ORIGINAL ARTICLE

Designing of an epitope-based peptide vaccine against walking pneumonia: an immunoinformatics approach

P. Ambili Unni[1] · A. M. Mohamed Thoufic Ali[2] · Madhusmita Rout[1] · A. Thabitha[1] · S. Vino[3] · S. Sajitha Lulu[1]

Abstract

Mycoplasma pneumoniae is a substantial respiratory pathogen that develops not only pneumonia but also other respiratory diseases, which mimic viral respiratory syndromes. Nevertheless, vaccine development for this pathogen delays behind as immunity correlated with protection is now predominantly unknown. In the present study, an immunoinformatics pipeline is utilized for epitope-based peptide vaccine design, which can trigger a critical immune response against *M. pneumoniae*. A total of 105 T-cell epitopes from 12 membrane associated proteins and 7 T-cell epitopes from 5 cytadherence proteins of *M. pneumoniae* were obtained and validated. Thus, 18 peptides with 9-mer core sequence were identified as best T-cell epitopes by considering the number of residues with > 75% in favored region. Further, the crucial screening studies predicted three peptides with good binding affinity towards HLA molecules as best T-cell and B-cell epitopes. Based on this result, visualization, and dynamic simulation for the three epitopes (WIHGLILLF, VILLFLLLF, and LLAWMLVLF) were assessed. The predicted epitopes needs to be further validated for their adept use as vaccine. Collectively, the study opens up a new horizon with extensive therapeutic application against *M. pneumoniae* and its associated diseases.

Keywords *Mycoplasma pneumoniae* · Epitope · Peptide · Vaccine · Immunoinformatics

Introduction

Mycoplasma pneumoniae is significant respiratory pathogen that produces diverse severity from mild upper respiratory tract infection to severe atypical pneumonia (walking pneumonia), predominantly in children and young adults. The disease habitually progresses with the symptoms of respiratory tract and lung infections, mortality rate of the disease is approximately 10–30% in the hospitalized community [1–4]. The organism also produces a wide spectrum of extrapulmonary manifestations like neurological, dermatological, congenital tract infections, etc, [5, 6]. Several reports have been published on the network of *M. pneumoniae* with community-acquired infections [6, 7]. Even though wide array of data from the investigating reports are available, the effective therapeutic management of *M. pneumoniae* infections still remains as a challenge.

Mycoplasma is cell wall deficient simplest self-limiting bacteria, capable of cell-free survival [8, 9]. *M. pneumoniae* has an extremely small genome (approx. 816 kilo base-pairs) responsible for its limited biosynthetic capabilities and lethargic replication rate [10]. *M. pneumoniae* is transmitted by the respiratory droplets via close contact [11] with respiratory tract mucosa leading to various effects like local inflammation and additional manifestations. The primary phase of *M. pneumoniae* respiratory tract infection implicates cytadherence of the organism towards the ciliated columnar epithelium, which defends the organism from local cytotoxic effects and mucociliary clearance [6].

A specialized organelle mediates the cellular adherence to sialoglycoproteins and sulphated glycolipids. It has a central

---

✉ S. Sajitha Lulu
ssajithalulu@vit.ac.in

1 Department of Biotechnology, School of Bio Sciences and Technology, VIT University, Vellore, Tamil Nadu 632014, India

2 Department of Integrative Biology, School of Bio Sciences and Technology, VIT University, Vellore, Tamil Nadu 632014, India

3 Department of Bio-Sciences, School of Bio Sciences and Technology, VIT University, Vellore, Tamil Nadu 632014, India

core which consists of dense filaments and a tip-like structure comprised of adhesins and accessory proteins [12, 13]. P1-adhesin, P30-adhesin and P65-adhesin are the crucial proteins involved in receptor binding. The proteins A, B, and C as well as high molecular protein-1 (HMW-1), HMW-2, HMW-3 act as accessory proteins, interacting with P1, P30, P65 and P165 adhesins, and thus facilitating cytadherence [14, 15]. These proteins could act as potential druggable targets for the effective therapeutic and vaccine approaches against *M. pneumoniae*.

Mycoplasma pneumoniae is highly susceptible to macrolides, tetracyclines, and the new quinolone antibiotics. Macrolide resistance condition have been reported from China and from European countries such as France and Germany and other parts of the world including the USA [16–19]. Macrolide resistance in *M. pneumoniae* is associated with 23SrRNA genes on nucleotide mutations which could increase minimal inhibitory concentrations to azithromycin, erythromycin and clarithromycin. A successful vaccine development will not only enable us to avert the critical situations of *M. pneumoniae* infection but also serve as preventive approach against disastrous outbreaks, predominantly in closed community settings [20].

Currently, epitope-based vaccine design, has currently drawn much attention in infectious diseases treatment [21–23]. Immunization based on the epitope-based vaccines is dynamic in stimulation of humoral and/or cellular arms of the immune system [24]. These kind of vaccines are comprised of extremely immunogenic T and B cell epitopes, which aggravate the cytotoxic T cells (CTL), B or T helper cells to particular epitopes [25, 26]. T helper cells and B cells play a significant role in stimulation of a protective immune response in numerous bacterial infections; therefore, determination of peptides that lead to B and T cell responses is critical for the design of potent epitope-based peptide vaccines [27, 28]. The epitope-based peptide vaccines have some prospective advantages, such as cost effective production, ability to select the type of immunity and to increase safety [23]. Immunoinformatics or computational biology approaches have added an inevitable contribution towards designing epitope-based peptide vaccines.

In the present study, we have designed the epitope-based peptide vaccine from membrane-associated proteins and cytadherence proteins of *M. pneumoniae* with the help of immunoinformatics approach. The presented vaccine strategy could serve as promising pathogen specific candidates with extensive therapeutic applications against pneumonia and its associated infectious disease.

Materials and methods

Dataset construction

We used the dataset of 12 membrane associated proteins reported by Gupta et al. [29] and 5 cytadherence proteins of *M. pneumonia* from Uniprot database (http://www.uniprot.org/). The dataset was then improved by adding the information's like protein accession ID, protein family name, subcellular location and its corresponding references.

Unraveling protein antigenicity and allergenicity

To identify the antigenicity, all the 17 proteins were submitted to VaXiJen v 2.0 server which helps in the prediction of potent antigens and subunit vaccines with preset parameters [30]. Similarly, the allergenicity prediction for all the 17 proteins were done by AllerHunter server [31] by opting plain sequence format and choosing bacteria as the target organism. The 17 proteins with highest antigenicity scores were selected for further evaluation. The allergens and non-allergens with high sensitivity and specificity for the dataset were thus predicted.

Primary and secondary structure analysis

SOPMA server [32] was used to identify secondary structural elements and physiochemical properties of all the proteins. The major properties extracted include transmembrane helices, solvent accessibility, globular regions, random coil and coiled coil region.

