Infection, Genetics and Evolution 74 (2019) 103936

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Research paper

Immunoinformatics approaches for designing a novel multi epitope peptide vaccine against human norovirus (*Norwalk virus*)

Kazi Faizul Azim[a,b], Mahmudul Hasan[a,c], Md. Nazmul Hossain[a,b,*], Saneya Risa Somana[a], Syeda Farjana Hoque[a], Md. Nazmul Islam Bappy[a], Anjum Taiebah Chowdhury[a], Tahera Lasker[a]

[a] Faculty of Biotechnology and Genetic Engineering, Sylhet Agricultural University, Sylhet 3100, Bangladesh
[b] Department of Microbial Biotechnology, Sylhet Agricultural University, Sylhet 3100, Bangladesh
[c] Department of Pharmaceuticals and Industrial Biotechnology, Sylhet Agricultural University, Sylhet 3100, Bangladesh

ARTICLE INFO

Keywords:
Reverse vaccinology
Norovirus
Capsid protein VP1
Protein VP2
TLR8 receptor
Molecular docking

ABSTRACT

Norovirus is known as a major cause of several acute gastroenteritis (AGE) outbreaks each year. A study was conducted to develop a unique multi epitope subunit vaccine against human norovirus by adopting reverse vaccinology approach. The entire viral proteome of *Norwalk virus* was retrieved and allowed for further in silico study to predict highly antigenic epitopes through antigenicity, transmembrane topology screening, allergenicity assessment, toxicity analysis, population coverage analysis and molecular docking approach. Capsid protein VP1 and protein VP2 were identified as most antigenic viral proteins which generated a plethora of antigenic epitopes. Physicochemical properties and secondary structure of the designed vaccine were assessed to ensure its thermostability, hydrophilicity, theoretical PI and structural behavior. Molecular docking analysis of the refined vaccine with different MHCs and human immune TLR8 receptor demonstrated higher binding interaction as well. Complexed structure of the modeled vaccine and TLR8 showed minimal deformability at molecular level. The designed construct was reverse transcribed and adapted for *E. coli* strain K12 prior to insertion within pET28a(+) vector for its heterologous cloning and expression, and sequence of vaccine constructs showed no similarity with human proteins. However, the study could initiate in vitro and in vivo studies regarding effective vaccine development against human norovirus.

1. Introduction

Norovirus is responsible for nonbacterial gastroenteritis in adults and a major cause of several acute gastroenteritis (AGE) outbreaks each year. In children (under 5 years), noroviruses are the third most common etiological cause of diarrheal mortality after rotavirus and possibly, enteropathogenic *Escherichia coli* (Lanata et al., 2013).The prevalence of norovirus disease is approximately five times higher in children (21% per year) compared with the whole population (4.5% per year) (Phillips et al., 2010a). The symptoms of Norovirus infections are characterized by vomiting, abdominal cramps and fever. However, some elderly and young children may develop intense symptoms resulting even in death (Phillips et al., 2010b). The genus *Norovirus* has one species called *Norwalk virus* belongs to the family of Caliciviridae. Among the seven genogroups (GI-GVII), three of them (GI, GII and GVI) are responsible for most of the diseases in human (Hedlund et al., 2000; Fankhauser et al., 2002). Several reports revealed that GII is the predominant norovirus genogroup circulating worldwide and an analysis from 2005 to 2016, it was found that 91.7% of sequences were GII, 8.2% were GI, and < 0.1% were GIV (Beek et al., 2018). GII.4 is found to be the most prevalent genotype in many areas of the globe, and over the past two decades, a new GII.4 strain has emerged in every 2–4 years which sometimes involved with an increased norovirus-virulence (Siebenga et al., 2007; Siebenga et al., 2009).

Noroviruses are genetically and antigenically diverse group of single stranded RNA virus which presents serious challenges for developing a broadly immunogenic vaccine. Structurally, they are non-enveloped with a single-stranded positive sense RNA genome (7.5–7.7 kb).The norovirus genome contains three open reading frames (ORF), whose ORF1 encodes six nonstructural proteins, ORF2 encodes VP1, P2

Abbreviations: AGE, Acute Gastroenteritis; NCBI, National center for Biotechnology information; IEDB, Immune epitope database and analysis resource; CTLs, Cytotoxic T lymphocytes; MGL, Molecular graphics laboratory; MHC, Major histocompatibility complex

* Corresponding author at: Department of Microbial Biotechnology, Faculty of Biotechnology and Genetic Engineering, Sylhet Agricultural University, Sylhet 3100, Bangladesh.
E-mail address: nhossain.mib@sau.ac.bd (Md. N. Hossain).

https://doi.org/10.1016/j.meegid.2019.103936
Received 14 April 2019; Received in revised form 18 June 2019; Accepted 20 June 2019
Available online 21 June 2019

subdomain accounts for most of the antigenicity of each virus, and ORF3 responsible for encoding the minor capsid VP2 (Kroneman et al., 2013; Vinje, 2015). Currently, no licensed vaccine against norovirus has been reported (Bernstein et al., 2015). However, several norovirus vaccine candidates are under developmental stage right now which have either progressed to clinical trials or have numerous preclinical investigations (Cortes-Penfield et al., 2017; Mattison et al., 2018). Two norovirus vaccines are under clinical trial include, bivalent GI.1/GII.4 intramuscular VLP vaccines (in phase IIb clinical trial) and a monovalent GI.1 oral pill recombinant adenovirus vaccine (phase I trials). Apart from that, recently developed norovirus vaccine candidates contain non-infectious recombinant VP1 proteins as recombinant adenoviruses (Mattison et al., 2018). Researchers have developed vaccines based on expressing norovirus VP1 using recombinant adenovirus serotype 5 vectors; one is based on a GI.1 norovirus sequence (Phase I clinical trials) while the other, based on a GII.4 sequence. In a nutshell, the progress of vaccines developments highlight the importance of vaccine coverage beyond GII.4 strains and suggest candidate vaccines include representative VLPs from at least both genogroups (Aliabadi et al., 2015).

Norovirus infections are extremely difficult to prevent and control due to their environmental stability, low infectious dose and high shedding titre. There are several possible barriers to the development of an effective norovirus vaccine. Development of efficient and long-lasting norovirus vaccines has been hampered greatly by their uncultivability (Jones et al., 2015). Hence, other common vaccine technologies, including killed or live-attenuated viruses, have not been pursued for norovirus (Mattison et al., 2018). Circulating noroviruses are antigenically diverse and continually evolving, which could limit the durability of protection conferred by a vaccine that does not elicit broadly neutralizing antibodies. Immunization of mice with norovirus like particles (VLPs) has demonstrated that GI VLPs fail to produce antibodies with blocking activity against GII noroviruses and vice versa (Malm et al., 2015). Still to date, the VP1 candidate vaccine was developed but the robust immune response to heterogenous strains have not been documented in experimental studies (Cortes-Penfield et al., 2017). Therefore, it is crying need to find out the suitable polyvalent vaccine candidates against norovirus rendering immunity to several genogroups. *In-silico* study for the drug-development has opened a new era where one can design drugs using genomics, proteomics as well as immunoinformatics approach, minimizing time and cost required in development process (Mohammed et al., 2017; Dash et al., 2017; Adhikari et al., 2018; Joy et al., 2019). Various genome based technologies have enabled functionally blind selection of vaccine candidates and allowed prediction of antigens without the requirement to grow the pathogen in vitro (Rappuoli, 2000; Sette and Fikes, 2003; Agrawal and Raghava, 2018).Hence, we herein applied some immunoinformatics approaches to design a highly immunogenic and thermostable, multi-epitope peptide vaccine against norovirus utilizing their capsid proteins.

2. Materials and methods

In the present study, attempts were taken to develop a novel multi-epitope peptide vaccine against human norovirus. The schematic diagram, summarizing the protocol over in silico vaccinomics strategy to design a chimeric subunit vaccine is elucidated in Fig. 1.

2.1. Retrieval of protein sequences and antigenicity screening

The NCBI (National Center for Biotechnology Information) and ViralZone (https://viralzone.expasy.org/194) server were used for the selection and comprehensive study of human norovirus (Genome ID: 5619) (https://www.ncbi.nlm.nih.gov/genome/genomes/5619?). The entire viral proteome of human norovirus *(Norwalk virus)* was retrieved from UniProtKB (https://www.uniprot.org/uniprot/?query+database), consisting three proteins named genome polyprotein, capsid protein VP1 and protein VP2. The VaxiJen v2.0 (http://www.ddg-pharmfac.net/vaxijen/) was used for the investigation of protein immunogenicity to find the most potent antigenic proteins (Doytchinova and Flower, 2007a). Only the structural proteins were prioritized based on their antigenic score and various physicochemical parameters were analyzed using ExPASy's secondary structure prediction tool ProtParam (Gasteiger et al., 2003; Das et al., 2016).

2.2. Identification of homologous proteins sets and T-Cell epitope prediction

Homologous sequences of the antigenic proteins (i.e. capsid protein VP1 and protein VP2) were retrieved from the NCBI database by using BLASTp tool (Supplementary File 1). Both the proteins were used as query and the searches were restricted to norovirusgenogroup GI (Taxid: 122928) and GII (Taxid: 122929). Multiple sequence alignment (MSA) was performed by using Clustal Omega (Sievers and Higgins, 2014; Chenna et al., 2003), and further those common regions were employed for T-Cell epitopes prediction (Vita et al., 2014). MHC-I (http://tools.iedb.org/mhci/) and MHC-II prediction tool (http://tools.iedb.org/mhcii/) prediction tool of Immune Epitope Database were used to predict the MHC-I binding and MHC-II binding peptides respectively.

2.3. Transmembrane topology and antigenicity screening of T-Cell epitope

TMHMM (http://www.cbs.dtu.dk/services/TMHMM/) and VaxiJen v2.0 server (http://www.ddg-pharmfac.net/vaxijen/) were used to predict the transmembrane helices in proteins (Krogh et al., 2001) and determine epitope antigenicity (Doytchinova and Flower, 2007b), respectively. The most potent antigenic epitopes were selected and allowed for further analysis.

2.4. Population coverage, allergenicity and toxicity analysis of predicted epitopes

HLA distribution varies among different ethnic societies and geographic domains around the world. Population coverage for each individual epitope was analyzed by IEDB population coverage calculation tool (http://tools.iedb.org/population/) (Vita et al., 2014). The fraction of individuals predicted to respond to the proposed epitope set were calculated on the basis of HLA genotypic frequencies and MHC binding (Bui et al., 2006). Moreover, Four servers named AllergenFP (Dimitrov et al., 2014), AllerTOP (http://www.ddg-pharmfac.net/AllerTop/) (Dimitrov et al., 2013), Allermatch (http://www.allermatch.org/allermatch.py/form) (Fiers et al., 2004) and PA^{3}P (http://www.Ipa.saogabriel.unipampa.edu.br:8080/pa3p/) (Chrysostomou and Seker, 2014) were used to predict the allergenicity of the proposed epitopes for vaccine construction. Epitopes found to be non-allergenic were then allowed to assess the toxicity level by ToxinPred server (http://crdd.osdd.net/raghava/toxinpred/) (Gupta et al., 2013).

