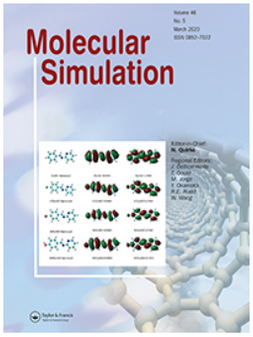

Molecular Simulation

ISSN: 0892-7022 (Print) 1029-0435 (Online) Journal homepage: https://www.tandfonline.com/loi/gmos20

Utilising capsid proteins of poliovirus to design a multi-epitope based subunit vaccine by immunoinformatics approach

Nitin Bohra, Santanu Sasidharan, Shweta Raj, S. N. Balaji & Prakash Saudagar

To cite this article: Nitin Bohra, Santanu Sasidharan, Shweta Raj, S. N. Balaji & Prakash Saudagar (2020) Utilising capsid proteins of poliovirus to design a multi-epitope based subunit vaccine by immunoinformatics approach, Molecular Simulation, 46:5, 419-428, DOI: 10.1080/08927022.2020.1720916

To link to this article: https://doi.org/10.1080/08927022.2020.1720916

MOLECULAR SIMULATION
2020, VOL. 46, NO. 5, 419–428
https://doi.org/10.1080/08927022.2020.1720916

Taylor & Francis
Taylor & Francis Group

Utilising capsid proteins of poliovirus to design a multi-epitope based subunit vaccine by immunoinformatics approach

Nitin Bohra*, Santanu Sasidharan*, Shweta Raj, S. N. Balaji and Prakash Saudagar

Department of Biotechnology, National Institute of Technology, Warangal, Telangana, India

ABSTRACT

Current vaccines used for polio are live attenuated oral polio vaccine and inactivated polio vaccine. Recently, in India, the relapse of virulence was observed in attenuated viruses resulting in catastrophic effects. Therefore, the need for the development of multi-epitope subunit vaccine was realised and an immunoinformatics approach to design a multi-epitope subunit vaccine was conceived. Capsid proteins of all the three types of polio strains were utilised to predict major histocompatibility complex class-1 as well as class-2 epitopes. The subunit vaccine was designed with β-defensin at N-terminal followed by cytotoxic T-lymphocytes epitopes and helper T-lymphocytes epitopes connected by compatible linkers. The vaccine construct was further modelled and docked against TLR4 receptor. The high affinity of the construct towards the receptor was observed in the docking study and also substantiated by a 20 ns simulation of the complex. The vaccine construct was cloned in-silico for expression of the protein effectively in a prokaryotic system (*Escherichia coli* strain K12). Immuno-simulation of the construct was found to elicit immunoglobulin production effectively in the human body. This designed multi-epitope subunit vaccine is capable of immune response and further studies will help us understand the feasibility of this multi-epitope subunit vaccine.

ARTICLE HISTORY
Received 21 September 2019
Accepted 12 January 2020

KEYWORDS
Immunoinformatics; multi-epitope; poliovirus; vaccine; molecular dynamic simulation

Highlights

- A multi-epitope sub-unit vaccine for the treatment of polio has been proposed in this study.
- CTL and HTL epitopes from three strains stitched together were capable of eliciting immune response *in-silico*.
- Modelling, docking and molecular dynamics simulation exhibited a stable relationship with TLR receptor.
- Immuno-simulation of the construct elicited IgG and IgM antibodies along with T_H and T_c cells.

1. Introduction

Poliomyelitis commonly known as polio is caused by the poliovirus which belongs to the family of *Picornaviridae* and genus *Enterovirus*. A portray from 1580 to 1350 BC of an Egyptian priest with withered leg suggested that polio existed in early civilisations. Polio epidemics have crippled millions of people and has a history of thousands of years. In 1894, the first documented case of polio was registered in the United States [1]. Polio is a contagious viral infectious disease that transfers primarily through fecal-oral route. It is caused by an intestinal virus that attacks the nerve cells of the brain and spinal cord. Different types of paralysis have been recorded depending upon the type of nerve cells involved in it. The most common form of polio is spinal polio characterised by the asymmetrical paralysis involving legs usually. Another form of paralysis called bulbar polio causes weakness of muscles that are innervated by cranial nerved. A combination of both named as bulbospinal paralysis also occurs widely [2].

The virus protects its genetic material by employing the capsid that forms a protective covering. This capsid protein also enables the poliovirus to infect certain types of nerve cells. There are three serotypes depending upon the capsid proteins: type 1 (PV1), type 2 (PV2) and type 3 (PV3) [3]. All three serotypes are extremely virulent and have been known to exhibit similar symptoms. Among the three serotypes, PV1 has been the most commonly found serotype and also the most closely associated with the disease paralysis. Infection of individuals with the contemporary vaccines develop immunity. The most commonly observed immunity being in tonsils and gastrointestinal tract. The immunity conferred by one serotype is different from another and therefore full-fledged immunity against polio requires exposure of the individual to each of the serotype [4]. This calls for the need to design a multi-epitope vaccine against polio that can develop immunity against all three serotypes in an individual.

Polio vaccines have been used throughout the world to combat the virus extensively. The first developed vaccine was in 1952 by Jonas Salk that consisted of inactivated/dead poliovirus. An oral vaccine was later developed by Albert Sabin using attenuated virus in 1957. These vaccines include administration of heat-killed or live attenuated virus/bacteria which reduce the safety regarding the utilisation of such vaccines. But, there have been several relapses of the virus despite having no long term carrier and no primate reservoir in nature as of today. There is an immediate requirement of other forms of vaccine that do not involve any viral infectious particles. Subunit vaccines are an emerging vaccine type where the safety profile can be enhanced as they do not consist of the whole micro-organism but contain

CONTACT Prakash Saudagar ps@nitw.ac.in
*The authors contributed equally.

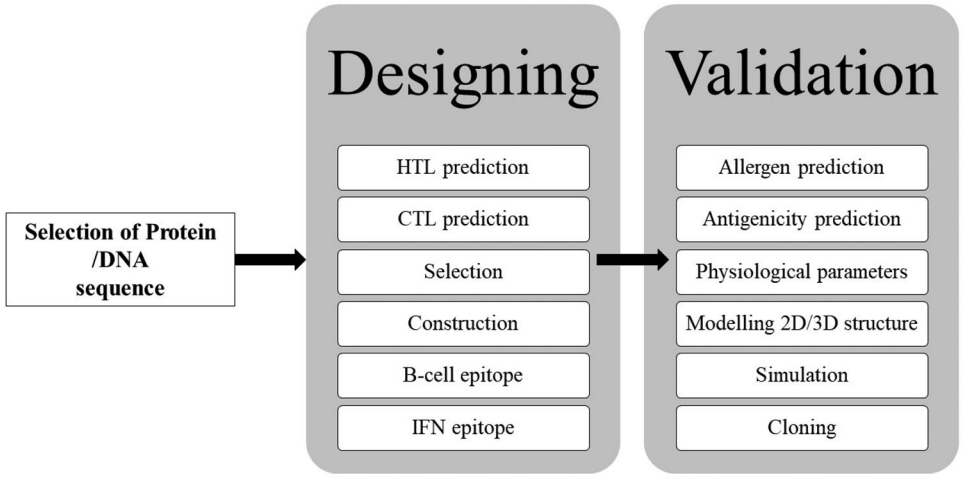

Figure 1. Workflow representation of the study is depicted in the figure. The study was carried out in two phases (i) Designing and (ii) Evaluation. Different works carried out under each section is shown in the figure.

only specific epitope regions of antigens present in them [5,6]. The immunogenicity of these antigens can be enhanced by the addition of adjuvants. Subunit vaccines allow the customisation of immune responses by means of optimised delivery systems, adjustment of vaccine size, and generation of adaptive immune responses such as humoral, cell-mediated immune responses along with enhanced efficacy of vaccine [7].