T cell epitope prediction from the conserved sequences

The T cell epitope prediction is highly crucial for rational vaccine design and was done using by EpiDOCK server [33], the first structure based server for the prediction of MHC class II binding. MHC class II present the peptides derived from endocytosed extracellular proteins and is encoded by three loci, HLA DR, HLA DQ and HLA DP. EpiDOCK predicts the binding of 23 most frequent human MHC class II alleles, 12 HLA DR, 6 HLA-DQ, and 5 HLA DP alleles. MHC class II present the peptides derived from endocytosed extracellular proteins and is encoded by three loci, HLA DR, HLA DQ and HLA DP. The binding cleft of MHC Class II proteins are open ended and allows longer peptides to bind. The FASTA format sequences of the proteins were given as input to identify the binding of 23 MHC class II alleles by assigning threshold to each

HLA molecules. Maximum number of binders as epitopes with MHC class II alleles were thus identified and subjected for the further analysis.

Modeling and validation of best selected epitopes

To perform docking and simulation study, all the best conserved epitopes were inserted to PEP-FOLD server [34], which is a de novo method designed to predict the peptide structures from amino acid sequences. The best peptide models provided by the server were selected for further analysis.

The stereochemical properties of predicted models were retrieved from RAMPAGE server http://mordred.bioc.cam.ac.uk/~rapper/rampage.php, which provides accurate information about the residues present in the favored, allowed and disallowed regions. The PDB coordinates of the predicted epitope models were given as input and Ramachandran plot was then constructed to validate the structure. The predicted epitope models were subjected to further analysis by considering a percentage score (> 75%) of residues present in the favored region.

Molecular docking study

The binding affinity between HLA molecules and predicted epitopes were identified through docking study using CLUS-PRO server [35], a web based program for computational docking of protein structures. The HLA proteins and predicted peptide structures were given as input to calculate the least binding energy score. Docking algorithm evaluate several complexes retaining a preset number with favorable surface complementarities. Filtering method is applied to these set of structures for selecting those with good electrostatic and desolvation free energies. Docking studies were conducted between the predicted epitopes and HLA DP (3LQZ), HLA DQ (1S9V), HLA DR (2G9H) molecules. The PDB coordinates for HLA DP, DQ and DR were obtained from the Research Collaboratory for Structural Bioinformatics (RCSB) protein database. Chain A and B of HLA DP, HLA DQ and HLA DR were considered for docking with predicted epitopes and the binding energies were analyzed.

Molecular dynamics simulation studies

All the molecular dynamic simulations were performed using GROMACS 4.5.2 package [36] with OPLS-AA force field [37]. Firstly, all the models were solvated with 1.09 nm simple point charge (SPC) water entrenched in the simulation boxes. After that, all the systems were examined and subjected to steepest descent energy minimization until attaining a tolerance of 100 kJ/mol. Subsequently, the solvent molecules were equilibrated with the stable protein at 300 K for a period; the entire system was gradually relaxed and heated up to 300 K. Finally, 15-ns MD simulations were carried out under the normal temperature and pressure along with coupling time constant of 1.0 ps. The particle mesh Ewald method was utilized to treat long-range electrostatic interactions. van der Waals force was maintained at 1.4 nm. The g_rmsd command was used to analyze RMSD trajectory file and corresponding graph was generated by using GRACE software.

Potential B cell epitope prediction

The module of Kolaskar and Tongaonkar [38] from the Immune Epitope Database (IEDB) (http://tools.immuneepitope.org/bcell/) was used to predict linear B-cell epitopes. The results predicted the epitopes with around 75% accuracy. Furthermore, the properties of B-cell epitopes, which includes Emini surface accessibility prediction [39], Parker hydrophilicity prediction [40] karplus and schulz flexibility prediction [41] and Bepipred linear epitope prediction [42] were predicted using IEDB. The study also confirmed that antigenic parts of the proteins are confined to beta turn regions [43] were predicted by chou and fasman (http://tools.iedb.org/bcell/) [44] beta prediction tool.

Ab-initio modeling and refinement

To predict the three-dimensional (3D) structure of the proteins of selected best predicted epitopes, I-TASSER server http://zhanglab.ccmb.med.umich.edu/I-TASSER/ [45] was used. FASTA format data of the proteins were given as input and the retrieved output comprises secondary structure prediction, predicted solvent accessibility of each residue, homologous template proteins detected by threading and structure alignments, structure based functional annotations for enzyme classification, gene ontology terms and protein ligand binding sites. All predictions were associated with a confidence score which shows the accuracy of predictions.

Protein–Ligand docking

Molecular docking analysis was performed to identify the inhibitory potential of the best epitopes predicted with the existing drugs for *M. pneumoniae* infection, which can be potential druggable targets. Azithromycin, telithromycin, tetracycline, doxycycline and roxithromycin drugs under the macrolide class of antibiotics were used in the study. Protein–ligand docking was performed using CLUSPRO server [35]. PDB coordinates for the drugs were obtained from PubChem (https://pubchem.ncbi.nlm.nih.gov/) and the energy minimization with MOPAC server (http://www.openmopac.net) and the docked complexes were visualized with PyMOL software [46].

Results

Structure prediction and analysis of antigenic protein

The choice of immunogen or epitope is the initial step for effective vaccine design; henceforth, to find out the most probable antigenic protein, the protein sequences of *M. pneumoniae* were retrieved for dataset construction. A total of 12 membrane associated proteins from the study reported by Gupta et al., and 5 cytadherence proteins from UniprotKB database were retrieved and tabulated in Table 1. The antigenicity of all proteins is identified from overall score generated by the VaxiJen server [30]. The scores of 12 membrane associated proteins and 5 cytadherence proteins from *M. pneumoniae* revealed their potentiality to induce immune response. Moreover, putative adhesion P1-like protein (UniprotKB ID: P75491) and the uncharacterized lipoprotein MG186 homolog protein (UniprotKB ID: P75265) showed significantly highest

antigenic scores of 0.9674 and 0.9342, respectively. All the query sequences could be categorized as non-allergens by using AllerHunter server [31]. Table 2 summarizes the antigenic score and allergenicity assessment of all the selected proteins. All the antigenic proteins were analyzed for the secondary structural features like alpha helixes, beta sheets, beta turns, and random coils are depicted in Table 3.

