8:e73957. <https://doi.org/10.1371/journal.pone.0073957>
- Huber SR, van Beek J, de Jonge J, Luytjes W, van Baarle D (2014) T Cell responses to viral infections—opportunities for peptide vaccination. *Front Immunol* 5:171
- Jin S, Joe A, Lynett J, Hani EK, Sherman P, Chan VL (2001) JlpA, a novel surface-exposed lipoprotein specific to *Campylobacter jejuni*, mediates adherence to host epithelial cells. *Mol Microbiol* 39(5):1225–1236
- Kanampalliwar AM, Soni R, Girdhar A, Tiwari A (2013) Reverse vaccinology: basics and applications. *J Vaccines Vaccin*. <https://doi.org/10.4172/2157-7560.1000194>
- Kaur H, Garg A, Raghava GPS (2007) PEPstr: a de novo method for tertiary structure prediction of small bioactive peptides. *Protein Pept Lett* 14:626–630
- Khan F, Srivastava V, Kumar A (2017) Epitope based peptide prediction from proteome of enterotoxigenic *E. coli*. *Int J Peptide Res Ther* <https://doi.org/10.1007/s10989-017-9617-1>
- Khan F, Srivastava V, Kumar A (2018) Computational identification and characterization of potential T-cell epitope for the utility of vaccine design against enterotoxigenic *Escherichia coli* <https://doi.org/10.1007/s10989-018-9671-3>
- Krogh A, Larsson B, Heijne G, Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 305:567–580
- Kumar A, Hays M, Lim F, Foster LJ, Zhou M, Zhu G, Miesner T (2015) Protective enterotoxigenic *Escherichia coli* antigens in a murine intranasal challenge model. *PLOS*. <https://doi.org/10.1371/journal.pntd.0003924>
- Larsen MV, Lundegaard C, Lamberth K, Buus S, Lund O, Nielsen M (2007) Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. *BMC Bioinform* 8:424
- Man SM (2011) The clinical importance of emerging *Campylobacter* species. *Nat Rev Gastroenterol Hepatol* 8(12):669–685
- Parkhill J, Wren BW, Mungall KL, Kettle JM, Churcher CM, Basham D, Chillingworth T, Davies RM, Feltwell T, Holroyd S, Jagels K, Karlyshev AV, Moule S, Pallen MJ, Penn CW, Quail MA, Rutherford KM, Barrell BG (2000) The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. *Nature* 2000:665–668
- Peters B, Sette A (2005) Generating quantitative models describing the sequence specificity of biological processes with the stabilized matrix method. *BMC Bioinform* 31 6:132
- Riddle S, Guerry P (2016) Status of vaccine research and development for *Campylobacter jejuni*. *Vaccine* 34(26):2903–2906

- Robinson J, Halliwell JA, Hayhurst JH, Flicek P, Parham P, Marsh SGE (2015) The IPD and IMGT/HLA database: allele variant databases. *Nucleic Acids Res* 43:D423–D431
- Sachdeva G, Kumar K, Jain P, Ramachandran S (2005) SPAAN: a software program for prediction of adhesins and adhesin-like proteins using neural networks. *Bioinformatics* 21:483–491
- Sali A, Potterton L, Feng Y, Herman V, Martin K (1995) Evaluation of comparative protein modeling by Modeller. *Protein* 23:318–326
- Sette A, Rappuoli R (2010) Reverse vaccinology: developing vaccines in the era of genomics. *Immunity* 33(4):530–541
- Singh H, Raghava GPS (2001) ProPred: prediction of HLA-DR binding sites. *Bioinformatics* 17:1236–1237
- Singh SP, Deeksha S, Bhartendu NM (2017) Genome-wide identification of novel vaccine candidates for *Plasmodium falciparum* malaria using integrative bioinformatics approaches. *3 Biotech* 7:318
- Singh SP, Khan F, Mishra BN (2010) Computational characterization of *Plasmodium falciparum* proteomic data for screening of potential vaccine candidates. *Hum Immunol* 71(2):136–143
- Srivastava M, Gupta SK, Abhilash PC, Singh N (2011) Structure prediction and binding sites analysis of curcin protein of *Jatropha curcas* using computational approaches. *J Mol Model* 18(7):2971–2979. <https://doi.org/10.1007/s00894-011-1320-0>
- Tang H, Liu XS, Fang YZ, Pan L, Zhang ZW, Zhou P, Lv JL, Jiang ST, Hu WF, Zhang P, Wang YL, Zhang YG (2012) The epitopes of foot and mouth disease. *Asian J Anim Vet Adv* 7:1261–1265
- Trott O, Olson AJ (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *J Comput Chem* 31:455–461
- Vanessa J, Paul S, Andreatta M, Marcatili P, Peters B, Morten N (2017) NetMHCpan-4.0: improved peptide–MHC class I interaction predictions integrating eluted ligand and peptide binding affinity data. *J Immunol* <https://doi.org/10.4049/jimmunol.1700893>
- Vita R, Zarebski L, Greenbaum JA, Emami H, Hoof I, Salimi N, Damle R, Sette A, Peters B (2010) The immune epitope database 2.0. *Nucleic Acids Res* 38:D854–D862
- Wallis MR (1994) The pathogenesis of *Campylobacter jejuni* Br J Biomed Sci 51(1):57–64
- Wiederstein and Sippl (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res* 35:W407–W410
- Xiang Z, He Y (2009) Vaxign: a web-based vaccine target design program for reverse vaccinology. *Procedia Vaccinol* 1(1):23–29
- Yang X, Yu X (2009) An introduction to epitope prediction methods and software. *Rev Med Virol* 19(2):77–96. <https://doi.org/10.1002/rmv.602>
- Yongqun ZX, Mobley HL (2010) Vaxign: the first web-based vaccine design program for reverse vaccinology and applications for vaccine development. *J Biomed Biotechnol* <https://doi.org/10.1155/2010/297505>
- Young KT, Davis LM, Dirita VJ (2007) *Campylobacter jejuni*: molecular biology and pathogenesis. *Nat Rev Microbiol* 5(9):665–679
- Yu NY, Wagner JR, Laird MR, Melli G, Rey S, Lo R, Dao P, Sainnapp SC, Ester M, Foster LJ, Brinkman FS (2010) PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics* 26(13):1608–1615