2.5. Conservancy and cluster analysis of MHC restricted alleles

Conservancy level determines the efficacy of epitope candidates to confer broad spectrum immunity. To reveal the conservancy pattern, homologous sequence sets of the selected antigenic proteins were retrieved form the NCBI database by using BLASTp tool (Supplementary File 1). Further, the epitope conservancy analysis tool (http://tools.iedb.org/conservancy/) at IEDB was selected for the analysis of conservancy pattern. Moreover, Structure dependent clustering methods have been proven efficient to identify the super-families of MHC proteins with similar binding specificities. MHC cluster v2.0 (Thomsen et al., 2013) server was used to produce pictorial tree-based graphics and generate highly instinctive heat-map of the functional association among MHC variants.

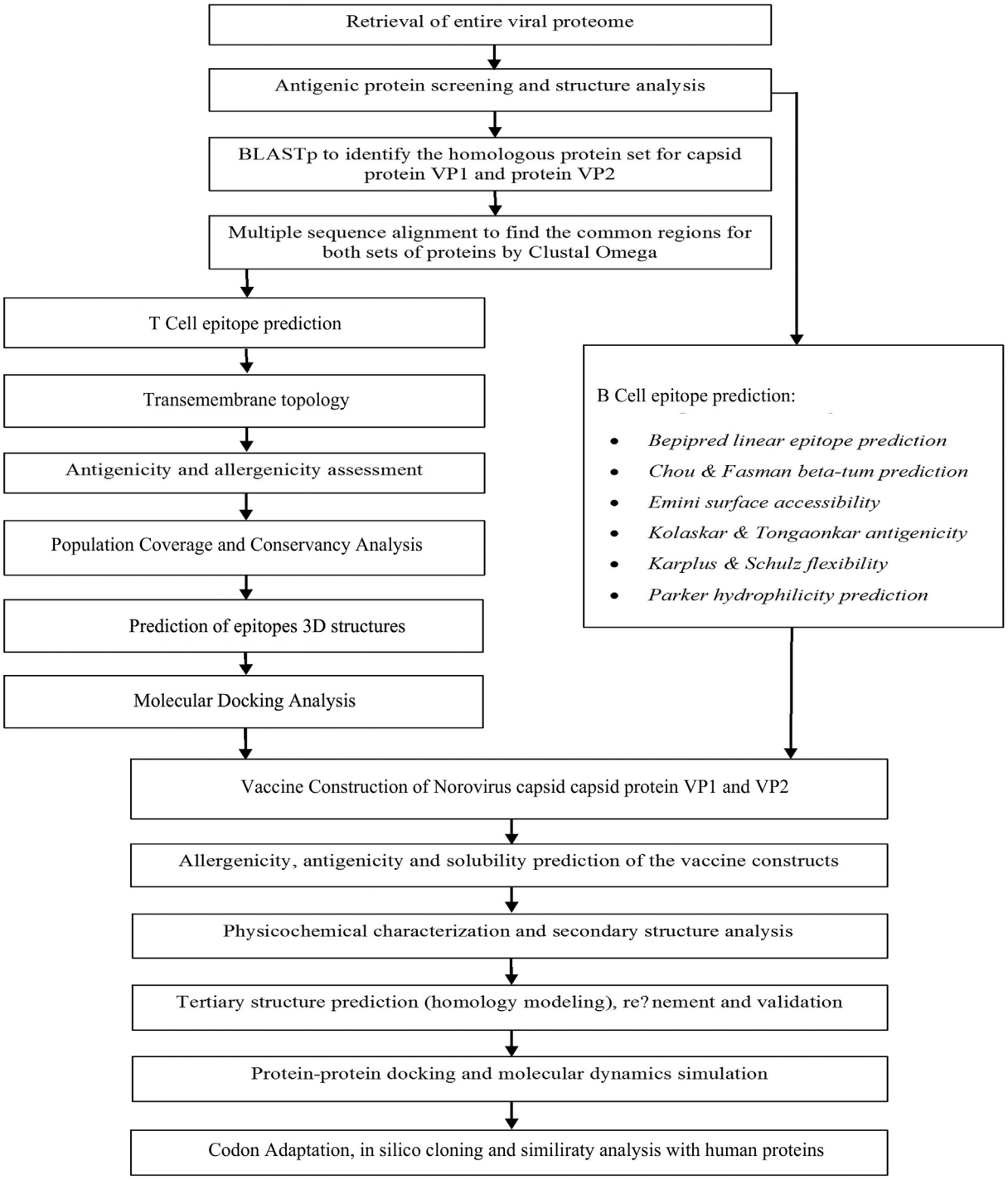

Fig. 1. Flow chart summarizing the protocols for the prediction of epitope based vaccine candidate by in silico reverse vaccinology technique.

2.6. Prediction of 3D structures for superior epitopes and analysis of molecular docking

Top ranked epitopes were subjected to PEP-FOLD server which aims to predict peptide structures (Maupetit et al., 2010; Kaur et al., 2007; Wang et al., 2011). In this study, the predicted superior CTL and HTL epitopes were able to bind with different large number of HLA alleles. Depending on the available structures deposited in Protein Data Bank (PDB) database, HLA-A*11:01 and HLA-DRB1*04:01 were selected for docking analysis with MHC class I and class II binding epitopes respectively. The purpose of the molecular docking analysis was to show the proposed epitopes could interact with at least one MHC molecule at minimum binding energy. MGLTools were used to visualize and analyze the molecular structures of biological compounds (Sanner, 1999). The grid box was set to 28 A°, 18 A°, 20 A°(x, y and z) with default value of 1.0 A° spacing by AutodockVina at 1.00- °A spacing. The exhaustiveness parameter was kept at 8.00 and the number of outputs was set at 10 (Morris et al., 2009; Sanner et al., 2002). Output PDBQT files were converted in PDB format using Open Babel. The docking interaction was visualized with PyMOL molecular graphics system (https://www.pymol.org/).

2.7. Identification of B-Cell epitope

B cell epitopes were predicted for both capsid protein VP1 and protein VP2 to find the potential antigens that would interact with B lymphocytes and initiate immune response. Several tools from IEDB were used to identify the B cell antigenicity depending on some algorithms i.e. Kolaskar and Tongaonkar antigenicity scale (Kolaskar and Tongaonkar, 1990), Emini surface accessibility prediction (Emini et al., 1985), Karplus and Schulz flexibility prediction (Karplus and Schulz, 1985), Bepipred linear epitope prediction analysis (Jespersen et al., 2017), Chou and Fasman beta turn prediction (Chou and Fasman, 1978) and Parker hydrophilicity prediction (Parker et al., 1986).

2.8. Construction of vaccine molecules

Vaccine sequences started with an adjuvant followed by the top CTL epitopes for capsid protein (VP1 and VP2), top HTL epitopes and BCL epitopes respectively, in the similar way. Three vaccine sequences were constructed (i.e. V1, V2 and V3) each associated with different adjuvants including beta defensin (a 45 mer peptide), L7/L12 ribosomal protein and HABA protein (*M. tuberculosis*, accession number: AGV15514.1) (Rana and Akhter, 2016). PADRE sequence and different linkers, for instances, EAAAK, GGGS, GPGPG and KK linkers were also used to construct effective vaccine molecules.

2.9. Allergenicity, antigenicity and solubility prediction of different vaccine constructs

AlgPred v.2.0 (Saha and Raghava, 2006) and AllerTOP v.2.0 (Dimitrov et al., 2013) servers were used to predict the non-allergic behavior of the vaccine constructs. In order to suggest the superior vaccine candidate, VaxiJen v2.0 server (Doytchinova and Flower, 2007b) was utilized for determining the probable antigenicity of the designed vaccine constructs through an alignment independent algorithm. The solubility of the proposed vaccines was analyzed via Proso II server (Smialowski et al., 2006) and Protein-sol software (Hebditch et al., 2017; Chan et al., 2013).

2.10. Physicochemical characterization and secondary structure analysis

ProtParam, a tool from Expasy's server (http://expasy.org/cgi-bin/protpraram)(Gasteiger et al., 2003; Hasan et al., 2015a) was used to functionally characterize the vaccine proteins. Physicochemical properties including molecular weight, aliphatic index, isoelectric pH, hydropathicity, GRAVY values, instability index and estimated half-life were investigated. The PSIPRED v3.3 (Kosciolek and Jones, 2014) and NetTurnP 1.0 program (Petersen et al., 2010; Thaysen-Andersen and Packer, 2012) were used to predict the alpha helix, beta sheet and coil structure of the vaccine protein.

2.11. Prediction, refinement, validation and disulfide engineering of vaccine constructs

The RaptorX server (Kallberg et al., 2014; Peng and Xu, 2011) was applied for determining 3D structure of designed vaccine constructs based on the degree of similarity between target protein and available template structure from PDB. Refinement was performed to enhance the accuracy of the predicted structures using ModRefiner (Xu and Zhang, 2011) followed by FG-MD refinement server (Zhang et al., 2011). The refined protein structure was validated through Ramachandran plot assessment at RAMPAGE (Lovell et al., 2002; Hasan et al., 2015b). Again, an online tool, DbD2 was used to design disulfide bonds for the predicted vaccine constructs (Craig and Dombkowski, 2013). Residues in the highly mobile region of the protein exhibit the potential to be mutated with cysteine that allowed disulfide engineering of the refined structure. Pairs of residues with proper geometry and ability to form a disulfide bond were detected while individual amino acids mutated to cysteine.

2.12. Protein-protein docking and molecular dynamics simulation

Owing to determine the binding affinity of the vaccine constructs with different HLA alleles and TLR-8 immune receptor, ClusPro 2.0 (Comeau et al., 2004), hdoc (Macalino et al., 2018; Extra-Kangueane and Nilofer, 2018) and PatchDock server (Schneidman-Duhovny et al., 2005) were applied. Inflammation mediated by ssRNA virus are involved with immune receptors present over the lymphocytes mainly by TLR-7 and TLR-8 (Heil et al., 2004; Cros et al., 2010). The 3D structure of human TLR-8 immune receptor was retrieved from RCSB protein data bank. Desirable complexes were identified according better electrostatic interaction and free binding energy following refinement via FireDock server. The iMODS server were used to explain the collective motion of proteins via analysis of normal modes (NMA) in internal coordinates (Lopez-Blanco et al., 2014). Essential dynamics is a powerful tool and alternative to costly atomistic simulation that can be compared to the normal modes of proteins to determine their stability (Aalten et al., 1997; Wuthrich et al., 1980; Cui and Bahar, 2007). The server predicted the direction and magnitude of the immanent motions of the complex in terms of deformability, eigenvalues, B-factors and covariance. Structural dynamics of the protein-protein complex was investigated (Awan et al., 2017; Prabhakar et al., 2016).