Identification of T-cell epitopes

Antigenic peptide vaccines consist of T-cell epitopes. The prime prerequisite for an efficient peptide to enact as a T-cell epitope is that it should bind to major histocompatibility complex (MHC) protein in order to promote immune response. Hence, T-cell epitopes in 12 membrane associated proteins and 5 cytadherence proteins of *M. pneumoniae* were predicted. A total of 113 antigenic peptides with 9-mer core sequences in the 12 membrane associated proteins and 5 cytadherence proteins of *M. pneumoniae* were identified to be T-cell epitopes using EpiDock server (S1

Table 1 Membrane associated and cytadherence proteins of *M. pneumoniae*

S. no	Accession number	Protein family name	Sub-cellular location	References
1	P75556	Transmembrane	Outer membrane	Gupta et al. [27]
2	P75588	Transmembrane	Extra-cellular	
3	Q50327	Hydrolases	Outer membrane	
		Sodium binding		
		Transmembrane		
4	P75056	Lipid binding protein	Extra-cellular	
5	P75340	Lipid binding protein	Outer membrane/extra cellular	
6	P75194	Lipid binding protein	Extra-cellular	
		Zinc binding		
		Transferases-Glycosyltransferases		
		Electrochemical potential -driven transporters		
7	P75198	Lipid binding protein	Extra-cellular	
		Hydrolases		
8	P75195	Lipid binding protein	Extra-cellular	
9	Q50336	Hydrolases	Extra-cellular	
10	P54125	Zinc binding	Extra-cellular	
		Hydrolases(acting on peptide bonds)		
		Hydrolases(acting on ester bonds)		
11	P75265	Hydrolases	Extra-cellular	
		Nucleases		
		Endonucleases		
12	P75491	Cytadherence protein	Extra-cellular	Chaudhry et.al [49]
13	Q50365	Cytadherence high molecular protein 1	Cytadherence protein	
14	P75471	Cytadherence high molecular protein 2	Cytadherence protein	
15	Q50360	Cytadherence high molecular protein 3	Cytadherence protein	
16	P75330	Cytadherence protein	Cytadherence protein	
17	P0CJ81	Cytadherence protein	Cytadherence protein	

Table 2 Antigenicity and allergenicity prediction

S. no	Protein	Antigenic score	Allergenicity prediction
1	P75556	0.4925	Not allergic
2	P75588	0.6274	Not allergic
3	Q50327	0.6902	Not allergic
4	P75056	0.3793	Not allergic
5	P75340	0.796	Not allergic
6	P75194	0.3509	Not allergic
7	P75198	0.4009	Not allergic
8	P75195	0.3714	Not allergic
9	Q50336	0.4753	Not allergic
10	P54125	0.2442	Not allergic
11	P75265	0.9342	Not allergic
12	P75491	0.9674	Not allergic
13	Q50365	0.5857	Not allergic
14	P75471	0.4811	Not allergic
15	Q50360	0.5348	Not allergic
16	P75330	0.7167	Not allergic
17	P0CJ81	0.4978	Not allergic

Table 3 Secondary structure prediction for all proteins

S. no	Protein	Alpha helix (%)	Beta sheet (%)	Turn (%)	Coil (%)
1	P75556	42.62	21.75	8.45	27.18
2	P75588	22.89	18.66	8.71	49.75
3	Q50327	78.26	8.70	1.93	11.11
4	P75056	43.92	19.38	9.48	27.22
5	P75340	31.88	15.36	9.57	43.19
6	P75194	31.12	24.21	9.51	35.16
7	P75198	21.64	27.33	10.71	40.32
8	P75195	22.19	25.50	6.62	45.70
9	Q50336	36.54	19.26	10.20	33.99
10	P54125	56.14	16.20	9.33	18.33
11	P75265	39.87	13.62	9.97	36.54
12	P75491	11.18	24.73	10.11	53.98
13	Q50365	32.85	7.30	8.03	51.82
14	P75471	34.81	8.89	10.37	45.93
15	Q50360	29.47	19.74	7.56	43.22
16	P75330	83.50	3.63	2.70	10.18
17	P0CJ81	13.84	18.45	8.04	59.67

Table). The server identifies likely overlapping 9-mers on the given sequence by using docking score-based quantitative matrix (DS-QM) and generates the score by predicting peptide-MHC class II binding based on sensitivity, specificity, accuracy and area under curve (AUC) using coordinates sensitivity/1-specificity.

Superior immune response predominantly depends on successful recognition of the epitopes by HLA molecules with substantial affinity. A reliable T-cell epitope should interact with maximum number of HLA alleles in order to induce strong immune response. Therefore, the best peptides from each protein with highest number of binding HLA-DP, HLA-DQ and HLA-DR alleles were selected as putative T-cell epitope candidates. The details of predicted putative T-cell epitopes along with their respective binding HLA-DP, HLA-DQ and HLA-DR alleles are shown in Table 4.

Structure prediction and validation of putative T-cell epitopes

The 3D structures of WIHGLILLF epitope (Fig. 1), VILLFLLLF epitope (Fig. 2), and LLAWMLVLF (Fig. 3) were predicted and their scrambled versions were generated through the PEP-FOLD de novo modeling server and validated using RAMPAGE web-based tool.

Through RAMPAGE analysis, the number of amino acid residues (in percentage) was obtained in the favored

Table 4 Putative T-cell epitopes with interacting MHC class II molecules

Protein	Epitope	No. of binding allele
P75556	FKTNLNLAF	17
	KITVYKLIF	17
	KLNDPKPNF	17
	NDPKPNFWF	17
P75588	**WIHGLILLF**	17
Q50327	LLAFVILLF	17
	VILLFLLLF	17
P75056	NKEDAKELF	16
P75340	LLDAAIPAF	17
P75194	FAVLEILLF	17
P75198	SAVLELPLF	17
P75195	FAILEIPLF	17
Q50336	VKQMALDAF	16
	QDQQIKDHF	16
P54125	IQQNQKPDF	17
	AYITSIPHF	17
P75265	RDYTTKNEF	15
	FKSNPKQTF	15
P75491	QLGGWLVTF	15
	LGAGEKSRF	15
P75330	PPRRKLKLF	16
	LLAWMLVLF	16
P0CJ81	KDKLDNAIF	16
	NLQIMKQNF	16
Q50365	TEENPEQIF	15
P75471	IKNLENEIK	11
Q50360	DKERAKLAK	12
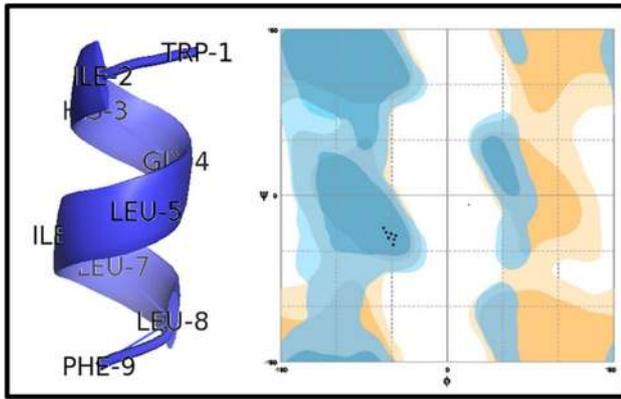
Fig. 1 Model and Validation of peptide WIHGLILLF
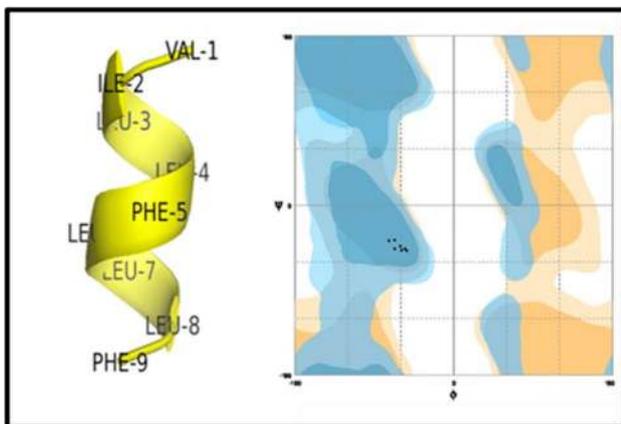
Fig. 2 Model and Validation of peptide VILLFLLLF
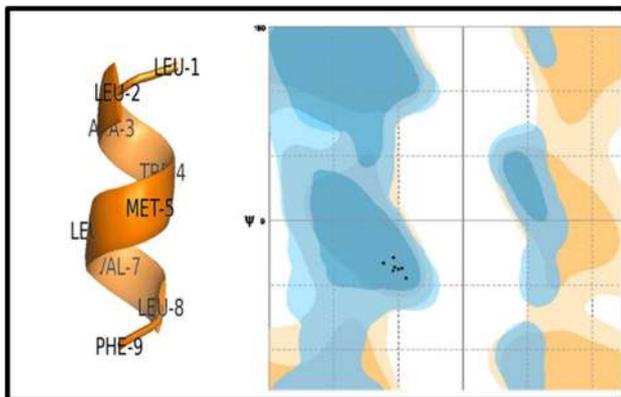
Fig. 3 Model and Validation of peptide LLAWMLVLF