In this study, the authors utilise immunoinformatics tools to develop a multi-epitope sub-unit vaccine which is capable of generating humoral, cell-mediated and innate immune responses as it contains B-cell and T-cell inducing epitopes. We retrieved the protein sequence of all structural proteins of poliovirus strains in FASTA format from the UniProt database. These epitopes were found and joined together by appropriate linkers to enhance the epitope presentation and efficient separation of epitopes by proteasomal degradation inside the human body. Also, an adjuvant (β-defensin) (Mohan et al. [21]) was added at the N-terminal of the above sequence of epitopes and linkers were also combined to enhance the immunogenicity of the vaccine. Vaccine properties such as allergenicity, antigenicity, and physicochemical properties were also predicted. This structure was subjected to docking with TLR-4 to explore the binding interactions and molecular dynamics simulation was conducted for 20 ns to understand the stability of the interactions. TLR-4 was taken as the immune receptor of choice because, its expression increases with immune cells when poliovirus enters inside the host [8]. Finally, the translation efficiency of the vaccine model in an expression vector pET28a(+) was confirmed by *in-silico* cloning. The workflow of the study is depicted in Figure 1.

2. Material and methods

2.1. Collection of poliovirus proteins for multi-epitope vaccine construction.

A detailed survey of the literature was conducted to identify crucial strains and proteins of *poliovirus*. The amino acid sequence of six strains of *poliovirus* Type 1 (*strain Mahoney, strain Sabin*), *poliovirus* Type 2 (*strain Lansing, strain W-2*) and *poliovirus* Type 3 (*strain 23127, strain P3*) were retrieved from UNIPROT database in FASTA format [9]. Capsid proteins make the outside envelope of poliovirus in which virus genome is embedded. Capsid proteins are selected because they play an important role in host cell surface receptor binding, cell entry and infectivity. From all the strains five capsid proteins VP0, VP1, VP2, VP3 and VP4 were selected from UNIPROT database for multi-epitope vaccine construction (Table 1). All these curated protein sequences were then used for various epitope predictions.

2.2. MHC class II specific helper T-lymphocyte epitopes prediction

Major histocompatibility complex plays a crucial role in the presentation of antigen to activate T-lymphocytes. Class II MHC activates the $CD4^+$ T-lymphocytes (or HTL) which play a crucial role in humoral and adaptive immunity [10]. Therefore, it was essential to identify HTL epitopes for the activation of the immune response. HTL epitopes of above-selected capsid proteins of all the poliovirus strains were predicted using immune epitope database (IEDB) (https://tools.iedb.org/mhcii/) server. The server is based on IEDB recommended prediction method which uses the Consensus approach [11], combining NN-align [12], SMM-align [13], CombLib [14] and Sturniolo approach [15]. FASTA format of protein sequences were submitted and IEDB recommended prediction method for the full reference set of 27 alleles of human species covering >99% of the world population were used. The selection of HTL epitopes was based on IC_{50} values and percentile rank, the peptide having low IC_{50} values will have the highest affinity whereas the percentile rank of each predicted peptides was determined by comparing it to 5 million randomly generated peptides from SWISSPRO and least rank represented the highest affinity [16].

Table 1. The table shows the strains that were selected for the construction of the multi-epitope sub-unit vaccine construct.

TYPE-1	TYPE-2	TYPE-3
Strain Mahoney(P03300)	Strain W-2(P23069)	Strain P3 (P03302)
Strain Sabin (P03301)	Strain Lansing(P06210)	Strain 23127 (P06209)

2.3. MHC class 1 specific cytotoxic T-lymphocyte epitopes prediction

$CD8^+$ T-lymphocytes (or CTL) are activated by MHC class I and play a vital role in the cell-mediated immune response. They are responsible to eliminate virus-infected cells, cancerous cells and other damaged cells by identifying antigen epitopes presented by MHC class I complex on the cell surface. CTL epitopes had to be selected carefully for the activation of the cellular immune response. NetCTL 1.2 server (https://www.cbs.dtu.dk/services/NetCTL/) was used to predict CTL epitopes of all capsid proteins of the selected poliovirus strains [17]. The server is predicted using neural networks trained using 81 different human MHC alleles and 41 animal alleles which includes monkey, cattle, pig and mouse.

Three components were utilised to predict CTL epitopes namely MHC class-1 peptide binding affinity, proteasomal C terminal cleavage, and TAP transport efficiency [18]. Among the parameters, three different supertypes A2, A3, and B7 were selected to cover approximately 88% of the population, weight on C terminal cleavage, weight on TAP transport efficiency, and the threshold for epitope prediction was at default values of 0.15, 0.05, and 0.75 respectively [19].

2.4. Construction of multi-epitope vaccine and physicochemical properties evaluation

All the predicted HTL and CTL epitopes were checked for overlapping sequences and then stitched together using suitable linkers. To join HTL epitopes with each other GPGPG linker was used whereas AAY linker was utilised to join CTL epitopes [20]. At the N terminal of the constructed peptide, β defensin adjuvant sequence was added using EAAAK linker. β defensin provides innate host immune response for viral and microbial infection, recruit naïve T-cell and immature dendritic cell at the place of infection using CCR6 (Chemokine Receptor 6) receptor [21]. ProtParam server (https://web.expasy.org/protparam/) was used for the analyses of various physical and chemical properties of the constructed multi-epitope subunit vaccine [22]. ProtParam sever only require the primary protein sequence to predict the following parameters which include a number of amino acid, theoretical pI, molecular weight, amino acid composition, instability index, aliphatic index, extinction coefficient, atomic composition, half-life, and GRAVY (grand average of hydropathicity) value [22].

2.5. B-cell epitope prediction

B-cell epitopes are crucial to produce antibodies for encountered antigens and develop memory response for a future encounter with the same antigens. Therefore, B-cell epitopes in a multi-epitope subunit vaccine was determined for understanding its efficacy. Two different types of B-cell epitopes were identified i.e. linear (or continuous) B-cell epitopes and conformational (or discontinuous) B-cell epitopes. Linear B-cell epitopes were checked for the constructed multi-epitope based subunit vaccine using B-cell epitope prediction server (BCPred) (https://ailab.ist.psu.edu/bcpred/predict.html) and similarly, for conformational B-cell epitopes ElliPro (https://tools.iedb.org/ellipro/) sever was employed [23–25]. BCPreds server is based on Support Vector Machine classifiers which have been trained using a dataset of 701 linear B-cell epitopes and 701 non- epitopes. ElliPro implements three algorithms sequentially to predict the antibody epitopes. In both the cases of B-cell prediction, default parameters were used and B-cell epitopes chimera tool showed conformational epitopes.