2.13. Codon adaptation, in silico cloning and similarity analysis with human proteins

E. coli strain K12 was selected as host for the cloning purpose of the vaccine construct V1. Due to the dissimilarity between the codon usage of human and *E. coli,* a codon adaptation tool (JCAT) was used to adapt the codon usage to the well characterized prokaryotic organisms for accelerating the expression rate in it. Rho independent transcription termination, prokaryote ribosome binding site and cleavage sites of restriction enzyme *Bgl*II and Apa1 were avoided while using the server (Grote et al., 2005). The optimized sequence of vaccine protein V1 was reversed followed by conjugation with BglII and Apa1 restriction site at the N-terminal and C-terminal sites respectively. SnapGene (Solanki and Tiwari, 2018) restriction cloning module was used to insert the adapted sequence into pET28a(+) vector between the BglII (401) and Apa1(1334). At last, human *sequence similarity analysis of the proposed vaccine with human proteins was done by using NCBI protein-protein Blast (*https://blast.ncbi.nlm.nih.gov/Blast.cgi*)*, and here blast was done against *Homo sapiens* (taxid: 9606) dataset.

3. Results

3.1. Retrieval of protein sequences and antigenicity screening

The entire viral proteomeof human norovirus (*Norwalk virus*) (Proteome ID: UP000000826) was retrieved from UniProtKB (https://www.uniprot.org/uniprot/?query+database). All the sequences belong to the origin of United States. Only the structural proteins were selected for further investigation. The VaxiJen server was used to find out the most potent antigenic proteins. Among three viral proteins, capsid proteinVP1 (Accession ID: Q83884) and proteinVP2 (Accession ID: Q83885) were identified, having better immunogenic potential with total prediction score of 0.555 and 0.510 respectively. Different physiochemical behaviors of the proteins were analyzed by ProtParam tool as described in Supplementary Table 1.

3.2. Identification of homologous proteins sets and T-Cell epitope prediction

Two sets of homologous proteins, one for capsid protein VP1 and another for protein VP2 were retrieved from the NCBI protein database

after BLASTp search (Supplementary File 1). The retrieved protein sequences belonged to the norovirus genogroup GI and GII. Multiple sequence alignment tool of Clustal Omega determined the common fragments with varying lengths from both sets of proteins which were further utilized to predict T-Cell epitopes for vaccine design. A plethora of immunogenic epitopes were identified to be potential T cell epitopes for capsid protein VP1(Supplementary File 2) and protein VP2 (Supplementary File 3) that can bind to different large number of HLA-A and HLA-B alleles using MHC class-I and MHC class II binding predictions of IEDB. Epitopes that could interact with maximum number of HLAs with a high binding affinity were selected.

3.3. Transmembrane topology and antigenicity screening of T-Cell epitope

Top epitopes (MHC-I and MHC-II binding peptides) for both proteins having the capacity to elicit strong T-cell response were selected as putative T cell epitope candidates according to their topology screening by TMHMM and antigenic scoring (AS) by Vaxijen server (Supplementary Table 2).

3.4. Population coverage, allergenicity and toxicity analysis of predicted epitopes

Population coverage was demonstrated for each individual epitope by IEDB population coverage calculation tool analysis resource. Here, all the indicated alleles were identified as optimum binders with the suggested epitopes and used to determine the population area covered by them (Fig. 2). Two different population coverages were calculated from CTL and HTL populations for MHC class I and MHC class II restricted peptides respectively. Epitopes, found to be non-allergen for human were identified according to the allergenicity assessment via 4 servers (i.e. AllerTOP, AllergenFP, PA^{3}P, Allermatch) (Supplementary File 4). However, epitopes, predicted as toxin by ToxinPred server were removed from the predicted list of epitopes (Table 1).

3.5. Conservancy and cluster analysis of MHC restricted alleles

Several epitope candidates from both capsid protein VP1 and VP2 were found to be highly conserved within different strains of human *Norovirus* with maximum conservancy level of 100% (Table 1). The epitopes, showing conservancy pattern at a biologically significant level, were only allowed for further docking analysis. Clustering of both class I and class II HLA molecules identified to interact with the predicted epitopes were generated by MHCcluster v2.0 server. The server utilized the conventional phylogenetic method to generate the output (heat map), based on sequence data available for different HLA alleles (Supplementary Fig. 1).

3.6. Prediction of 3D structures for superior epitopes and analysis of molecular docking

3D structures were predicted for top epitopes (6 from capsid protein VP1 and 6 from VP2) to analyze their interactions with different HLA alleles. The PEP-FOLD3 server modeled five 3D structures for each individual epitopes and the best one was identified for docking study. Result showed that 'FDLSLGPHLNPFLLH' epitope of capsid protein VP1 bound in the groove of the HLA-DRB1*04:01 with an energy of −8.7 kcal/mol. The demonstrated energy was −8.9 kcal/mol for epitope 'VPIPFASKQKQVQSS' containing the 9-mer core 'VPIPFASKQ' of protein VP2. On the contrary VP1-epitope 'SCIPPGFGS' was found to be superior in terms free binding energy while interacted with HLA-A*11:01 (−8.8 kcal/mol) (Table 2).

3.7. Identification of B-Cell epitope

B-cell epitopes of both capsid protein VP1 and protein VP2 were generated using six different algorithms from IEDB. For capsid protein VP1, Bepipred linear prediction method predicted that the peptide sequences from 221 to 247 and 510–527 amino acids are able to induce

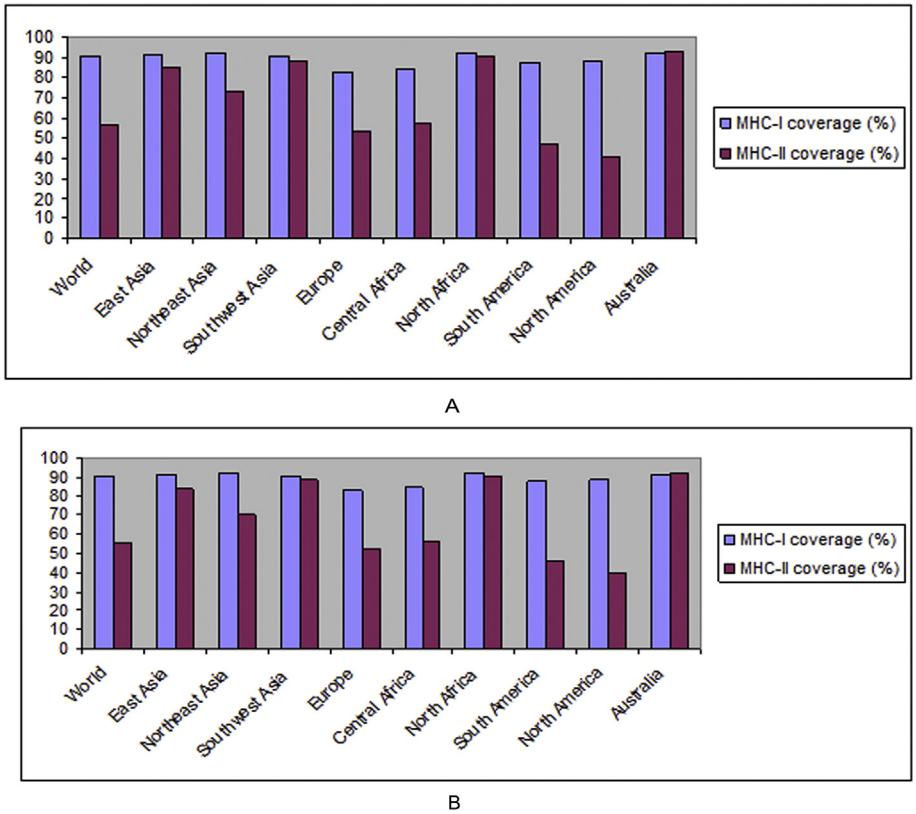

Fig. 2. Population coverage analysis of capsid protein VP1 (A) and protein VP2 (B).

Table 1
Allergenicity assessment, toxicity test and conservancy analysis of the predicted epitopes generated from protein VP1 and protein VP2.

Capsid protein VP1				Protein VP2			
Epitope (Outer Membrane)	Allergenicity	Toxicity	Conservancy	Epitope (Outer Membrane)	Allergenicity	Toxicity	Conservancy
PDFNFLFLV	Non Allergen	Non Toxin	100%	SSRFGNLSP	Non Allergen	Non Toxin	26.6%
FDLSLGPHL	Non Allergen	Non Toxin	100%	VPIPFASKQ	Non Allergen	Non Toxin	83.3%
VLFDLSLGP	Non Allergen	Non Toxin	100%	PNYSPSSIS	Non Allergen	Non Toxin	86.6%
LFDLSLGPH	Non Allergen	Non Toxin	100.%	ALGAGIQVG	Non Allergen	Non Toxin	100%
SCIPPGFGS	Non Allergen	Non Toxin	98.0%	FGNLSPYHA	Non Allergen	Non Toxin	56.6%
FDLSLGPHLNPFLLH	Non Allergen	Non Toxin	100%	NSSRFGNLSPYHAEA	Non Allergen	Non Toxin	26.6%
SCIPPGFGSHNLTIA	Non Allergen	Non Toxin	95.5%	VESQNSSRFGNLSPY	Non Allergen	Non Toxin	30.0%
IVSCIPPGFGSHNLT	Non Allergen	Non Toxin	38.5%	SNPNYSPSSISRTTS	Non Allergen	Non Toxin	60.0%
PFLSLGPHLNLLHLS	Non Allergen	Non Toxin	98.0%	VPIPFASKQKQVQSS	Non Allergen	Non Toxin	56.6%
VSCIPPGFGSHNLTI	Non Allergen	Non Toxin	95.5%	GFMSVPIPFASKQKQ	Non Allergen	Non Toxin	83.3%

the strong immune responses as B cell epitopes. Emini surface accessibility prediction was done which indicated 223–232 and 462–471 amino acid residues to be more accessible. Chou and Fasman beta-turn prediction method predicted regions from 73 to 79 and 193–199 as potential Beta-turn regions. Karplus and Schulz flexibility prediction method displayed that the region of 190–196 and 488–494 amino acid residues are the most flexible regions. Kolaskar and Tongaonkar antigenicity result confirmed the region from 233 to 241 and 497–516 as highly antigenic while Parker hydrophilicity prediction indicated 193–199 and 518–524 amino acid residues to be more potent (Fig. 3). In case of VP2, bepipred linear prediction method predicted that the peptide sequences from 32 to 43 and 107–169 amino acids are able to induce better immune responses as B cell epitopes. The regions from 31 to 50 and 128–134 amino acid residues were more accessible according to Emini surface accessibility prediction algorithm. Chou and Fasman beta-turn prediction method predicted regions from 140 to 146 and 183–189 as potential Beta-turn regions. Karplus and Schulz flexibility prediction method indicated the region of 30–36 and 128–134 amino acid residues as a most flexible region. Kolaskar and Tongaonkar antigenicity result confirmed the region from 86 to 97 & 188–197 as highly antigenic while Parker hydrophilicity prediction indicated 45–51 and 157–163 amino acid residues to be more preferred (Fig. 3). Allergenicity pattern of the predicted B-cell epitopes is shown in Supplementary Table 3.