region and displayed in Table 5. Hence, the epitopes FKTNLNLAF, KITVYKLIF, WIHGLILLF, LLA-FVILLF, VILLFLLLF, LLAFVILLF, LLDAAIPAF, FAVLEILLF, SAVLELPLF, FAILEIPLF, VKQMALDAF,

Table 5 Validation of predicted T cell epitopes by RAMPAGE analysis

Protein	Epitope	No. of binding allele	Ramachandran plot score (%)
P75556	FKTNLNLAF	17	100
	KITVYKLIF	17	100
	KLNDPKPNF	17	57.10
	NDPKPNFWF	17	71.4
P75588	**WIHGLILLF**	17	100
Q50327	LLAFVILLF	17	100
	VILLFLLLF	17	100
P75056	NKEDAKELF	16	71.4
P75340	LLDAAIPAF	17	85.7
P75194	FAVLEILLF	17	100
P75198	SAVLELPLF	17	85.7
P75195	FAILEIPLF	17	100
Q50336	VKQMALDAF	16	85.7
	QDQQIKDHF	16	100
P54125	IQQNQKPDF	17	100
	AYITSIPHF	17	28.60
P75265	RDYTTKNEF	15	71.4
	FKSNPKQTF	15	71.4
P75491	QLGGWLVTF	15	85
	LGAGEKSRF	15	57.10
P75330	PPRRKLKLF	16	85.7
	LLAWMLVLF	16	100
P0CJ81	KDKLDNAIF	16	85.7
	NLQIMKQNF	16	100
Q50365	TEENPEQIF	15	71.4
P75471	IKNLENEIK	11	71.4
Q50360	DKERAKLAK	12	100

QDQQIKDHF, IQQNQKPDF, PPRRKLKLF, LLAWM-LVLF, KDKLDNAIF, NLQIMKQNF, and DKERAKLAK, showed the number of residues with > 75% in favored region were subjected to docking studies with HLA molecules.

Binding energy determination of HLA-epitope interaction

Using CLUSPRO server, the binding models for the predicted T cell epitopes to their respective HLA-DP, -DQ and –DR molecules were generated and displayed in Table 6. The predicted epitopes, WIHGLILLF with HLA-DP (Fig. 4), VILLFLLLF with HLA-DQ (Fig. 5), and LLA-WMLVLF with HLA-DR (Fig. 6), showed binding energies of − 915.8 kcal/mol, − 982.1 kcal/mol, and − 893 kcal/mol, respectively. The 2D structure representation and ligplot analysis (Figs. 7, 8, 9) depicts their hydrophobic interactions.

Table 6 Docking score of epitopes with HLA -DP, -DQ and –DR

Protein	Peptide	Binding energy (Kcal/mol)		
		DP	DQ	DR
P75556	FKTNLNLAF	646.2	− 653.3	− 128.3
	KITVYKLIF	− 776.8	− 758.4	− 773.5
P75588	**WIHGLILLF**	− 915.8	− 949.2	− 630.8
Q50327	LLAFVILLF	− 832.8	− 881.7	− 834.8
	VILLFLLLF	− 827.9	− 982.1	− 855.5
P75340	LLDAAIPAF	− 674.3	− 812.2	− 678.7
P75194	FAVLEILLF	− 750.7	− 870.7	− 755.2
P75198	SAVLELPLF	− 718.8	− 847.9	− 719.7
P75195	FAILEIPLF	− 791.3	− 911.3	− 850.2
Q50336	VKQMALDAF	− 600	− 678.6	− 624.1
	QDQQIKDHF	− 592.9	− 672.4	− 582.7
P54125	IQQNQKPDF	− 537.6	− 569	− 465.7
P75330	PPRRKLKLF	− 883.1	− 806.8	− 889.9
	LLAWMLVLF	− 846.1	− 953.8	− 893
P0CJ81	KDKLDNAIF	− 596.4	− 704.5	− 624.7
	NLQIMKQNF	− 629.1	− 603.6	− 643.1
Q50360	DKERAKLAK	− 503.6	− 467.5	− 501.2
P75491	QLGGWLVTF	− 783.5	− 937.2	− 699.8
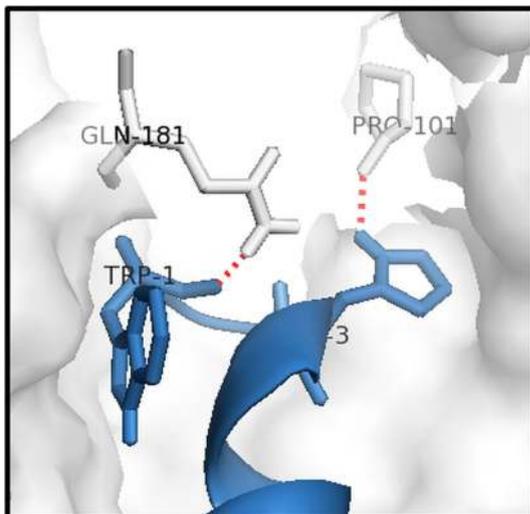
Fig. 4 Docking of peptide (WIHGLILLF) derived from protein (Uniprot ID: P75588) with HLA DP

Molecular dynamics simulation of protein and protein-peptide complex

The dynamic behavior of HLA proteins (DP, DQ and DR) as well as the HLA-peptide complexes was studied using Molecular Dynamic Simulation studies. We analyzed RMSD, and total energy in the native protein as well as in protein-peptide complexes. (Fig. 10) shows the RMSD of HLA DP was equilibrated after 5 ns with RMSD value of

0.15 nm, while protein-peptide complex was equilibrated at 5 ns with RMSD value of 0.15 nm. Figure 11 depicts the RMSD value of HLA DQ was equilibrated after 12.5 ns with RMSD value of 0.16 nm, while protein-peptide complex was equilibrated at 12.5 ns with RMSD value of 0.17 nm. RMSD of HLA DR was equilibrated and further RMSD value of 0.17 nm at 6 ns, while protein-peptide complex was equilibrated at 5.5 ns with RMSD value of 0.157 nm and showed in Fig. 12.