2.6. Antigenicity and allergenicity prediction

The allergenicity of constructed vaccine was evaluated by the AllerTOP v.2.0 server (https://www.ddg-pharmfac.net/AllerTOP/) [26]. The method used for evaluation of allergenicity is based on auto cross-covariance (ACC) transformation of protein sequences into uniform equal length vectors and the classification of proteins was based on the k-nearest neighbour algorithm (kNN, $k = 1$) which contains a training set of 2427 allergens and non-allergens from different species [27]. Similarly, the antigenicity of the final vaccine construct was predicted using ANTIGENpro tool (https://scratch.proteomics.ics.uci.edu/). This web server tool is based on the sequence of input protein, free from its alignment and also independent of pathogen identity to predict antigenicity and its corresponding probability [28]. VaxiJen v.2.0 server (https://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html) [29] was also used to check the antigenic property of vaccine construct. VaxiJen server predicts antigenicity based on the physicochemical properties of the protein and has a very high accuracy of more than 85% [29].

2.7. Codon optimisation *and* in silico *cloning*

To produce the vaccine protein as an expressed protein in *E. coli,* reverse translation along with codon optimisation was performed using Java Codon Adaptation Tool (JCAT) [30]. The tool uses a method to adapt the usage of codons for prokaryotic organisms. The codon optimisation is required to express foreign genes in prokaryotic hosts. The server provides both CAI-values and GC content. The CAI-values were calculated according to A Carbone et al. [31]. These properties are necessary to achieve high vaccine expression inside the host. The sequence was optimised for expression in *Escherichia coli* K12 strain. pETa(+) vector is a well-known vector used for expression of proteins seamlessly, with a His-tag and lac operon. The DNA sequence was therefore, cloned *in-silico* into pET28(a+) vector between BamHI and HindIII after the addition of restriction sites into the vaccine sequence.

2.8. Modelling, docking and molecular dynamic simulation

The amino acid sequence of the projected vaccine was utilised to predict the 3D structure of the vaccine. I-TASSER server was employed for the automated prediction of 3D structure. The server runs on iterative threading assembly simulations [32]. The quality of the models were analysed

using the C-score and the model with higher confidence has a higher C-score. The refinement of the structure was later performed using GalaxyRefine [33]. The server is known to utilise mild and aggressive relaxation methods to refine the whole protein structure. The GalaxyRefine significantly improves the quality of the initial models with high probability (>50%). The best model was extracted and validated by SAVES sever. SAVES server houses different tools like ERRAT, Verify-3D and Ramachandran plots. The potential errors in the refined model was analysed and necessary refinement was performed accordingly. The interaction of TLR4 receptor (PDB id: 3FXI) and protein vaccine was further investigated using the ClusPro server [34]. The server is a fully automated protein-protein docking server that performs the docking algorithm in three steps and provides the results in the form of scores. MD simulation of the protein vaccine-TLR4 complex was performed to assess the stability of the docked complex. The simulation was done using Gromacs v5.1.4 package and under similar conditions as mentioned in previous work [35]. The simulations were carried out in Amber 99 force field. The receptor-vaccine complex was then solvated by TIP4 water molecules inside an dodecahedron box. After neutralisation and energy minimisation, temperature and pressure equilibration was carried out at 300 K and 1 bar respectively for 1 ns. Parrinello-Rahman and Modified Berendsen thermostat were utilised respectively for NPT and NVT equilibration. The 20 ns simulation study provided the opportunity to establish enough contact between the vaccine-receptor complex and trajectory analysis was conducted for all the time frames.

2.9. Immune simulation

The immunogenicity and the immune response profile of the multi-epitope peptide were characterised using the C-ImmSim server. C-ImmSim server is used for immune predictions using a position-specific scoring matrix (PSSM) and machine learning techniques for prediction of the immune epitope. The server simulates three separate compartments simultaneously that represent the three separate anatomical regions of the mammals i.e. (i) bone marrow which is the site of hematopoietic cells which produces new lymphoid and myeloid cells; (ii) thymus, which is the site of naïve T-cell selection to avoid auto-immunity; (iii) a tertiary lymphatic organ, like lymph node is simulated [36]. Three injections were administered at a time $t = 0$, 17 and 56 days. All other simulation parameters were set at default. The Simpson Index D which is a measure of diversity was also analysed from the plot.

3. Results and discussion

3.1. Selection of poliovirus proteins for vaccine construction

Capsid proteins are the structural proteins of poliovirus which takes part in building its outer envelope and also plays an important role in cell adhesion and cell entry. A poliovirus genome contains nearly 2209 amino acids which code for its structural and non-structural proteins. During detailed literature and database survey six strains of poliovirus were identified. The FASTA format sequences of poliovirus strains were collected from the UNIPROT database and from them, only capsid protein sequences (VP0, VP1, VP2, VP3, and VP4) were obtained for all the strains (Table 1).

3.2. Helper T and cytotoxic T lymphocyte epitopes selection

MHC class II specific HTL epitopes for capsid proteins of each poliovirus type were predicted using the IEDB database [16]. The cutoff value of 0.40 of the percentile rank was set to select the 15-mer epitopes for each capsid proteins of the poliovirus strains. Selected epitopes were then subjected for the analyses of overlapping sequences and all such sequences were merged together to form a final epitope. Total of eight HTL epitopes from strains 3, 1, 2, 1, and 1 for VP0, VP1, VP2, VP3, and VP4 respectively was shortlisted for the final vaccine construct (Table 2). NetCTL server was utilised to predict MHC class I specific CTL epitopes. For each capsid protein three supertypes A2, A3, and B7 were selected to predict CTL epitopes (Lund O et al. [19]). Top five epitopes for each protein and supertype were chosen and only common CTL epitope among all the six strain was selected. Total of 17 CTL epitopes were then shortlisted for the final subunit vaccine construct (Table 3).

3.3. Sequence construction and B-cell epitope selection

All the predicted HTL and CTL epitopes were then joined using suitable linker molecule to construct final multi-epitope subunit vaccine (Figure 2). Two different linkers GPGPG for HTL epitopes and AAY for the CTL epitopes were used to join selected epitopes. With the help of EAAAK linker β defensin adjuvant of the length 45 amino acid was added at N-terminal of the vaccine. After joining all the epitopes and adjuvant the final size of the multi-epitope vaccine was 451 amino acids (Figure 2). BCPRED server was utilised to predict the linear B-cell epitopes. Total of nine 20-mer B-cell epitopes was shortlisted which has a score of more than 0.8 and showing more affinity for B-cell receptor (Table 4). Similarly, 13 conformational B-cell epitopes were mapped using the ElliPro server which has a minimum score of 0.5 and a maximum distance of 6Å.

Table 2. The predicted helper T-lymphocyte epitopes using the server is represented in the table with their percentile score and position.