3.8. Construction of vaccine molecules

A total 3 constructs were designed in the study, each comprising a protein adjuvant and PADRE peptide sequence. The rest of the constructs were occupied by the T-cell and B-cell epitopes and their respective linkers. PADRE sequence was also introduced to enhance the efficacy and potency of the vaccine proteins. Each constructs consist of 12 CTL epitopes and 10 BCL epitopes where CTL, HTL and BCL epitopes were conjugated together by GGGS, GPGPG and KK linkers respectively (Supplementary Table 4). The designed vaccine constructs V1, V2 and V3 were403, 488 and 517 residues long.

3.9. Allergenicity, antigenicity and solubility prediction of different vaccine constructs

AlgPred server demonstrated the non-allergic behavior of the final vaccine constructs. Vaccine construct V1 with beta-defensin adjuvant and construct V3 with HBHA adjuvant were found to be non-allergic in nature. Immunogenicity of the vaccine proteins was further determined using VaxiJen 2.0 server. Results indicated V1 as the most potent vaccine candidate with better antigenic nature (0.60) that can elicit strong immune response (Supplementary Table 4). All three constructs showed solubility above threshold value 0.6 (0.77, 0.73, 0.68 for V1, V2 and V3 respectively). However, Proso II server revealed that construct V1 has the best chance to be soluble upon over expression in *E. coli* which was further confirmed via Protein-sol software. The server calculated the surface distribution of charge, hydrophocity and the stability at 91 different combinations of pH and ionic strength.

3.10. Physicochemical characterization and secondary structure analysis

The ProtParam tool from Expasy server was used for the characterization of final vaccine protein according to different physicochemical parameters. Molecular weight of the designed construct V1 was 41.67 kDa which confirmed its good immunogenic potential. The theoretical pI 10.33 implied that the protein will have net negative charge above this pI and vice versa. The extinction coefficient was 19,940 at 0.1% absorption, assuming all cysteine residues are reduced. The estimated half-life was predicted to be 1 h within mammalian reticulocytes in vitro while > 10 h in *E. coli* in vivo. Hydrophilic behavior and thermostability of the protein was represented by the GRAVY value and aliphatic index which were − 0.499 and 60.82 respectively. Various physicochemical features classified the protein as a stable one with the capacity to initialize a strong immunogenic reaction in the body. The predicted structure secondary confirmed to have 21.33% alpha helix, 7.7% sheet and 70.9% coil structure (Supplementary Fig. 2).

Table 2
Binding energy of suggested T-cell epitopes with selected class I and class II MHC molecules generated from molecular docking analysis.

Epitopes	MHC allele	Binding energy (kcal/mol)	Epitopes	MHC allele	Binding energy (kcal/mol)
PDFNFLFLV	MHC-A*11:01	−8.6	FDLSLGPHLNPFLLH	HLA-DRB1*04:01	−8.7
FDLSLGPHL		−8.7	SCIPPGFGSHNLTIA		−8.2
SCIPPGFGS		−8.8	PFLSLGPHLNLLHLS		−7.7
VPIPFASKQ		−8.4	SNPNYSPSSISRTTS		−8.5
PNYSPSSIS		−8.4	VPIPFASKQKQVQSS		−8.9
ALGAGIQVG		−8.3	GFMSVPIPFASKQKQ		−7.6

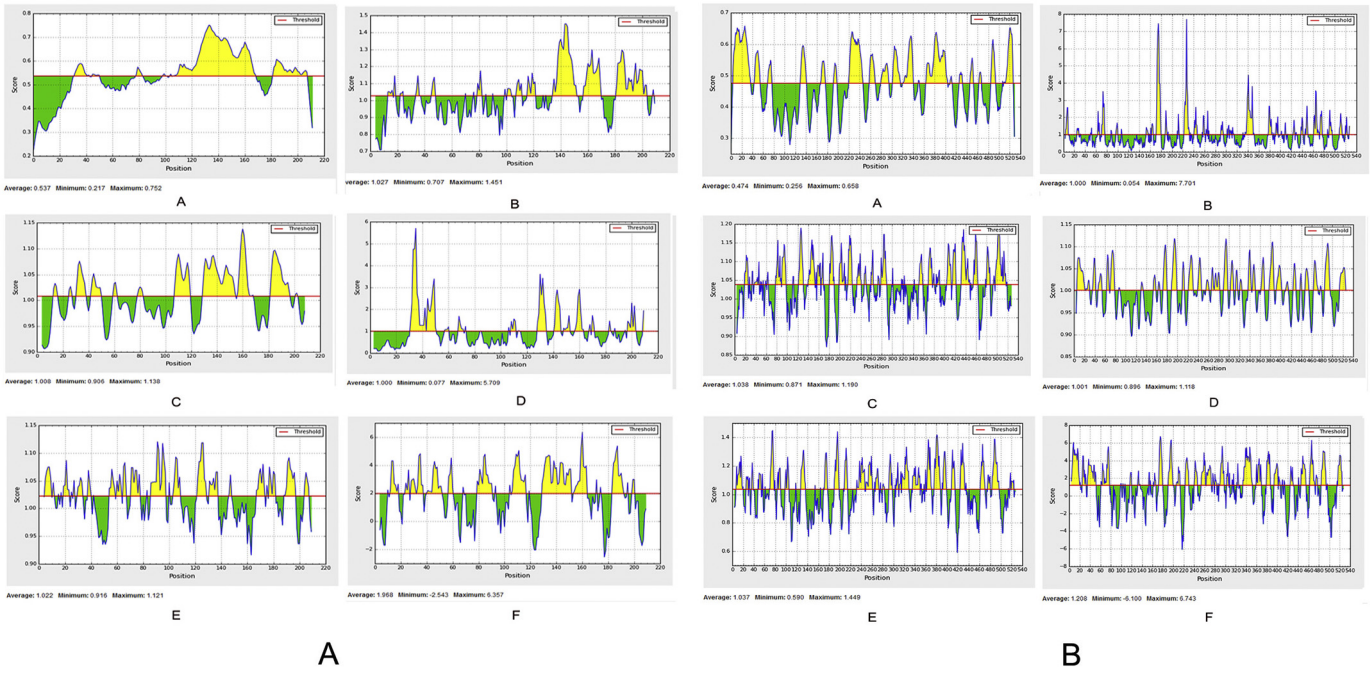

Fig. 3. Prediction of B cell epitopes and intrinsic properties for protein VP1 (A) and VP2 (B) using different scales (A: Linear, B: Beta-turn, C: Flexibility, D: Surface Accessibility, E: Antigenicity, F:Hydrophilicity). For each graph, X-axis and Y-axis represent the position and score. Residues that fall above the threshold value are shown in yellow color while the highest peak in yellow coloridentifies most favored position. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.11. Prediction, refinement, validation and disulfide engineering of vaccine constructs

RaptorX predicted the tertiary structure of the construct V1 composed of 2 domains. Homology modeling was performed by detecting and using 1kj6A from RCSB Protein Data Bank (http://www.rcsb.org/structure/1KJ6) as best suited template for Vaccine protein V1 (Fig. 4A and Fig. 4B). All 403 amino acid residues were modeled with only a few residues in the disordered region. The quality of the 3D model was defined by *P* value which was $2.03e^{-03}$. The lower P value ensured good quality of the modeled structure. Refinement was performed using ModRefiner followed by FG-MD refinement server to enhance the accuracy of 3D modeled structure of predicted vaccine V1. During Ramachandran plot analysis, 93% residues were in the favored, 5.1% residues in the allowed and 1.9% residues in the outlier region prior to refinement. However, after refinement 98.1% and 1.9% residues were in the favored and allowed region respectively while no residues occupied in the outlier region (Fig. 4C). Modeled tertiary structure of designed construct V2 and V3 have been shown in Supplementary Fig. 3. A total 18 pairs of amino acid residues were identified with the potential to form disulfide bonds by DbD2 server. However, after investigating them according to energy, chi3 and B-factor parameter, only 2 pairs (i.e. PRO 318-LYS 350, TYR 340-SER 344) were compatible for disulfide bond formation (Supplementary Fig. 4). All these residues were replaced with cysteine residue. The value of chi3 considered for the residue screening was between −87 to +97 while the energy considered was < 2.2.

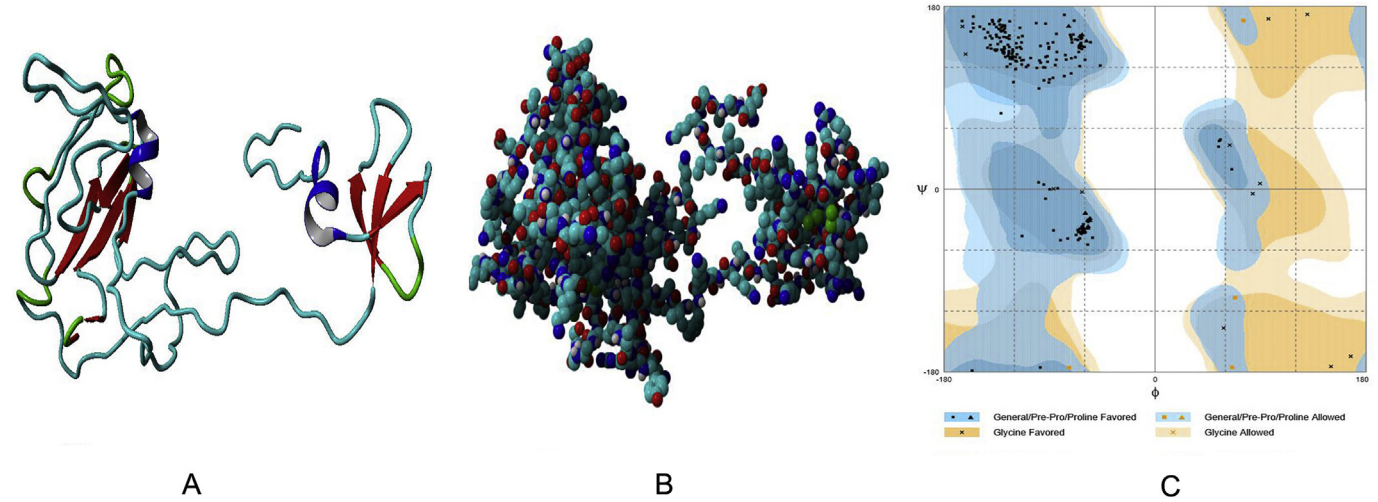

Fig. 4. Tertiary structure prediction and validation of vaccine protein V1, (A) Cartoon format, (B) Ball structure, (C) Validation of the 3D structure of vaccine protein V1 by Ramachandran plot analysis.