Total energy of protein-peptide complex was higher than native protein in the case of HLA DP- WIHGLILLF (Fig. 13), while the total energy of HLA DQ- VILLFLLLF complex was found to be lower than the HLA-DQ protein (Fig. 14), however HLA DR- LLAWMLVLF complex showed higher energy than HLA DR Protein (Fig. 15).

Identification of B-cell epitopes

Prediction and identification of B-cell epitopes in the target antigens is one of the main steps in epitope-driven vaccine design. Putative B-cell epitopes have various features which are essential for the successful recognition by B-cells. These characteristics include hydrophilicity, surface accessibility and beta turn prediction. Thus, to obtain B-cell epitope candidates in the predicted putative T-cell epitopes of proteins P2, P3, and P30, in silico identification of the B-cell epitopes based on the IEDB database was carried out.

Kolasker and Tongaonkar antigenicity prediction tool evaluated the proteins for B cell epitopes investigating the physico-chemical properties of amino acids and their affluence in known B cell epitopes. Further, the results revealed that the average antigenicity prospensity value of predicted epitopes of P2 protein was 1.037 with a minimum of 0.873 and a maximum of 1.285, and of P3 protein was 1.048 with a minimum of 0.925 and a maximum of 1.235, whereas of P30 protein was 1.005 with a minimum of 0.889 and maximum of 1.188 (Fig. 16).

The surface accessibility and hydrophilicity regions are also the key features for predicting the B-cell epitopes. The Emini surface accessibility and Parker hydrophilicity prediction tools were used and the results are displayed with graphical interpretation from the Emini surface tool for P2 protein, P3 protein, and P30 protein (Fig. 17), respectively. The results from parker hydrophilicity prediction tool for P2 protein, P3 protein, and P30 protein (Fig. 18), respectively.

The beta turns in a protein are commonly surface accessible and hydrophilic in nature. Chou and Fasman Beta turn prediction was performed for P2, P3 and P30 (Fig. 19) proteins to find beta turn regions in the query protein sequence as beta turns have a substantial effect to induce antigenicity. Generated results were revealed that a region from 345 to 355 for P2 membrane protein, a region from 113 to 115 for P3 membrane protein, and similarly
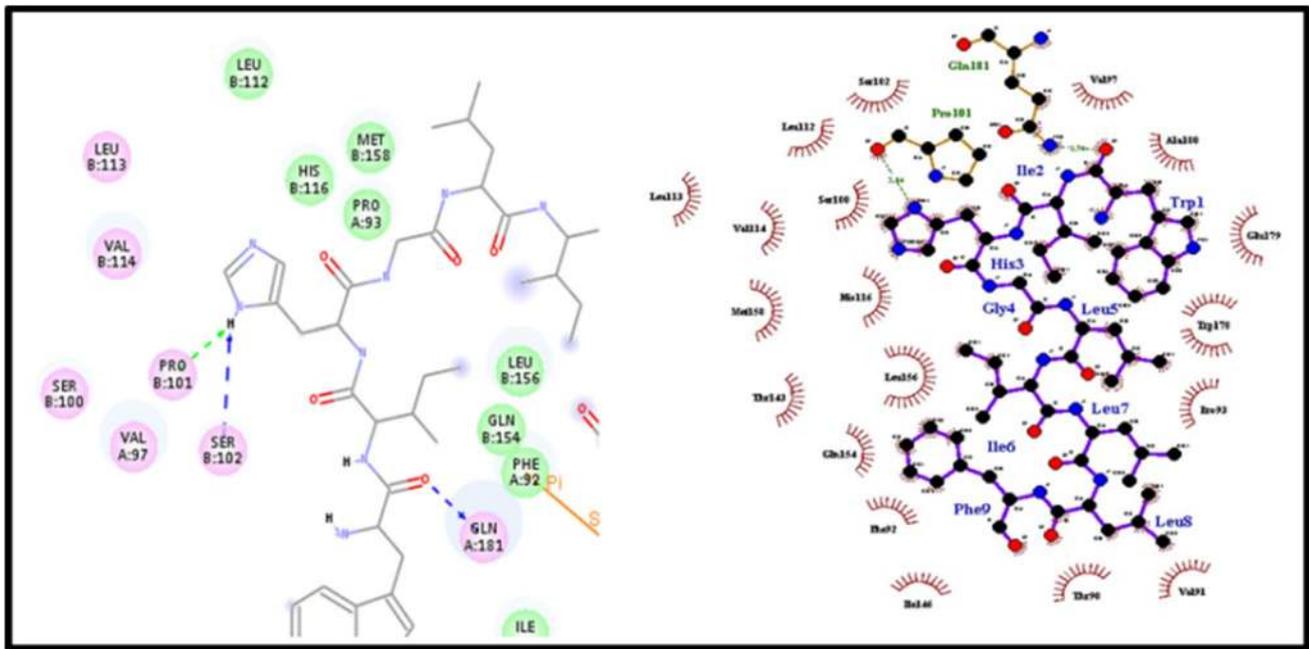
Fig. 5 2D structure representation and hydrophobic interaction of HLA-DP with the peptide WIHGLILLF
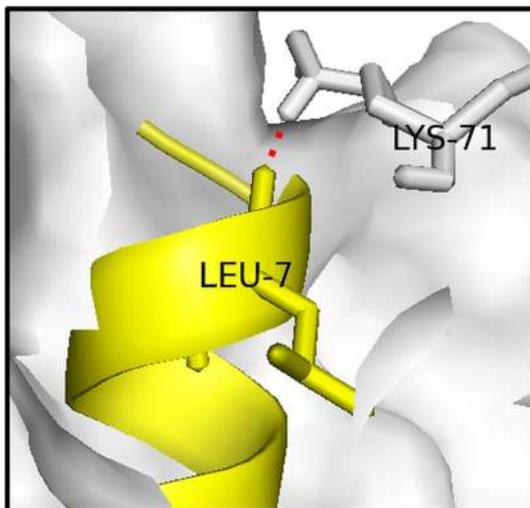
Fig. 6 Docking of peptide (VILLFLLLF) derived from protein (Uniprot ID: Q50327) with HLA DQ

a region from 250 to 254 for P30 adhesin protein with constant predicted B turn region.