Protein	Sequence	Position	Score
VP0	AASKQDFSQDPSKFTEPIK	39–57	0.37
VP0	PIKDVLIKTAPTLNSPNIE	55–73	0.11
VP0	PHQIINLRTNNSATIVLPYVN	261–281	0.2
VP1	KDTVQLRRKLEFFTYSRFDMEFTFVVT	111–138	0.4
VP2	NNWGIAILPLAPLNFAS	225–241	0.3
VP2	GEFKGSFTLDTNATNPARN	154–174	0.1
VP3	PGSNQYLTADNFQSPCALP	08–26	0.55
VP4	YTTINYYRDSASNAASKQD	26–44	0.83

Notes: The following epitopes were then selected for the construction of the subunit vaccine.

Table 3. The predicted CTL epitopes using the server for each capsid proteins that were further selected and taken up for the construction of multi-epitope based subunit vaccine.

Protein	A2	A3	B7
VP0	YLGRSGYTV	TINYTTINY	DPSKFTEPI
VP0	VLIKTAPTL	–	–
VP1	FDMELTFVV	VVNDHNPTK	FARGACVAI
VP1	AVPYYGPGV	–	–
VP2	LVLPYVNSL	GLFGQNMYY	LPYVNSLSI
VP2	YLGRAGYTV	NMFYHYLGR	–
VP3	FLFCGSMMA	YTHWAGSLK	DPRLSHTML
VP4	VLIKTAPAL	–	–

Table 4. The table represents the predicted linear B-cell epitopes of the constructed multi-epitope based subunit vaccine and their respective scores.

Position	Epitope	Score
373	ASNAASKQDGPGPGPGSNQY	1
281	VVTGPGPGAASKQDFSQDPS	1
320	KTAPTLNSPNIEGPGPGPHQ	1
397	NFQSPCALPGPGPGNNWGIA	1
419	PLAPLNFASGPGPGGEFKGS	1
352	VLPYVNGPGPGYTTINYYRD	1
182	YAVPYYGPGVAAYVVNDHNP	0.999
239	KAAYDPRLSHTMLGPGPGKD	0.998
66	KTAPTLAAYTINYTTINYAA	0.943

3.4. Antigenicity and allergenicity assessment

The designed subunit vaccine was evaluated using AllerTOP v.2.0 server. The server predicted no allergen nature for the multi-epitope subunit vaccine and also no IgE inducing epitope were reported. The probability of antigenicity predicted by ANTIGENpro was reported as 0.7687 which means a decreased probability in the antigenic nature of constructed subunit vaccine. Similarly, the VaxiJen server also reported the value of 0.6792 depicting the lowered probability of antigenic nature.

3.5. Physiochemical parameters evaluation

The nine different physiochemical parameters were evaluated using the ProtParam tool. The molecular weight of the subunit vaccine was reported as 48.44 kDa. The ideal vaccine should have a molecular weight between 40 and 50 kDa for the high uptake through lymphatic system and therefore, the vaccine comes under the ideal range. The theoretical pI of the candidate vaccine was reported as 9.26 which depict the basic nature of the subunit vaccine. The instability index of less than 40 represents the stable nature of the vaccine. The reported instability index of candidate vaccine was 32.50 and hence, the subunit vaccine can be considered as the stable vaccine. Also, the half-life of the multi-epitope vaccine in the mammalian reticulocytes was predicted as 30 h, which means it will take 30 h to remain half of its initial concentration inside the cell. Also, the aliphatic index of the vaccine was calculated which represents the thermos-stability at the variable temperature. The value of 73.28 represents the vaccine is thermostable at variable temperature. The grand average of hydropathicity value represents whether the candidate vaccine is hydrophilic or hydrophobic, positive score means hydrophobic in nature whereas negative score represents hydrophilic nature. The obtained value of −0.113 represents the hydrophilic nature of the constructed multi-epitope subunit vaccine.

3.6. Modelling and polio vaccine construct

The sequence of the construct was codon optimise by employing the JCAT tool for expression in *Escherichia coli* (strain K12). The sequence had a CAI value of 1 and a total GC content of 54%. The sequence was then restricted using BamHI and HindIII and ligated into pET28a(+) vector. The graphic map of vector with the construct is represented using SnapGene viewer in Figure 3A. This was performed to His-tag the construct for easy purification of the construct caccine. The polio vaccine sequence was modelled to study its interaction with T-cell receptor. The model with higher C-score was selected from the I-TASSER server. The polio vaccine was modelled based on top ten threading templates which were PDB id: 5mquA; 5jzg; 1h8tA; 1kb0A; 1dgi; 5c9a; 1d3eA; 5nf2A; 1pvcA and 1pov. The model was analysed using SAVES server and was found to have only 76% residues in the favoured region. Following this, the GalaxyRefine server was utilised and the final model was obtained. Ramachandran plot analysis of the final model was found to have 88.4 % residues in the favoured region, 8.7% in the allowed region and 3% in the outlier region

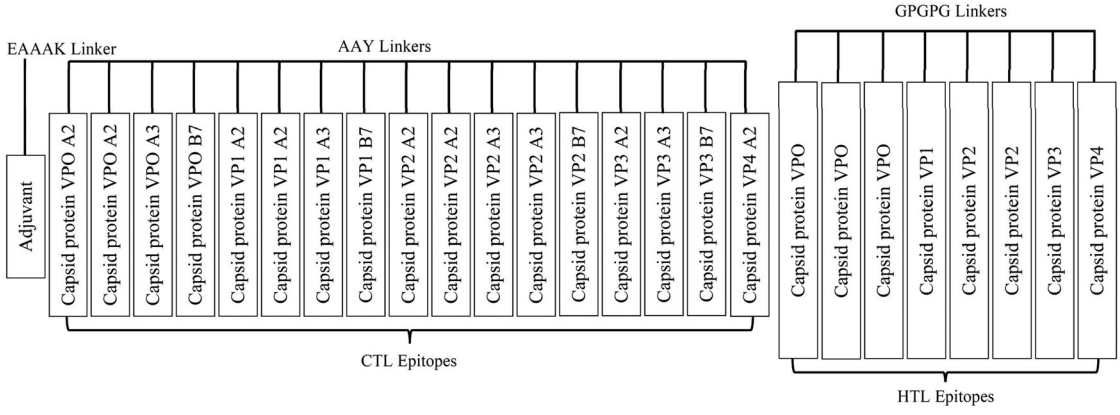

Figure 2. Construction of the multi-epitope vaccine. The figure represents the order of the predicted T-helper and T-cytotoxic epitopes that were chosen from the servers. The adjuvant is added at the N-terminal followed by the CTL epitopes joined together by AAY linkers. The HTL epitopes are present at the C-terminal and they are linked by the GPGPG linkers appropriately.