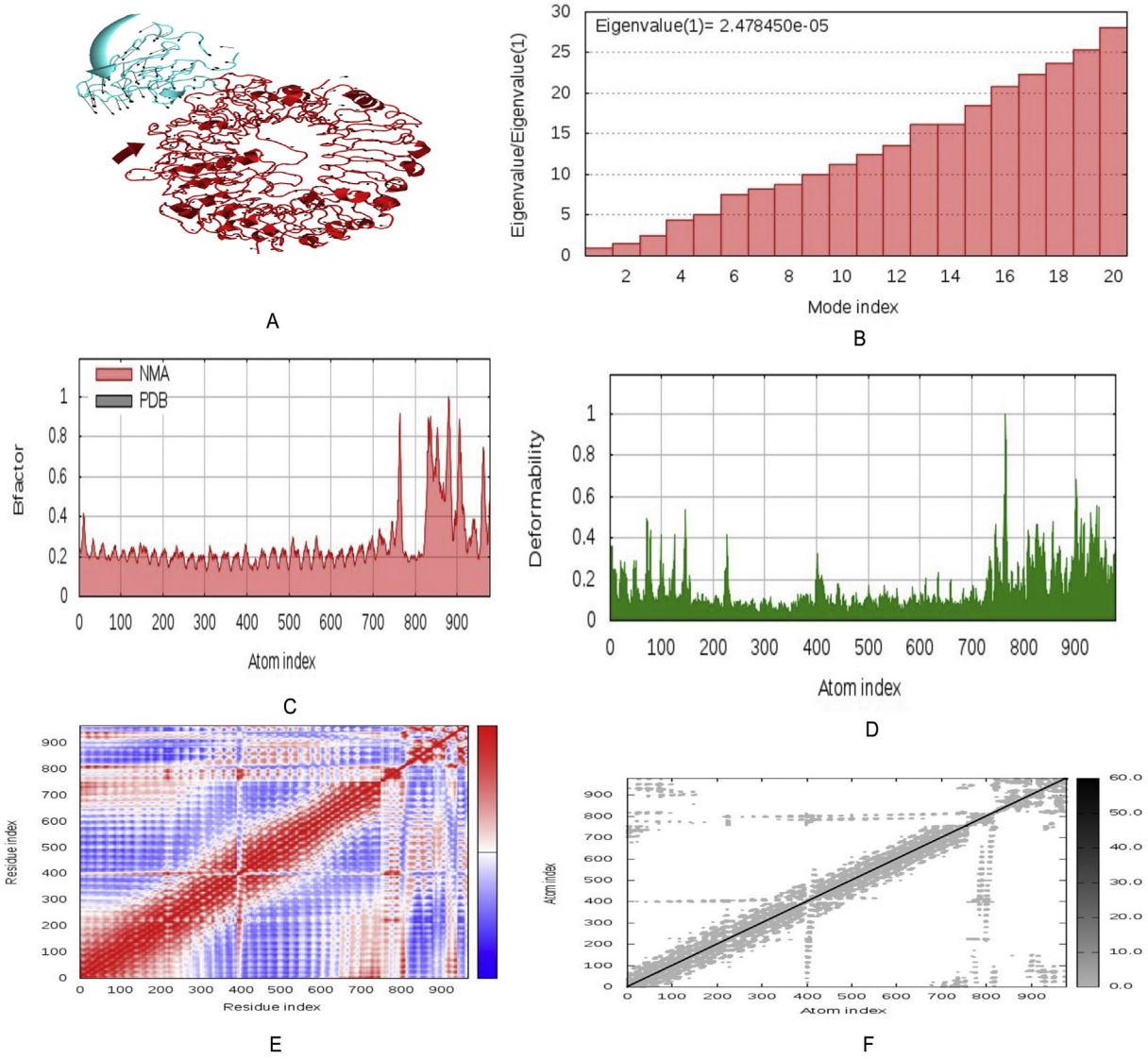

Fig. 5. Molecular dynamics simulation of vaccine protein V1-TLR8 complex; stability of the protein-protein complex was investigated through mobility (A), eigenvalue (B), B-factor (C), deformability (D), covariance (E) and elastic network (F) analysis.

3.12. Protein-protein docking and molecular dynamics simulation

Docking study was conducted between three vaccine constructs (i.e. V1, V2, V3) and different HLA alleles. Construct V1 showed biologically significant results and found to be superior in terms of free binding energy (Table 3). Besides, docking was also performed to evaluate the binding affinity of predicted vaccine V1 with TLR8 immune receptor of human via ClusPro, HDOC and PatchDock web servers. ClusPro generated thirty protein-ligand complexes (clusters) as output along with respective free binding energy. The lowest energy was −1183.9 for cluster 3 (Supplementary Fig. 5). The HDOC server predicted the binding energy for the docked complex was −300, while FireDock output refinement of PatchDock server showed the lowest global energy of −8.52 for solution 10. Normal mode analysis allowed the demonstration of large scale mobility andthe stability of proteins. The iMODS server performed such analysis based on internal coordinates of the protein-protein complex (Fig. 5A). In the 3D model, the direction of each residue was given by arrows and the length of the line represented the extent of mobility. The vaccine protein V1 and TLR8 were oriented towards each other. The B-factor values deduced from normal mode analysis, was analogous to RMS (Fig. 5C). Hinges in the chain indicated the probable deformabilty of the complex measured by contortion of each individual residues (Fig. 5D). The eigenvalue found for the complex was $2.4784e^{-05}$ (Fig. 5B). The variance associated to each normal mode was inversely linked to the eigenvalue (Kovacs et al., 2004). Covariance matrix explained the coupling between pairs of residues where correlated, uncorrelated or anti-correlated motions were represented via red, white and blue colors respectively (Fig. 5E). The result also generated an elastic network model (Fig. 5F) which identified the pairs of atoms connected via springs. Each dot in the diagram was indicated one spring between the corresponding pair of atoms and colored based on the degree of stiffness.

3.13. Codon adaptation, in silico cloning and similarity analysis with human proteins

Codon adaptation was performed due to the difference in the protein expression system of human and *E. coli*. Vaccine protein V1 was transcribed reversely. Codon adaptation index (CAI) of the adapted codons indicated the higher proportion of most abundant codons. The CAI was found 0.951 and the GC content of the optimized codons (53.43%) was also significant. The construct did not hold restriction sites for *Bgl*II and *Apa*I which ensured its safety for the cloning purposes. The optimzed codons were incorporated into pET28a(+) vector

Table 3
Docking score of vaccine construct V1, V2 and V3 with different HLA alleles including HLA-DRB3*02:02 (3C5J), HLA-DRB5*01:01 (1H15), HLA-DRB1*01:01 (2FSE), HLA-DRB3*01:01 (2Q6W), HLA-DRB1*04:01 (2SEB) and HLA-DRB1*03:01 (1A6A).

Vaccine construct	HLA alleles PDB ID's	Global energy	Hydrogen bond energy	ACE	Score	Area
V1	3C5J	−30.45	−5.78	8.80	17,036	2246.70
	1H15	−41.18	−9.20	7.58	16,716	2620.90
	2FSE	−5.43	−1.95	0.71	19,796	3091.6
	2Q6W	−13.83	−2.87	1.87	16,404	2145.60
	2SEB	−21.77	−0.82	8.21	16,912	2389.6
	1A6A	−0.68	−3.80	2.77	15,836	2577.8
V2	3C5J	−2.92	0.00	0.56	22,456	5683.1
	1H15	3.54	−0.32	0.92	17,880	3565.6
	2FSE	−24.52	−2.16	−0.86	18,200	2816.5
	2Q6W	−12.61	−2.82	4.38	17,146	2568.6
	2SEB	−27.91	−1.98	5.69	20,392	2834.9
	1A6A	9.58	0.00	0.13	17,702	2408.4
V3	3C5J	−23.12	−4.23	6.93	20,378	2797.1
	1H15	−19.97	−4.70	4.00	23,418	3743.5
	2FSE	−28.33	−6.67	3.84	19,880	3540.5
	2Q6W	−33.22	−0.38	−2.16	18,280	2446.0
	2SEB	−6.22	−1.45	3.36	19,006	2764.3
	1A6A	−9.55	−1.12	6.28	18,558	2498.8

along with BglII and ApaI restriction sites. A clone of 5655 base pair was obtained including the 1219 bp desired sequence and the rest belonging to the vector. The desired region was shown in red color in between the pET28a(+) vector sequence (Supplementary Fig. 6). Sequence similarity analysis of the proposed vaccine with human proteins revealed that there was no similarity between predicted vaccine constructs and human proteins (Supplementary File 5).

4. Discussion

The importance of developing new vaccines for combating the ever rising global burden of disease cannot be ignored. Norovirus is an emerging virus that may be attributed to the leading cause of viral gastroenteritis in people of all ages worldwide. This highly infectious virus cause significant secondary spread of infection within close communities. Therefore, it is important to take preventive measures against it. To date, there is no specific treatment for norovirus (Bernstein et al., 2015). In this study, we aimed to design an epitope based peptide vaccine against human norovirus through immunoinformatics strategies by utilizing various bioinformatics tools. Advances in genomics have changed the concepts and approaches to vaccine candidate selection and design (Ahluwalia et al., 2017). Now, it is possible to use genomic based approaches to aid selection of vaccine candidates and structure based design to optimize the chosen immunogens. Conventional approaches for vaccine development are not always feasible as microorganisms are sometimes difficult to cultivate and in some instances to attenuate resulting in undesirable or adverse immune responses (Purcell et al., 2007). Moreover, most of the methods so far utilized to acquire and purge the target antigen were unsuccessful, resulting in less suitable vaccine candidates (Rappuoli, 2000). That is why vaccine development is now being shifted to the most effective and less time consuming prescreening programs (Mora et al., 2003; Rappuoli et al., 2016).Reverse vaccinology, a novel approach to combine immunogenomics and immunogenetics with bioinformatics has been used tremendously to introduce new vaccines (Poland et al., 2009; Hasan et al., 2019a).