Several experimental data revealed that the region of a peptide interaction with the antibody tends to be flexible. Hence, Karplus Schulz flexibility prediction tool was identified the flexible regions on all the query protein. The region from 345 to 350 for P2 membrane protein, region from 116 to 122 for P3 membrane protein, and similarly

122 to 124 for P30 adhesin protein (Fig. 20) is considerably the most favorable region in flexibility prediction analysis.

Biepipred tool is based on the Hidden-Markov model, to determine Linear B cell epitopes. The tool was used to eliminate the fact that single scale amino acid propensity profile cannot consistently predict the antigenic epitopes all the time, and to acquire a better result from epitope prediction tools than the receiver operating characteristics (ROC) plot. The Bepipred predicted epitopes on all the proteins are depicted in Fig. 21.

Subsequently, cross processing all the data obtained from previous B-cell epitope prediction tools, the region from 345 to 355 for P2 membrane protein, 115 to 122 for P3 membrane protein, and 122–139 for P30 adhesin protein are found to be the best capable region to induce B-cell response.

Model building and refinement

The models of all the three proteins (P75588, Q50327 and P75330) were generated by I-TASSER and validated by Ramachandran plot analysis. Further, the modelled proteins were subjected for docking studies with the available therapeutics (antibiotics) used in the treatment of *M. pneumoniae* infection to compare the binding efficiency of nonameric peptides.
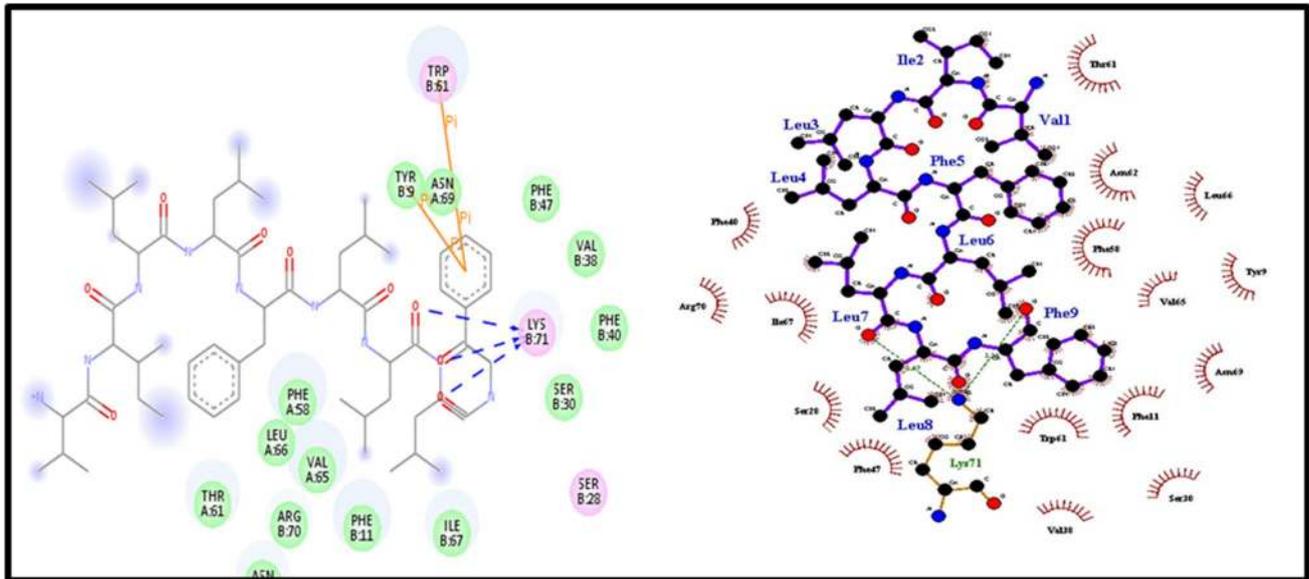
Fig. 7 2D structure representation and hydrophobic interaction of HLA-DQ with the peptide VILLFLLLF
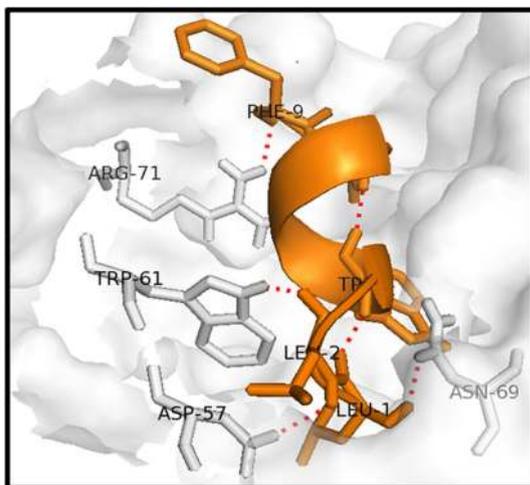
Fig. 8 Docking of peptide (LLAWMLVLF) derived from protein (Uniprot ID: P75330) with HLA DR

Protein–Ligand docking analysis

The ligand information for docking analysis showed as log file in Table 7. The proteins identified as the best binders with MHC II were docked with existing drugs to identify binding affinity displayed in Table 8. Claithromycin, azithromycin, doxycycline, telithromycin and tetracycline were docked against P75588, Q50327 and P75330. The binding energy of the nonameric peptide derived from P75588 was − 915.8 kcal/mol while the conventional drugs which are targeted against the disease inducing outer membrane protein, p75588 showed binding energy of in the range of

− 110.3 kcal/mol to − 127.6 kcal/mol. The nonameric peptide derived from Q50327, ATP synthase subunit b exhibited a binding energy of − 982.1 kcal/mol with HLA DQ, while the drugs showed binding energy in the range of − 101.1 kcal/mol to − 121.7 kcal/mol. Similarly, the epitope derived from p30 adhesin protein showed binding energy of − 893 kcal/mol while the drugs exhibited binding energies in the range of − 105.5 kcal/mol to − 149.8 kcal/mol. From the docking scores, it is evident that existing drugs used for the treatment of infection has least binding affinity to the antigenic outermembrane proteins but the peptide derived from these antigenic proteins exhibits better binding affinity with HLA, which critically confirms that nonameric peptides can be used as effective therapeutic agents for the treatment of the disease caused by *M. pneumoniae*.

Discussions

The prevalence of bacterial infections such as pneumonia, meningitis, tuberculosis and so on, direct towards an urgent requirement of more enhanced and rapidly disbursable vaccines for the same. For these pathogens, immunity related with protection remains largely unknown. Insight into the gaps of protective immunity against the pathogen create vaccine development for the newly emerging infectious diseases more crucial and challenging [47].