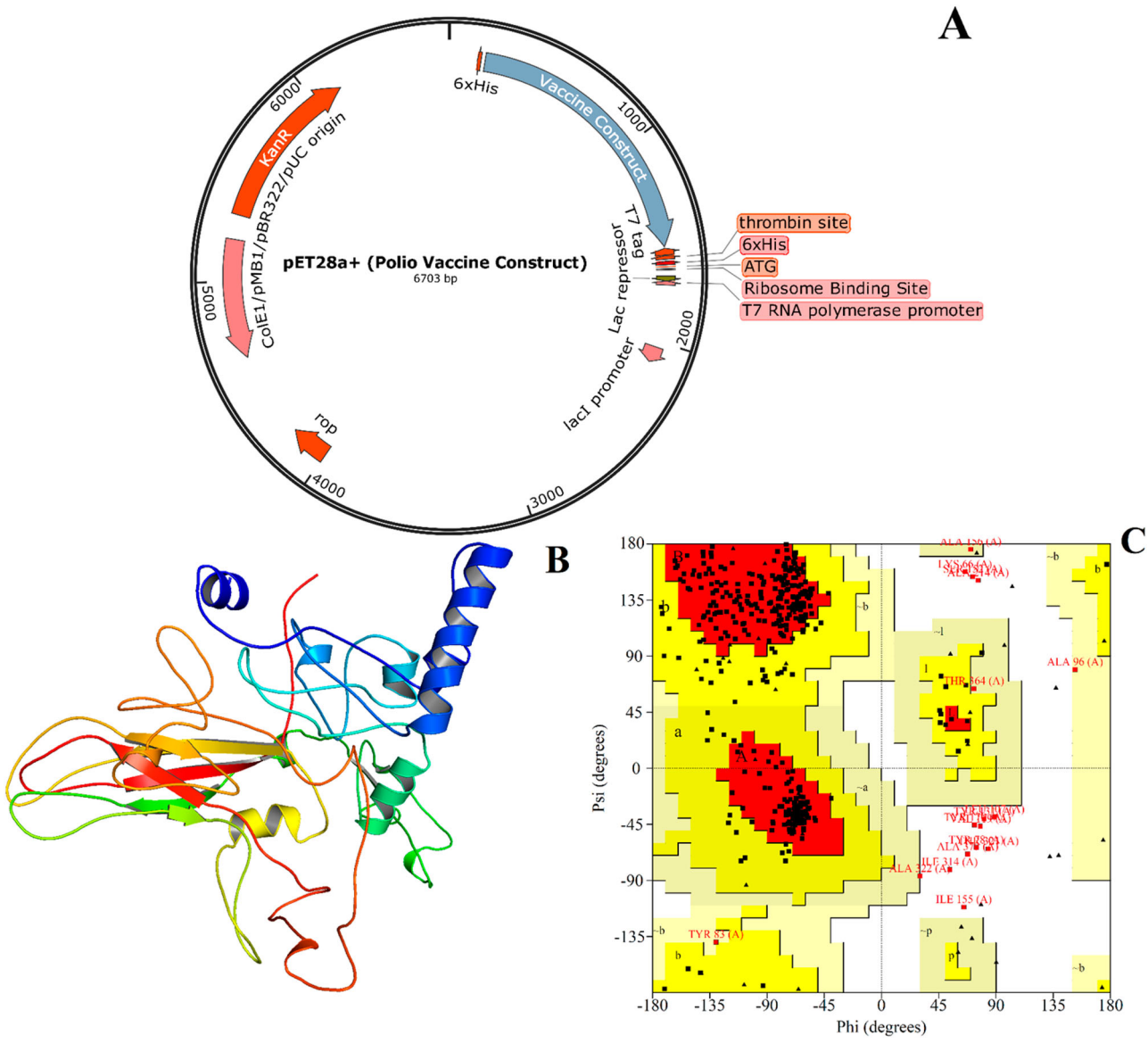

Figure 3. (Colour online) Modelling and Validation of vaccine construct. (A) Graphical map of protein construct between BamHI and HindIII of pET28a(+). The vector houses His tag in the N-terminal that can be purified by affinity chromatography. (B) Secondary structure representation of vaccine construct after refinement. (C) Ramachandran plot of refine vaccine construct exhibiting 88% amino acids in the favoured region and 9% in the allowed region along with 3% in the disallowed region.

Figure 3C. The model shown in Figure 3B passed the Errat quality and Verify-3D analysis. The model was then taken up for further docking and simulation studies.

3.7. Docking and molecular dynamics simulation of construct

The modelled vaccine construct was docked with the human TLR4 receptor protein. ClusPro server, used for protein-protein docking, showed the high binding affinity of modelled construct towards TLR4 protein chain. The model with the lowest binding energy represented in Figure 4A was then chosen for simulation. The receptor-vaccine construct was then prepared for molecular dynamic simulation for 20 ns. Initial energy minimisation and equilibration steps displayed stable structure during the 1 ns equilibration simulation. After confirming the stability of the structure and absence of structural anomalies, the production MD simulation for 20 ns was performed. To understand the stability of the complex, RMSD analysis was conducted. The RMSD analysis found that the receptor-vaccine remained stable throughout the simulation as seen in the plot in Figure 4B. The average RMSD values were calculated as 0.3548 ± 0.08 nm. Once the structure was found to be stable for 20 ns, the hydrogen bonds between the receptor and the vaccine was analysed to understand the affinity of the complex. The simulation showed an average of 15 hydrogen bonds between the receptor and the construct during the simulation run as represented in Figure 4C. The electrostatic energy of the complex was found to be $2.5 \times 10^5 \pm 885$ kJ/mol and the Lennard Jones energy was calculated to be $6.3 \times 10^5 \pm 885$ kJ/mol. The simulation analysis showed that the receptor-vaccine complex was stable throughout the simulation time and also formed sufficienthydrogen bonds (around 15 in average)between them, thus, proving the stability and affinity of the vaccine construct.