In the present study, the entire viral proteome of human norovirus (*Norwalk virus*) was retrieved from UniProtKB and the physiochemical properties of the proteins were analyzed. The VaxiJen server was used to assess the antigenicity of all the retrieved protein sequences in order to find out the most potent antigenic protein. Among the three viral proteins, capsid protein VP1 (Accession ID: Q83884) and protein VP2 (Accession ID: Q83885) were identified as the best immunogenic protein candidates and allowed for further analysis. GI and GII contain majority of the norovirus strains associated with the human diseases (Bok and Green, 2012).Hence, the homologous protein sequences for both proteins belonging to these two genogroups were retrieved via BLASTp analysis. Multiple sequence alignment (MSA) enabled the identification of common regions for both sets of protein. Only the common fragments were used to predict the CTL and HTL epitopes in this study. Most antigens trigger both, B cell and T cell response (Hasan et al., 2019b). Vaccine induces production of antibodies that are synthesized by B cells and mediates effector functions by binding specifically to a toxin or a pathogen (Cooper and Nemerow, 1984; Hossain and Rahman, 2014). Cytotoxic CD8 + T lymphocytes (CTL) restrict the spread of infectious agents by recognizing and killing infected cells or secreting specific antiviral cytokines (Garcia et al., 1999; Sharmeen et al., 2012). Thus, T-Cell epitope-based vaccination is a unique process of eliciting strong immune response against infectious agents such as viruses (Shrestha, 2004). Numerous immunogenic epitopes of capsid protein VP1 and protein VP2 were generated to be potential T-Cell epitopes that can bind a large number of HLA-A and HLA-B alleles with a very high binding affinity using the MHC-I and MHC-II binding predictions of IEDB with recommended methods. Top ten epitopes, bound with the highest number of HLA alleles were selected as putative T cell epitope candidates (Table 1) based on their protein transmembrane topology screening (Krogh et al., 2001) and VaxiJen score (Doytchinova and Flower, 2007b).

Allergenicity is one of the prominent obstacles in vaccine development. Today, most vaccines stimulate the immune system into an allergic reaction (McKeever et al., 2004). According to the WHO/FAO, if a sequence has an identity of at least six contiguous amino acids over a window of 80 amino acids (0.35% sequence identity) to a known allergens), it is considered to be potentially allergenic. In this study, AllerTOP, AllergenFP, PA^{3}P, Allermatch were used for the allergenicity assessment of the predicted T-cell epitopes of both proteins. Population coverage is another important factor in the development of a vaccine. Multiple population coverages were simultaneously generated and average population coverage was demonstrated. Results showed that > 90% population of the world can be covered by the predicted T-cell epitopes of capsid protein VP1 and protein VP2. As MHC super families play a vital role in vaccine design and drug development, MHC cluster analysis was also performed to determine the functional relationship between MHC variants. To ensure effective binding between predicted epitope and HLA molecules, a docking study was performed using MGLTools. HLA-A*11:01 and HLA-DRB1*04:01 was selected for docking analysis with MHC class I and class II binding epitopes respectively. All the finalized epitopes showed a lower binding energy which was biologically significant (Table 3). CTL epitope 'SCIPPGFGS' from capsid protein VP 1 was found to be best as it bound in the groove of the HLA-A*11:01 with an energy of −8.8 kcal/mol. Again, 'VPIPFASKQ' and 'PNYSPSSIS' were found best for protein VP 2 and their binding energy was −8.4 kcal/mol.

For B-cell epitope prediction, Emini surface accessibility prediction, Chou & Fasman beta-turn prediction and Schulz & Karplus flexibility prediction method was employed to identify the easily accessible, most potent beta-turns and most flexible regions respectively. Hydrophilicity and antigenicity were also predicted as well for both proteins. From all above, the most potent B cell epitopes for capsid protein VP1 and protein VP2 were identified as vaccine candidates against human norovirus (Supplementary Table 6). The final vaccine proteins were constructed using the promiscuous epitopes and protein adjuvants along with PADRE sequence. Beta defensin adjuvant can act as an agonist to TLR1, TLR2 and TLR4 receptor while L7/L12 ribosomal protein and HBHA protein are agonists to only TLR4 (Rana and Akhter, 2016). PADRE sequence was incorporated along with the adjuvant peptides to overcome the problems caused by highly polymorphic HLA alleles.

Literature studies revealed that PADRE containing vaccine construct showed better CTL responses than the vaccines lacked it (Wu et al., 2010).Individual epitopes were linked together via suitable linker to ensure effective immune response. EAAAK linkers were used to join the adjuvant and CTL epitopes. Similarly, GGGS, GPGPG and KK linkers were used to conjugate the CTL, HTL and BCL epitopes respectively. Linkers were utilized to ensure effective separation of individual epitopes in vivo (Saadi et al., 2017; Karkhah et al., 2017; Hajighahramani et al., 2017; Pandey et al., 2016).The constructed vaccines were further checked for their non-allergic behavior and immunogenic potential (Supplementary Table 4).

However, construct V1 was superior in terms of antigenicity and VaxiJen score. Results also ensured its higher solubility during heterologous expression in the *E. coli*. Moreover, the final vaccine constructs were functionally characterize according to molecular weight, aliphatic index, isoelectric pH, hydropathicity, instability index, GRAVY values, estimated half-life and various physicochemical properties. Physicochemical properties of vaccines such as surface charge, shape, hydrophobicity, pH and composition play a pivotal role to determine their activity and potency. Researchers also correlated the physicochemical properties of vaccines with immune responses (Bastola and Lee, 2019). Different physicochemical analysis of our designed vaccine classified it as a stable one.

To strengthen our prediction, we checked the interactions between designed vaccine constructs and different HLA molecules (i.e. DRB1*0101, DRB3*0202, DRB5*0101, DRB3*0101, DRB1*0401, and DRB1*0301). Again construct V1 was found to be best considering the free binding energy. Moreover, docking analysis was also performed to explore the binding affinity of vaccine protein V1 and human TLR8 receptor to evaluate the efficacy of used adjuvant. Molecular dynamics study was conducted to determine the complex stability. Structural dynamics had been investigated previously using subsets of atoms and covariance analysis (Aalten et al., 1997). Literature studies linked the stability of macromolecules with correlated fluctuations of atoms (Clarage et al., 1995; Caspar, 1995). In the present study, essential dynamics was compared to the normal modes of proteins to determine its stability through iMODS server. The analysis revealed negligible chance of deformability for each individual residue, as location of hinges in the chain was not significant and thereby strengthening our prediction. Finally, the designed vaccine construct V1 was reverse transcribed and adapted for *E. coli* strain K12 prior to insertion within pET28a(+) vector for its heterologous cloning and expression. Moreover, Similarity analysis is very important since vaccine could not trigger immune protection, instead the body develop immune tolerance against the vaccine because their sequence are similar with human body. But, the study found that the designed vaccine constructs (V1, V2 and V3) did not show similarity with human proteins which could be considered as the further efficacy of the designed vaccine molecules.

In case of conventional vaccines, the virulent form of attenuated vaccine may reverse or inactivation process may fail to suppress the virulence property (Hasson et al., 2015; Kaufmann et al., 2014). Moreover, some recent clinical trials raised safety concerns while emphasizing on conventional approach for vaccine development (Tameris et al., 2013).Sometimes the cellular lines used for the preparation of the vaccine may be contaminated by microorganisms that do not produce any cytopathic effect while undergoing replication along with the vaccine virus (Merten, 2002; Stratton et al., 2003). If the quality control process fails to detect these contaminations it may be proved hazardous for the host. However, the recombinant subunit vaccine does not carry any risk of emergence of virulence property as it comprises the antigenic parts only (not the entire virus itself). Besides, a number of tools have been used in this study to validate the nontoxic and non-allergic behavior of the designed vaccine construct. However, the predicted in silico results were based on different analysis of sequence and various immune databases. We suggest further wet lab based analysis using model animals for the experimental validation of our predicted vaccine candidates.

5. Conclusion

The genomic-based technologies continue to transform the field of vaccinology through aiding selection of potential vaccine candidates and facilitating the optimization of the chosen immunogens. In this study, we have designed overall vaccine models with highly immunogenic epitopes and also suggested expression system which definitely can confer a good message to the scientific community to carry out further in vivo trials. Thus the present study prompts future vaccine development against human norovirus and other emerging infectious diseases.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.meegid.2019.103936.

Funding

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

Conflict of interests

The authors declare that they have no conflicts of interest.

Authorship contributions

Kazi Faizul Azim: Experiment design, data handling, data analysis, manuscript writing and draft preparation.

Mahmudul Hasan: Supervision of experimental design, project administration and reviewing.

Md. Nazmul Hossain: Idea generation, conceptualization, manuscript writing, reviewing and approval of final manuscript.

Saneya Risa Somana, Syeda Farjana Hoque: Data handling and data analysis.

Md. Nazmul Islam bappy, Anjum Taiebah Chowdhury, Tahera Lasker: Data handling.

Acknowledgments

Authors would like to acknowledge The Department of Microbial Biotechnology of Sylhet Agricultural University for their technical support for the implementation of the project.

References

Aalten, D.M.F., Groot, B.L., Findlay, J.B.C., Berendsen, H.J.C., Amadei, A., 1997. A comparison of techniques for calculating protein essential dynamics. J. Comput. Chem. 18 (2), 169–181.

Adhikari, U.K., Tayebi, M., Rahman, M., 2018. Immunoinformatics approach for epitope-based peptide vaccine design and active site prediction against polyprotein of emerging oropouche virus. J. Immunol. Res. 2018, 1–22.

Agrawal, P., Raghava, G.P., 2018. Prediction of antimicrobial potential of a chemically modified peptide from its tertiary structure. Front. Microbiol. 9, 2551.

Ahluwalia, P.K., Pandey, R.K., Sehajpal, P.K., Prajapati, V.K., 2017. Perturbed microRNA expression by mycobacterium tuberculosis promotes macrophage polarization leading to pro-survival foam cell. Front. Immunol. 8, 107.

Aliabadi, N., Lopman, B.A., Parashar, U.D., Hall, A.J., 2015. Progress toward norovirus vaccines: considerations for further development and implementation in potential target populations. Expert Rev. Vaccines 14 (9), 1241–1253.

Awan, F.M., Obaid, A., Ikram, A., Janjua, H.A., 2017. Mutation-structure function relationship based integrated strategy reveals the potential impact of deleterious missense mutations in autophagy related proteins on hepatocellular carcinoma (HCC): a comprehensive informatics approach. Int. J. Mol. Sci. 18 (1), 139.