The present study aims to screen and investigate the best antigenic proteins of the *M. pneumoniae*, and also to identify the T-cell and B-cell epitopes that were forged on the most antigenic protein by screening of vaccine epitopes. In
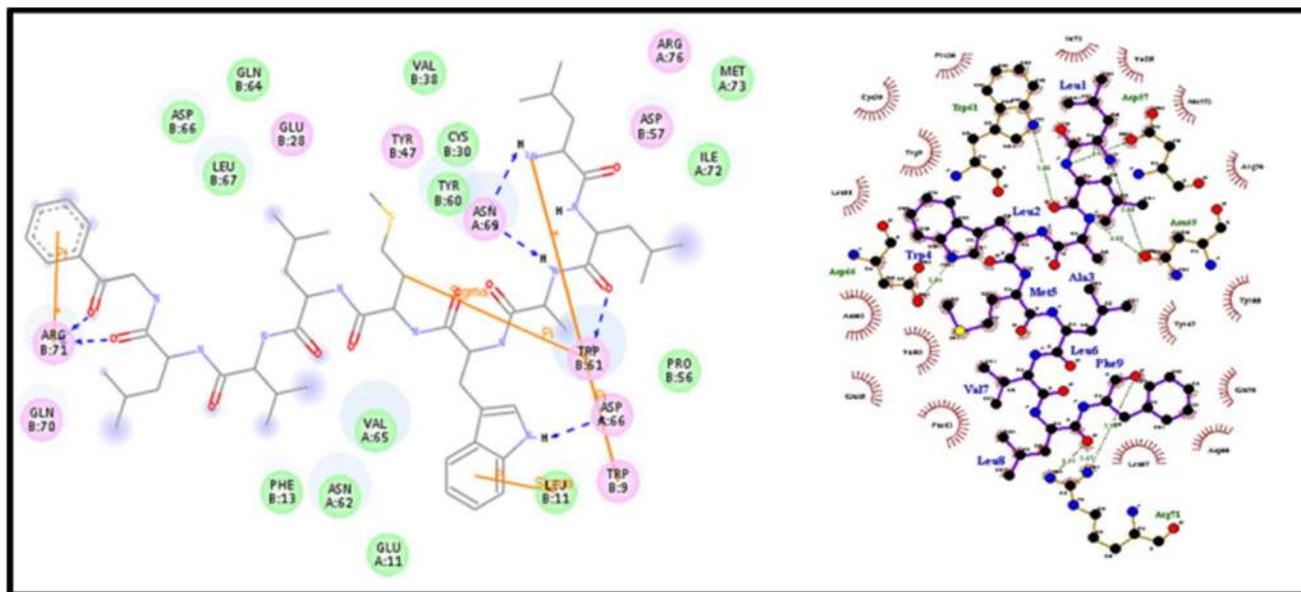
Fig. 9 2D structure representation and hydrophobic interaction of HLA-DR with the peptide LLAWMLVLF
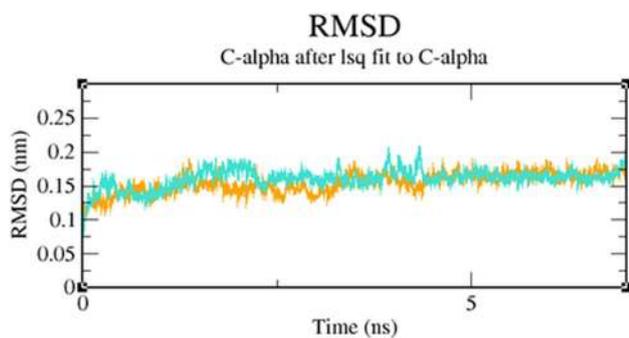
Fig. 10 Analysis of backbone RMSD of HLA-DP protein and protein-peptide complex (epitope and HLA-DP). The symbol coding scheme as follows: HLA-DP (orange color) and protein-peptide complex (turquoise blue color). (Color figure online)

Fig. 12 Analysis of backbone RMSD of HLA-DR protein and protein-peptide complex (epitope and HLA-DR). The symbol coding scheme as follows: HLA-DR (orange color) and protein-peptide complex (turquoise blue color). (Color figure online)
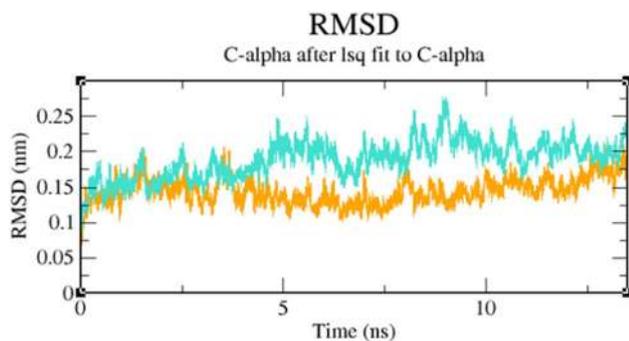
Fig. 11 Analysis of backbone RMSD of HLA-DQ protein and protein-peptide complex (epitope and HLA-DQ). The symbol coding scheme as follows: HLA-DQ (orange color) and protein-peptide complex (turquoise blue color). (Color figure online)
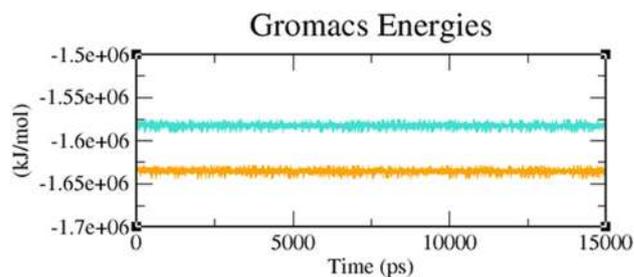
Fig. 13 Total energy of HLA-DP protein and protein-peptide complex (epitope and HLA-DP). The symbol coding scheme as follows: HLA-DP (orange color) and protein-peptide complex (turquoise blue color). (Color figure online)
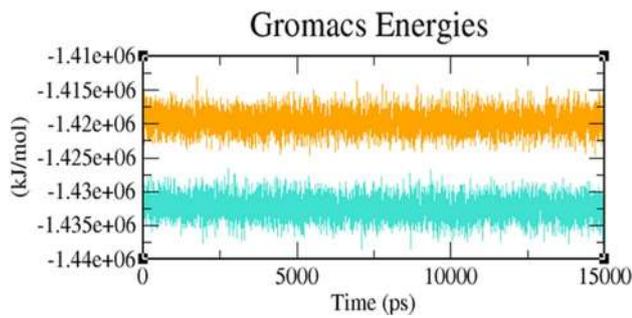
Fig. 14 Total energy of HLA-DQ protein and protein-peptide complex (epitope and HLA-DQ). The symbol coding scheme as follows: HLA-DQ (orange color) and protein-peptide complex (turquoise blue color). (Color figure online)
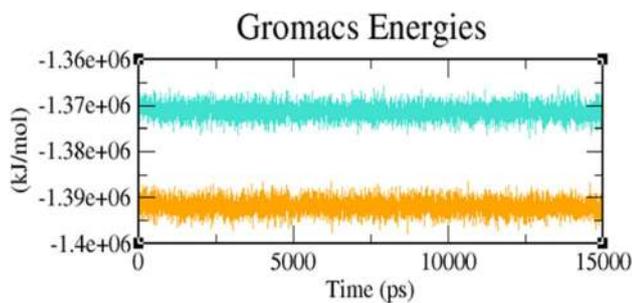
Fig. 15 Total energy of HLA-DR protein and protein-peptide complex (epitope and HLA-DR). The symbol coding scheme as follows: HLA-DR (orange color) and protein-peptide complex (turquoise blue color). (Color figure online)