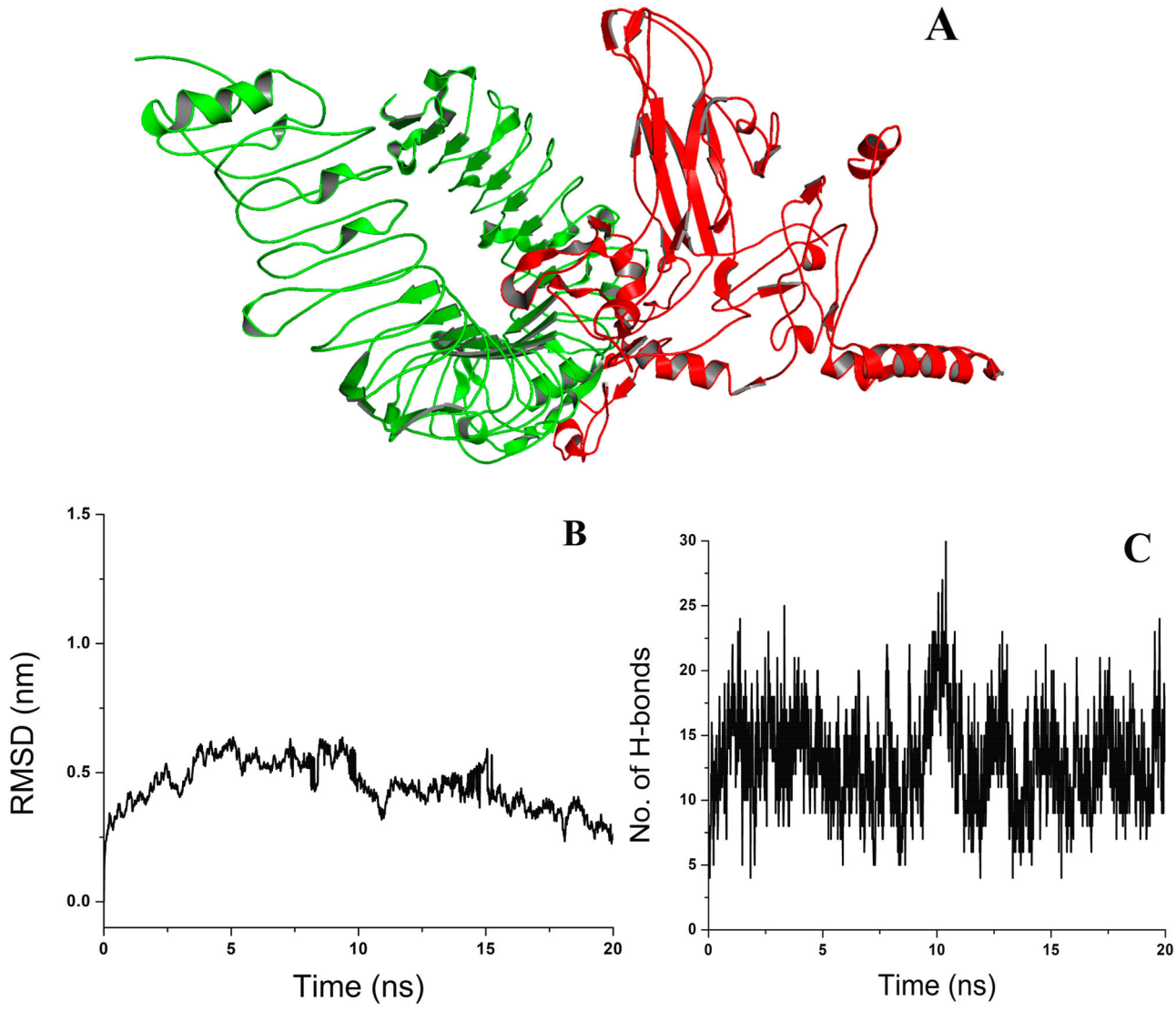

Figure 4. (Colour online) Docking and Simulation of polio vaccine construct with TLR4. (A) Docked pose of TLR4 receptor (in green) with vaccine construct (in red). (B) RMSD plot of receptor-vaccine construct displaying stable structure over 20 ns. (C) No. of hydrogen bonds formed over 20 ns was calculated to be an average of 15 corroborating the affinity of TLR4 towards the multi-epitope sub-unit vaccine.

3.8. Immune simulation

The immune simulation conducted on the C-ImmSim was consistent with the actual immune responses recorded as a general increase in the secondary response was observed. The primary response which is characterised by IgM was found to be very high in the study. The secondary and tertiary response which is characterised by the IgG1 + IgG2, IgM and IgG + IgM also had an increased response with respect to the decreasing antigen concentration. These results are analysed in Figure 5A,B. The trend indicated the development of an immune response on vaccine administration and subsequent clearance of the antigen on repeated exposures as found in Figure 5C,D. The repeated exposure of the antigen via injections triggered the levels of IgG1 to increase whereas the levels of IgM decreased. It was also observed that the levels of IFN-γ concentration and T_H cell population were maintained all through the course of the administration and exposure.

4. Discussion

Polio is a devastating disease that has been causing a huge loss of lives around the world. The novelty of this study is the construction of a multi-epitope vaccine for polio that confers immunity towards three strains of polio i.e. Type 1, Type 2 and Type 3. This study was conducted to design a multi-epitope sub-unit vaccineand it was carried out in two phases of designing and validation as shown in Figure 1. The designing of the vaccine started with the prediction of helper T epitopes and cytotoxic T cell epitopes. The capsid protein of three different strains of polio was taken up for study and fed to different servers for the epitope prediction. The HTL epitopes and CTL epitopes from capsid proteins were then chosen based on the top scores. The predicted epitopes were then selected and stitched together with the help of AAY and GPGPG linkers as shown in Figure 2. A β-defensin EAAAK liner was added to the N-terminal for better elicitation of the immune response. The B-cell epitope prediction found 9 epitopes of 20-mer in length