Bastola, R., Lee, S., 2019 May 15. Physicochemical properties of particulate vaccine adjuvants: their pivotal role in modulating immune responses. J. Pharm. Invest. 49 (3), 279–285.

Beek, J., Graaf, M., Al-Hello, H., et al., 2018. Molecular surveillance of norovirus, 2005–2016: an epidemiological analysis of data collected from the NoroNet network. Lancet Infect. Dis. 18, 545–553.

Bernstein, D.I., Atmar, R.L., Lyon, G.M., Treanor, J.J., Chen, W.H., Jiang, X., et al., 2015.

Norovirus vaccine against experimental human GII.4 virus illness: a challenge study in healthy adults. J. Infect. Dis. 211 (6), 870–878.
Bok, K., Green, K.Y., 2012. Norovirus gastroenteritis in immunocompromised patients. N. Engl. J. Med. 367 (22), 2126–2132.
Bui, H.H., Sidney, J., Dinh, K., Southwood, S., Newman, M.J., Sette, A., 2006. Predicting population coverage of T-cell epitope-based diagnostics and vaccines. BMC bioinformatics 7, 153.
Caspar, D.L., 1995. Problems in simulating macromolecular movements. Structure 3, 327–329.
Chan, P., Curtis, R.A., Warwicker, J., 2013. Soluble expression of proteins correlates with a lack of positively-charged surface. Sci. Rep. 3, 3333.
Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T.J., Higgins, D.G., Thompson, J.D., 2003. Multiple sequence alignment with the Clustal series of program. Nucleic Acids Res. 31 (13), 3497–3500.
Chou, P., Fasman, G., 1978. Prediction of the secondary structure of proteins from their amino acid sequence. Adv. Enzymol. 47, 45–148.
Chrysostomou, C., Seker, H., 2014. Prediction of protein allergenicity based on signal-processing bioinformatics approach. In: 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society.
Clarage, J.B., Romo, T., Andrews, B.K., Pettitt, B.M., Phillips, G.N., 1995. A sampling problem in molecular dynamics simulations of macromolecules. Proc. Natl. Acad. Sci. USA 92, 3288–3292.
Comeau, S.R., Gatchell, D.W., Vajda, S., Camacho, C.J., 2004. ClusPro: a fully automated algorithm for protein-protein docking. Nucl. Acids Res. 32, W96–W99.
Cooper, N.R., Nemerow, G.R., 1984. The role of antibody and complement in the control of viral infections. J. Invest. Dermatol. 83, 121–127.
Cortes-Penfield, N.W., Ramani, S., Estes, M.K., Atmar, R.L., 2017. Prospects and challenges in the development of a norovirus vaccine. Clin. Ther. 39 (8), 1537–1549.
Craig, D.B., Dombkowski, A.A., 2013. Disulfide by Design 2.0: a web-based tool for disulfide engineering in proteins. BMC Bioinform. 14, 346.
Cros, J., Cagnard, N., Woollard, K., Patey, N., Zhang, S.Y., Senechal, B., Puel, A., Biswas, S.K., Moshous, D., Picard, C., Jais, J.P., 2010. Human CD14dim monocytes patrol and sense nucleic acids and viruses via TLR7 and TLR8 receptors. Immunity 33 (3), 375–386.
Cui, Q., Bahar, I., 2007. Normal mode analysis theoretical and applications to biological and chemical systems. Briefing Bioinform. 8 (5), 378–379.
Das, K., Chakraborty, S., Hasan, M., Shovo, A.M., 2016. In Silico Analysis to Elect Superior Bacterial Alkaline Protease for Detergent and Leather Industries. 5. pp. 685–698.
Dash, R., Junaid, M., Akash, M.F.C., Islam, A., Hosen, S.M.J., 2017. In silico-based vaccine design against Ebola virus glycoprotein. Adv. Appl. Bioinform. Chem. 10, 11–28.
Dimitrov, I., Bangov, I., Flower, D.R., Doytchinova, I.A., 2013. AllerTOP v.2- a server for in silico prediction of allergens. J Mol. Model. 20, 2278.
Dimitrov, I., Naneva, L., Doytchinova, I.A., Bangov, I., 2014. AllergenFP: allergenicity prediction by descriptor fingerprints. Bioinformatics 30, 846–851.
Doytchinova, I.A., Flower, Darren R., 2007a. Identifying candidate subunit vaccines using an alignment-independent method based on principal amino acid properties. Vaccine 25, 856–866.
Doytchinova, I.A., Flower, D.R., 2007b. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinform. 8, 4.
Emini, E.A., Hughes, J.V., Perlow, D., Boger, J., 1985. Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. J. Virol. 55, 836–839.
Extra-Kangueane, P., Nilofer, C., 2018. Protein-protein docking: methods and tools. In: InProtein-Protein and Domain-Domain Interactions, pp. 161–168.
Fankhauser, R.L., Monroe, S.S., Noel, J.S., Humphrey, C.D., Bresee, J.S., Parashar, U.D., et al., 2002. Epidemiologic and molecular trends of [a]Norwalk-like viruses associated with outbreaks of gastroenteritis in the United States. J. Infect. Dis. 186, 1–7.
Fiers, M.W.E.J., Kleter, G.A., Nijland, H., Peijnenburg, A., Nap, J.P., Ham, R., 2004. Allermatch™, a webtool for the prediction of potential allergenicity according to current FAO/WHO Codex alimentarius guidelines. BMC Bioinform. 5, 133.
Garcia, K.C., Teyton, L., Wilson, I.A., 1999. Structural basis of T cell recognition. Annu. Rev. Immunol. 17, 369–397.
Gasteiger, E., Hoogland, C., GattikerA, Duvaud S., Wilkins, M.R., Appel, R.D., Bairoch, A., 2003. Protein identification and analysis tools on the ExPASy server. Nucleic Acids Res. 31, 3784–3788.
Grote, A., et al., 2005. JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. Nucleic Acids Res. 33, W526–W531.
Gupta, S., Kapoor, P., Chaudhary, K., Gautam, A., Kumar, R., 2013. Open source drug discovery consortium, Raghava GP in silico approach for predicting toxicity of peptides and proteins. PLoS One 8 (9), e73957.
Hajighahramani, N., Nezafat, N., Eslami, M., Negahdaripour, M., Rahmatabadi, S.S., Ghasemi, Y., 2017. Immunoinformatics analysis and in silico designing of a novelmulti-epitope peptide vaccine against Staphylococcus aureus. Infect. Genet. Evol. 48, 83–94.
Hasan, M., Joy, Z.F., Bhuiyan, E.H., Islam, M.S., 2015a. In Silico characterization and motif election of neurotoxins from Snake venom. Am. J. Biochem. Biotechnol. 11 (2), 84.
Hasan, M., Hakim, A., Iqbal, A., Bhuiyan, F.R., Begum, M.K., Sharmin, S., Abir, R.A., 2015b. Computational study and homology modeling of phenol hydroxylase: key enzyme for phenol degradation. Int. J. Comput. Bioinfo. In Silico Model 4 (4), 691–698.
Hasan, M., Azim, K.F., Begum, A., Khan, N.A., Shammi, T.S., Imran, A.S., Chowdhury, I.M., Urme, S.R., 2019a. Vaccinomics strategy for developing a unique multi-epitope monovalent vaccine against Marburg marburgvirus. Infect. Genet. Evol. 70, 140–157.
Hasan, M., Ghosh, P., Azim, K., Mukta, S., Abir, R.A., Nahar, J., Khan, M.M., 2019b. Reverse vaccinology approach to design a novel multi-epitope subunit vaccine against avian influenza A (H7N9) virus. Microb. Pathog. 130, 19–37.
Hasson, S.S., Al-Busaidi, J.K., Sallam, T.A., 2015. The past, current and future trends in DNA vaccine immunisations. Asian Pac. J. Trop. Biomed. 5 (5), 344–353.
Hebditch, M., Carballo-Amador, M.A., Charonis, S., Curtis, R., Warwicker, J., 2017. Protein–Sol: a web tool for predicting protein solubility from sequence. Bioinformatics 33, 3098–3100.
Hedlund, K.O., Rubilar-Abreu, E., Svensson, L., 2000. Epidemiology of calicivirus infections in Sweden, 1994-1998. J. Infect. Dis. 181 (Suppl. 2), S275–S280.
Heil, F., Hemmi, H., Hochrein, H., Ampenberger, F., Kirschning, C., Akira, S., Lipford, G., Wagner, H., Bauer, S., 2004. Species-specific recognition of single-stranded RNA via toll-like receptor 7 and 8. Science 303 (5663), 1526–1529.
Hossain, N., Rahman, M., 2014. Antagonistic activity of antibiotic producing *Streptomyces sp.* against fish and human pathogenic bacteria. Braz. Arch. Biol. Technol. 57 (2), 233–237.
Jespersen, M.C., Peters, B., Nielsen, M., Marcatili, P., 2017. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. Nucleic Acids Res. 45, W24–W29.
Jones, M.K., Grau, K.R., Costantini, V., Kolawole, A.O., De Graaf, M., Freiden, P., Graves, C.L., Koopmans, M., Wallet, S.M., Tibbetts, S.A., Schultz-Cherry, S., 2015. Human norovirus culture in B cells. Nat. Protoc. 10 (12), 1939.
Joy, Z.F., Purkaystha, A., Das, N.K., Al-Hakim, Chakrabarty S., Hasan, M., 2019. Screening for alternative sources of L-Asparaginase used in acute lymphoblastic leukaemia (All) treatment: an in Silico approach. Bioinform. Proteom. Open Access J. 3 (1), 1–9.
Kallberg, M., Margaryan, G., Wang, S., Ma, J., Xu, J., 2014. RaptorX server: a resource for template-based protein structure modeling. InProtein Struct. Predict. 1137, 17–27.
Karkhah, A., Saadi, M., Nouri, H.R., 2017 Apr 1. In silico analyses of heat shock protein 60 and calreticulin to designing a novel vaccine shifting immune response toward T helper 2 in atherosclerosis. Comput. Biol. Chem. 67, 244–254.
Karplus, P., Schulz, G., 1985. Prediction of chain flexibility in proteins. Naturwissenschaften 72, 212–213.
Kaufmann, S.H., McElrath, M.J., Lewis, D.J., Del Giudice, G., 2014. Challenges and responses in human vaccine development. Curr. Opin. Immunol. 28, 18–26.
Kaur, H., Garg, A., Raghava, G., 2007. PEPstr: a de novo method for tertiary structure prediction of small bioactive peptides. Protein Pept. Lett. 14, 626–631.
Kolaskar, A., Tongaonkar, P.C., 1990. A semi-empirical method for prediction of antigenic determinants on protein antigens. FEBS Lett. 276, 172–174.
Kosciolek, T., Jones, D.T., 2014. De novo structure prediction of globular proteins aided by sequence variation-derived contacts. PLoS One 9, e92197.
Kovacs, J., Chacon, P., Abagyan, R., 2004. Predictions of protein flexibility: first order measures. PROTEINS: structure. Funct. Bioinform. Proteins 56 (4), 661–668.
Krogh, A., Larsson, B., Von Heijne, G., Sonnhammer, E.L., 2001. Predicting transmembrane protein topology with a hidden markov model: application to complete genomes1. J. Mol. Biol. 305, 567–580.
Kroneman, A., Vega, E., Vennema, H., et al., 2013. Proposal for a unified norovirus nomenclature and genotyping. Arch. Virol. 158 (10), 2059–2068.
Lanata, C.F., Fischer-Walker, C.L., Olascoaga, A.C., Torres, C.X., Aryee, M.J., et al., 2013. Global causes of diarrheal disease mortality in children 5 years of age:a systematic review. PLoS One 8 (9), e72788.
Lopez-Blanco, J.R., Aliaga, J.I., Quintana-Orti, E.S., Chacon, P., 2014. iMODS: internal coordinates normal mode analysis server. Nucleic Acids Res. 42, W271–W276.
Lovell, S.C., Davis, I.W., Arendall, W.B., Bakker, P.I.W., Word, J.M., Prisant, M.G., Richardson, J.S., Richardson, D.C., 2002. Structure validation by Calpha geometry: phi,psi and Cbeta deviation. Proteins 50, 437–450.
Macalino, S., Basith, S., Clavio, N., Chang, H., Kang, S., Choi, S., 2018. Evolution of InSilico strategies for protein-protein interaction drug discovery. Molecules 23 (8), 1963.
Malm, M., Tamminen, K., Lappalainen, S., Uusi-Kerttula, H., Vesikari, T., Blazevic, V., 2015. Genotype considerations for virus-like particle-based bivalent norovirus vaccine composition. Clin. Vaccine Immunol. 22 (6), 656–663.
Mattison, C.P., Cardemil, C.V., Hall, A.J., 2018. Progress on Norovirus vaccine research: public health considerations and future directions. Exp. Rev. Vaccines 7 (9), 773–784.
Maupetit, J., Derreumaux, P., Tuffery, P., 2010. A fast method for large scale De Novo peptide and miniprotein structure prediction. J. Comput. Chem. 31, 726–738.
McKeever, T.M., Lewis, S., Smith, C., Hubbard, R., 2004. Vaccination and allergic disease: a birth cohort study. Am. J. Public Health 94 (6), 985–989.
Merten, O.W., 2002. Virus contaminations of cell cultures–a biotechnological view. Cytotechnology 39 (2), 91–116.
Mohammed, A.A., Hashim, O., Elrahman, K.A.A., Hamdi, A., Hassan, M.A., 2017. Epitope-based peptide vaccine design against Mokola rabies virus glycoprotein G utilizing in Silico approaches. Immunome Res. 13, 144.
Mora, M., Veggi, D., Santini, L., Pizza, M., Rappuoli, R., 2003. Reverse vaccinology. Drug Discov. Today 8 (10), 459–464.
Morris, G.M., Huey, R., Lindstrom, W., Sanner, M.F., Belew, R.K., Goodsell, D.S., et al., 2009. AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. J. Comput. Chem. 30, 2785–2791.
Pandey, R.K., Sundar, S., Prajapati, V.K., 2016. Differential expression of miRNA regulatesT cell differentiation and plasticity during visceral leishmaniasis infection. Front. Microbiol. 7, 206.
Parker, J., Guo, D., Hodges, R., 1986. New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites. Biochemistry 25, 5425–5432.
Peng, J., Xu, J., 2011. RaptorX: exploiting structure information for protein alignment by statistical inference. Proteins 79 (S10), 161–171.