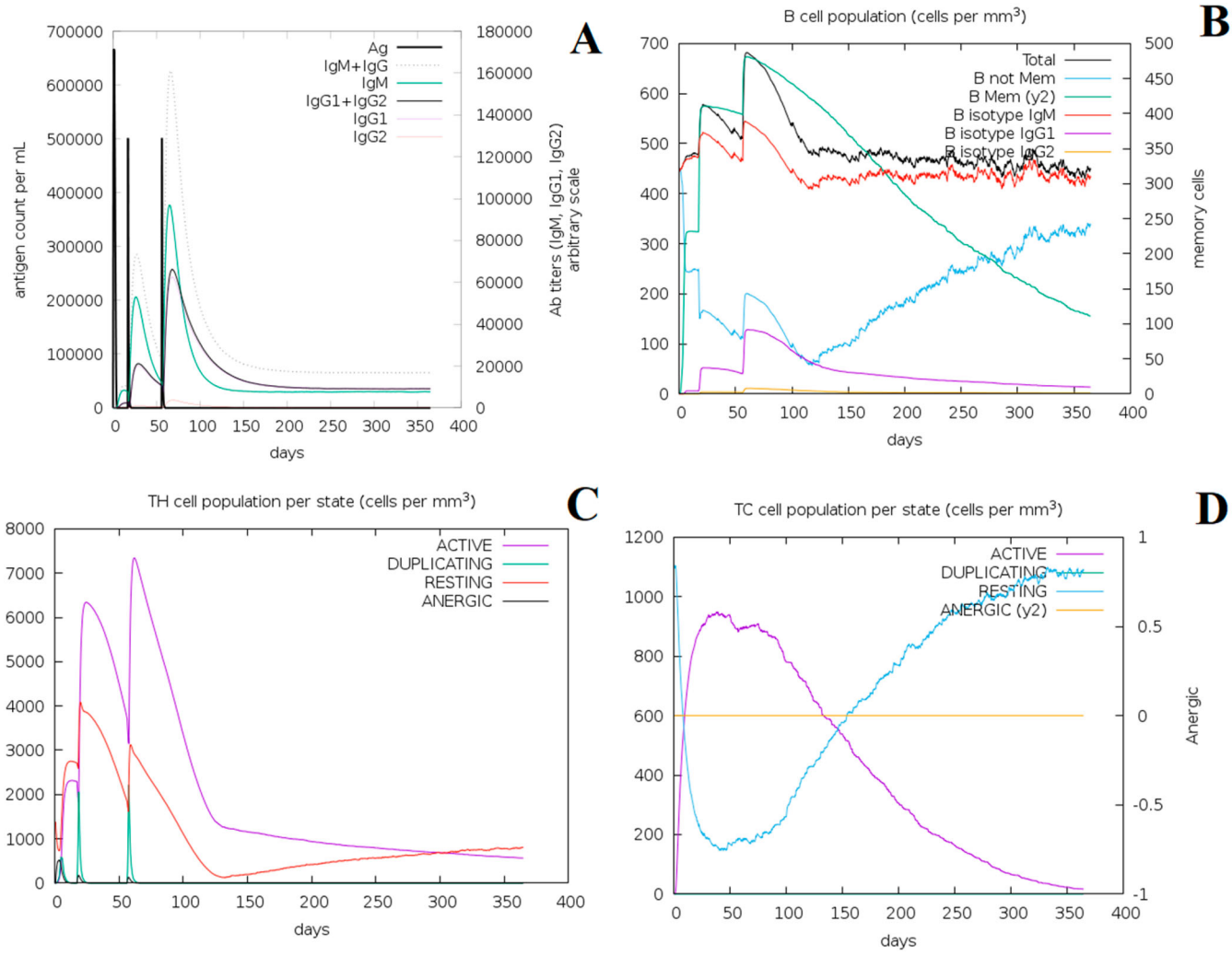

Figure 5. (Colour online) C-ImmSim Server results showing an *in-silico* immune simulation of the chimeric peptide. (A) Black like represents the antigen injection at various time intervals eliciting immunoglobulin production. Different sub-classes are represented as coloured peaks. (B) The evolution of the B-cell population is observed on the administration of three injections consequently. (C) The representation of the production of T-helper after the injections. (D) The production of T-cytotoxic cells on the administration of three consequent injections. RESTING refers to the cells that are not exposed to the antigens and ANERGIC represents the antigen tolerance of T-cells on repeated administration. The DUPLICATING Tc cell population is found to be zero when the vaccine is administered subsequently.

and 13 conformational B-cell epitopes which implied that the construct has affinity towards B-cell receptors.

The seconds part of the study involved validation of the construct for safety and efficacy. The construct with 451 amino acids was initially checked for allergenicity using different servers to confirm and substantiate the results. The construct was found to be a non-allergen in humans, and thus, confirmed to be safe for administration as predicted using in-silico tools. The physiological properties of the vaccine were determined to understand the nature of the protein and were found to have a molecular weight of 48.44 kDa with a pI value of 9.26. The molecular weight was well within the range of vaccines used before and the pI showed the basic nature of the vaccine [37,38]. The half-life period was also found to be 30 h for the vaccine. An average half-life time of around 20–30 h [39,40]. The instability index of the vaccine was analysed as 32.50 and these properties were well within the ranges of an ideal vaccine where instability index smaller than 40 is predicted to be stable and above 40 is considered to be unstable according to ExPasy server [41].

The amino acid sequence of the vaccine construct was then codon optimised for prokaryotic expression in *E. coli* K12 strain and cloned into His-tag pET28a(+) vector. The vaccine construct was further validated by modelling the 3D structure using the I-TASSER server. The server provided the best model using C-score which was then considered along with SAVES sever results. The best model for the vaccine construct was refined further and the model was validated using the Ramachandran plot. Ramachandran plots displays the residues that are energetically permitted by calculating the torsional angles of the amino acids. The residues that are within the range of permitted angles of each residue are categorised into favoured, allowed and disallowed region. The Ramachandran analysis showed 88% of the residues in the favoured region and 9% in the allowed region and around 3% in the disallowed region. The structure shows most of the residues in the favoured and allowed region and therefore was taken forward for docking and simulation. Once the structure was confirmed to be structurally refined, the affinity studies between the receptor TLR4 and the vaccine construct was performed. The

ClusPro protein-protein docking of the validated vaccine model against TLR4 receptor exhibited high binding affinity between each other. The receptor-vaccine complex with the lowest binding energy was extracted and was taken for simulation to study the stability of the complex. Molecular simulation studies usually provide information on the stability of the protein-protein complex as well as the hydrogen bond formation between the proteins. The authors found that the complex was highly stable throughout the simulation time of 20 ns. The average RMSD was 0.3548 ± 0.08 nm which proved that the receptor-vaccine complex is stable. The hydrogen bond analysis showed an average of 15 bonds during the 20 ns period. The stability and the affinity of the complex were also observed in the electrostatic and Lennard Jones potential energy analysed. To finally understand the capability of the sub-unit vaccine construct, the construct was immune simulated with three injections administered at $t = 0$, 17 and 56 days. The immune-simulation results were intriguing as it elicited the IgM and IgG immunoglobulins and also were successful in developing an immune response and clearing the antigen. The IgG1 + IgG2 and IgG + IgM responses were also increased and maintained even when the antigen concentration decreased with time. IgG1 concentration increased with repeated antigen exposure but the IgM concentration decreased with repeated exposure of antigen. The IFN-γ and T_H cell population remained constant even after repeated exposure of antigens. The development of successful memory cells was also seen as the sign of a good vaccine construct. The validation of the vaccine construct concludes that the vaccine construct design in this study is safe, stable and also exhibits affinity towards the toll-like receptors.

Concluding from the study, the authors present the design of a multi-epitope sub-unit vaccine that is predicted to be safe for administration in humans by in-silico tools. The vaccine construct from the study was capable of inducing an immune response and had stable interactions with toll-like receptors. The study also requires the real-time validation of the results, obtained in this work through immune-informatics analysis. Furthermore, safety and dosage administration studies should be thoroughly conducted which would involve the expression of the construct in a bacterial system, *in-vitro*/*in-vivo* evaluation of the vaccine and finally human trials. Only after taking all these experiments and results into consideration, the vaccine can be deemed safe and commercialised. In a nut shell, the vaccine construct in this study holds huge potential as a candidate to combat the huge crisis of polio burden pertaining in this world.

Acknowledgments

The authors thank Centre for Automation and Instrumentation (CAI), NIT Warangal for providing the necessary computational facility to carry out the work. The author SS and SR acknowledge research fellowship from NIT Warangal.