Petersen, B., Lundegaard, C., Petersen, T.N., 2010. NetTurnP–neural network prediction of beta-turns by use of evolutionary information and predicted protein sequence features. PLoS One 5 (11), e15079.
Phillips, G., Tam, C.C., Rodrigues, L.C., Lopman, B., 2010a. Prevalence and characteristics of asymptomatic norovirus infection in the community in England. Epidemiol. Infect. 138 (10), 1454–1458.
Phillips, G., et al., 2010b. Community incidence of norovirus-associated infectious intestinal disease in England: improved estimates using viral load for norovirus diagnosis. Am. J. Epidemiol. 171 (9), 1014–1022.
Poland, G.A., Ovsyannikova, I.G., Jacobson, R.M., 2009. Application of pharmacogenomics to vaccines. Future Med. 10 (5), 837–852.
Prabhakar, P.K., Srivastava, A., Rao, K.K., Balaji, P.V., 2016. Monomerization alters the dynamics of the lid region in campylobacter jejuniCstII: an MD simulation study. J. Biomol. Struct. Dyn. 34 (4), 778–779.
Purcell, A.W., McCluskey, J., Rossjohn, J., 2007. More than one reason to rethink the use of peptides in vaccine design. Nat. Rev. Drug Discov. 6, 404–414.
Rana, A., Akhter, Y., 2016. A multi-subunit based, thermodynamically stable model vaccine using combined immunoinformaticsandprotein structure based approach. Immunobiology 221, 544–557.
Rappuoli, R., 2000. Reverse vaccinology. Curr. Opin. Microbiol. 3, 445–450.
Rappuoli, R., Bottomley, M.J., D'Oro, U., Finco, O., De Gregorio, E., 2016. Reverse vaccinology 2.0: human immunology instructs vaccine antigen design. J. Exp. Med. 213 (4), 469–481.
Saadi, M., Karkhah, A., Nouri, H.R., 2017 Jul 1. Development of a multi-epitope peptide vaccine inducing robust T cell responses against brucellosis using immunoinformatics based approaches. Infect. Genet. Evol. 51, 227–234.
Saha, S., Raghava, G.P., 2006. AlgPred: prediction of allergenic proteins and mapping of IgE epitopes. Nucl. Acids Res. 34 (suppl_2), W202-9.
Sanner, M.F., 1999. Python: a programming language for software integration and development. J. Mol. Graph Model 17 (1), 57–61.
Sanner, M.F., Stoffler, D., Olson, A.J., 2002. ViPEr, a visual programming environment for Python. Proceedings of the 10th International Python Conference 1–15.
Schneidman-Duhovny, D., Inbar, Y., Nussinov, R., Wolfson, H.J., 2005. PatchDock and SymmDock: servers for rigid and symmetric docking. Nucl. Acids Res. 33, W363–W367.
Sette, A., Fikes, J., 2003. Epitope-based vaccines: an update on epitope identification, vaccine design and delivery. Curr. Opin. Immunol. 15 (4), 461–470.
Sharmeen, R., Hossain, M.N., Rahman, M.M., Foysal, M.J., Miah, M.F., 2012. In-vitro antibacterial activity of herbal aqueous extract against multi-drug resistant *Klebsiella sp.* isolated from human clinical samples. Int. Curr. Pharm. J. 1 (6), 133–137.
Shrestha, B., 2004. Role of CD8+ T cells in control of West Nile virus infection. J. Virol. 12, 8312–8321.
Siebenga, J.J., Vennema, H., Renckens, B., de Bruin, E., van der Veer, B., Siezen, R.J., Koopmans, M., 2007. Epochal evolution of GGII.4 norovirus capsid proteins from 1995 to 2006. J. Virol. 81 (18), 9932–9941.
Siebenga, J.J., Vennema, H., Zheng, D.P., et al., 2009. Norovirus illness is a global problem: emergence and spread of norovirus GII.4 variants, 2001-2007. J. Infect. Dis. 200, 802–812.
Sievers, F., Higgins, D.G., 2014. Russell DJ (ed.) multiple sequence alignment methods. Methods Mol. Biol. 1079, 105–116 Humana Press.
Smialowski, P., Martin-Galiano, A.J., Mikolajka, A., Girschick, T., Holak, T.A., Frishman, D., 2006. Protein solubility: sequence based prediction and experimental verification. Bioinformatics 23, 2536–2542.
Solanki, V., Tiwari, V., 2018. Subtractive proteomics to identify novel drug targets and reverse vaccinology for the development of chimeric vaccine against Acinetobacterbaumannii. Sci. Rep. 8 (1), 9044.
Stratton, K., Almario, D.A., Wizemann, T.M., McCormick, M.C., 2003. Immunization Safety Review: Vaccinations and Sudden Unexpected Death in Infancy. Institute of Medicine (US) Immunization Safety Review Committee, National Academies Press (US).
Tameris, M.D., Hatherill, M., Landry, B.S., Scriba, T.J., Snowden, M.A., Lockhart, S., Shea, J.E., McClain, J.B., Hussey, G.D., Hanekom, W.A., et al., 2013. Safety and efficacy of MVA85A, a new tuberculosis vaccine, in infants previously vaccinated with BCG: a randomised, placebo-controlled phase 2b trial. Lancet 381, 1021–1028.
Thaysen-Andersen, M., Packer, N.H., 2012. Site-specific glycoproteomics confirms that protein structure dictates formation of N-glycan type, core fucosylation and branching. Glycobiology 22 (11), 1440–1452.
Thomsen, M., Lundegaard, C., Buus, S., Lund, O., Nielsen, M., 2013. MHCcluster, a method for functional clustering of MHC molecules. Immunogenetics 65, 655–665.
Vinje, J., 2015. Advances in laboratory methods for detection and typing of norovirus. J. Clin. Microbiol. 53 (2), 373–381.
Vita, R., Overton, J.A., Greenbaum, J.A., Ponomarenko, J., Clark, J.D., Cantrell, J.R., et al., 2014. The immune epitope database (IEDB) 3.0. Nucleic Acids Res. 43, D405–D412.
Wang, Z., Eickholt, J., Cheng, J., 2011. APOLLO: a quality assessment service for single and multiple protein models. Bioinformatics 27, 1715–1716.
Wu, C.Y., Monie, A., Pang, X., Hung, C.F., Wu, T.C., 2010. Improving therapeutic HPV peptide-based vaccine potency by enhancing CD4+T help and dendritic cell activation. J. Biomed. Sci. 17, 88.
Wuthrich, K., Wagner, G., Richarz, R., Braun, W., 1980. Correlations between internal mobility and stability of globular proteins. Biophys. J. 32 (1), 549–560.
Xu, D., Zhang, Y., 2011. Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization. Biophys. J. 101, 2525–2534.
Zhang, J., Liang, Y., Zhang, Y., 2011. Atomic-level protein structure refinement using fragment-guided molecular dynamics conformation sampling. Structure 19, 1784–1795.