Disclosure statement

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

ORCID

Prakash Saudagar http://orcid.org/0000-0001-7211-0581

References

[1] Trevelyan B, Smallman-Raynor M, Cliff AD. The spatial dynamics of poliomyelitis in the United States: from epidemic emergence to vaccine-induced retreat, 1910–1971. Ann Assoc AM Geogr. 2005;95 (2):269–293.
[2] Atkinson W, Hamborsky J, McIntyre L, et al. 2017. Epidemiology and prevention of vaccine-preventable diseases. 13th ed. Public Health Foundation (US).
[3] Ohri LK, Marquess JG. Polio: will we soon vanquish an old enemy. Drug Benefit Trends. 1999;11(6):41–54.
[4] Kew OM, Sutter RW, de Gourville EM, et al. Vaccine-derived polioviruses and the endgame strategy for global polio eradication. Annu Rev Microbiol. 2005;59:587–635.
[5] Chander S, Pandey RK, Penta A, et al. Molecular docking and molecular dynamics simulation based approach to explore the dual inhibitor against HIV-1 reverse transcriptase and integrase. Comb Chem High Throughput Screen. 2017;20(8):734–746.
[6] Khatoon N, Pandey RK, Prajapati VK. Exploring Leishmania secretory proteins to design B and T cell multi-epitope subunit vaccine using immunoinformatics approach. Sci Rep. 2017;7(1):8285.
[7] Moyle PM, Toth I. Modern subunit vaccines: development, components, and research opportunities. ChemMedChem. 2013;8 (3):360–376.
[8] Vakili B, Eslami M, Hatam GR, et al. Immunoinformatics-aided design of a potential multi-epitope peptide vaccine against Leishmania infantum. Int J Biol Macromol. 2018;120:1127–1139.
[9] Consortium U. Uniprot: a worldwide hub of protein knowledge. Nucleic Acids Res. 2018;47(D1):D506–D515.
[10] Duvvuri VR, Duvvuri B, Alice C, et al. Preexisting CD4+ T-cell immunity in human population to avian influenza H7N9 virus: whole proteome-wide immunoinformatics analyses. PLoS One. 2014;9(3):e91273.
[11] Wang P, Sidney J, Kim Y, et al. Peptide binding predictions for HLA DR, DP and DQ molecules. BMC Bioinformatics. 2010;11(1):568.
[12] Nielsen M, Lund O. NN-align. An artificial neural network-based alignment algorithm for MHC class II peptide binding prediction. BMC Bioinformatics. 2009;10(1):296.
[13] Nielsen M, Lundegaard C, Lund O. Prediction of MHC class II binding affinity using SMM-align, a novel stabilization matrix alignment method. BMC Bioinformatics. 2007;8(1):238.
[14] Sidney J, Assarsson E, Moore C, et al. Quantitative peptide binding motifs for 19 human and mouse MHC class I molecules derived using positional scanning combinatorial peptide libraries. Immunome Res. 2008;4(1):2.
[15] Sturniolo T, Bono E, Ding J, et al. Generation of tissue-specific and promiscuous HLA ligand databases using DNA microarrays and virtual HLA class II matrices. Nat Biotechnol. 1999;17(6):555.
[16] Vita R, Overton JA, Greenbaum JA, et al. The immune epitope database (IEDB) 3.0. Nucleic Acids Res. 2014;43(D1):D405–D412.
[17] Larsen MV, Lundegaard C, Lamberth K, et al. Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. BMC Bioinformatics. 2007;8(1):424.
[18] Peters B, Bulik S, Tampe R, et al. Identifying MHC class I epitopes by predicting the TAP transport efficiency of epitope precursors. J Immunol. 2003;171(4):1741–1749.
[19] Lund O, Nielsen M, Kesmir C, et al. Definition of supertypes for HLA molecules using clustering of specificity matrices. Immunogenetics. 2004;55(12):797–810.
[20] Hajighahramani N, Nezafat N, Eslami M, et al. Immunoinformatics analysis and in silico designing of a novel multi-epitope peptide vaccine against Staphylococcus aureus. Infect Genet Evol. 2017;48:83–94.
[21] Mohan T, Sharma C, Bhat AA, et al. Modulation of HIV peptide antigen specific cellular immune response by synthetic α-and β-defensin peptides. Vaccine. 2013;31(13):1707–1716.

[22] Gasteiger E, Hoogland C, Gattiker A, et al. Protein identification and analysis tools on the ExPASy server. The proteomicsprotocols handbook. Totowa (NJ): Springer; 2005.
[23] Chen J, Liu H, Yang J, et al. Prediction of linear B-cell epitopes using amino acid pair antigenicity scale. Amino Acids. 2007;33 (3):423–428.
[24] EL-Manzalawy Y, Dobbs D, Honavar V. Predicting linear B-cell epitopes using string kernels. J Mol Recognition Inter. 2008;21 (4):243–255.
[25] Ponomarenko J, Bui H-H, Li W, et al. Ellipro: a new structure-based tool for the prediction of antibody epitopes. BMC Bioinformatics. 2008;9(1):514.
[26] Dimitrov I, Bangov I, Flower DR, et al. AllerTOP v. 2—a server for in silico prediction of allergens. J Mol Model. 2014;20(6):2278.
[27] Wold S, Jonsson J, Sjörström M, et al. DNA and peptide sequences and chemical processes multivariately modelled by principal component analysis and partial least-squares projections to latent structures. Anal Chim Acta. 1993;277(2):239–253.
[28] Magnan CN, Baldi P. SSpro/ACCpro 5: almost perfect prediction of protein secondary structure and relative solvent accessibility using profiles, machine learning and structural similarity. Bioinformatics. 2014;30(18):2592–2597.
[29] Doytchinova IA, Flower DR. Vaxijen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinformatics. 2007;8(1):4.
[30] Grote A, Hiller K, Scheer M, et al. JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. Nucleic Acids Res. 2005;33(suppl_2):W526–W531.
[31] Carbone A, Zinovyev A, Képes F. Codon adaptation index as a measure of dominating codon bias. Bioinformatics. 2003;19(16):2005–2015.
[32] Yang J, Yan R, Roy A, et al. The I-TASSER Suite: protein structure and function prediction. Nat Methods. 2015;12(1):7.
[33] Shin W-H, Lee GR, Heo L, et al. Prediction of protein structure and interaction by GALAXY protein modeling programs. Bio Design. 2014;2(1):1–11.
[34] Vajda S, Yueh C, Beglov D, et al. New additions to the ClusPro server motivated by CAPRI. Proteins Struct Funct Bioinformatics. 2017;85 (3):435–444.
[35] Sasidharan S, Saudagar P. Biochemical and structural characterization of tyrosine aminotransferase suggests broad substrate specificity and a two state folding mechanism in Leishmania donovani. FEBS Open Bio. 2019;9(10):1769–1783.
[36] Rapin N, Lund O, Bernaschi M, et al. Computational immunology meets bioinformatics: the use of prediction tools for molecular binding in the simulation of the immune system. PLoS One. 2010;5(4):e9862.
[37] Urrutia-Baca VH, Gomez-Flores R, De La Garza-Ramos MA, et al. Immunoinformatics approach to design a novel epitope-based oral vaccine against helicobacter pylori. J Comput Biol. 2019;26 (10):1177–1190.
[38] Ghahremanifard P, Afzali F, Rostami A, et al. Designing a novel multi-epitope T vaccine for "targeting protein for Xklp-2" (TPX2) in hepatocellular carcinoma based on immunoinformatics approach. Int J Pept Res Ther. 2019: 1–10. https://doi.org/10.1007/s10989-019-09915-2.
[39] Kalita P, Lyngdoh DL, Padhi AK, et al. Development of multi-epitope driven subunit vaccine against Fasciola gigantica using immunoinformatics approach. Int J Biol Macromol. 2019;138:224–233.
[40] Jenssen H, Aspmo SI. Serum stability of peptides. Peptide-based drug design. New York: Springer; 2008. p. 177–186.
[41] Guruprasad K, Reddy BB, Pandit MW. Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. Protein Eng Des Sel. 1990;4(2):155–161.