this study, an immunoinformatics-driven approach was utilized to screen significantly dominant immunogens against *M. pneumoniae*. The results showed that all the membrane associated and cytadherence proteins were better antigenic with highest antigenic scores. Moreover, most of the studies of pneumonia vaccine focused on the membrane associated and cytadherence proteins of *M. pneumoniae* [29, 48, 49], because of recognizing the attachment and stimulation of local damage at the cellular level are responsible for *M. pneumoniae* disease. Thus, the logical method is to avert the attachment and thereby prevent initiation of disease. Further, T-cell-based cellular immunity is important for eliminating *M. pneumoniae* infection, yet the vaccine against the membrane associated and cytadherence proteins mainly elicit neutralizing antibody response. Notably, high mutation rate of the membrane associated and cytadherence proteins of *M. pneumoniae* may lead to escape of neutralizing antibodies against the same proteins. Thus, an ideal target should be much conserved that elicit both neutralizing cellular immunity and antibody against *M. pneumoniae*, which is more

significant for an effective pneumonia vaccine development. The membrane associated and cytadherence proteins of *M. pneumoniae* are abundantly produced during the disease progression and exhibits strong conservancy and immunogenicity, which can act as a best immunogen to elicit both humoral and cell-mediated immune responses [29, 50].

A T-cell epitope is considered as strong and effective, if it is fine conserved among the selected proteins. To validate the selected epitopes WIHGLILLF, VILLFLLLF and LLAWMLVLF based on the docking score, which predicts nonameric peptide sequence binds to class II MHC.WIHGLILLF and VILLFLLLF interact with 17 MHC class II molecules out of 23 alleles, while LLAWMLVLF interacts with 16 MHC class II molecules out of 23 alleles. Investigating similar data from the other study revealed that this specific high binding affinity is completely desired because the epitope vaccine efficiency significantly relies on the specific interaction between epitope and HLA alleles [51]. Simultaneously, the selected epitopes are declared as non-allergen, an undeniable characteristics a vaccine must have. These results imply that the vaccine would be potential for the management of disease.

The epitopes were docked to evaluate the binding efficacy of all the epitopes with MHC class II alleles. The results revealed that the WIHGLILLF epitope highest binding affinity with specific HLA allele HLA-DP, VILLFLLLF epitope highest binding affinity with a specific HLA allele HLA-DQ, and LLAWMLVLF highest binding affinity with specific HLA allele HLA-DR. This computational analysis confirms the binding affinity of epitopes for the MHC-II molecules and upholds.

The best selected membrane associated proteins and cytadherence proteins were also searched for B cell epitopes as the B cell epitopes can activate both the primary and secondary immunity. Various tools from IEDB database generated results examining the protein based on the vital characteristics of B cell epitopes. Further, we cross referenced all data and the regions from 345 to 355 for P2 membrane protein, 115 to 120 for P3 membrane protein, and the regions form 135 to 139 for P30 adhesin protein, were found to be predicted as putative or B cell epitope by Bepired tool. The regions consist of flexible and beta turns. The region is hydrophilic and surface accessibility comparatively than other regions and evidenced to be antigenic. The 9-mer epitope WIHGLILLF from P2 adhesin protein, VILLFLLLF from P3 adhesin protein, and LLAWMLVLF from P30 adhesin protein are the most favorable as B cell epitopes according to the predicted results.

As suggested by VaXijen and Allerhunter all the proteins were found to be eligible targets for the vaccine development because of highest antigenicity and least/null allergenicity. The proteins from which 3 potent T cell epitopes WIHGLILLF, VILLFLLLF and LLAWMLVLF
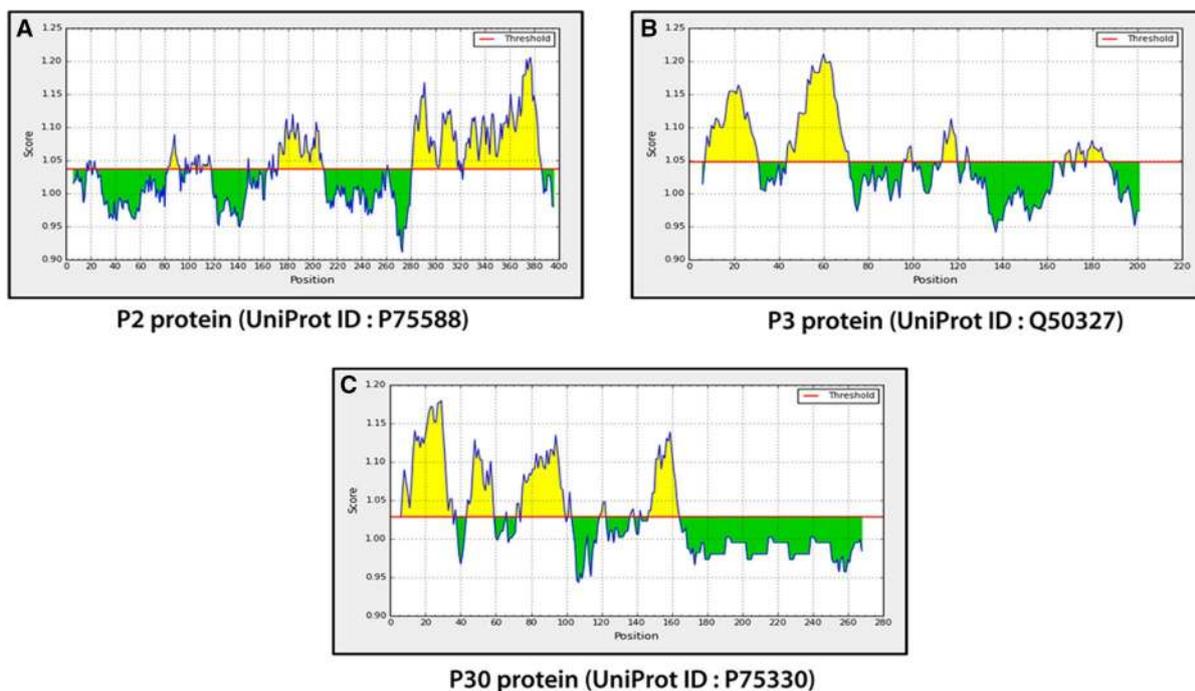
Fig. 16 Kolaskar and Tongaonkar antigenicity prediction of the best antigenic proteins. **a** P2 protein (Uniprot ID: P75588); The X-axis and Y-axis denote the sequence position and antigenic propensity, respectively. The threshold is 1.037. The regions above the threshold value are antigenic, shown in yellow. **b** P3 protein (Uniprot ID: Q50327); The X-axis and Y-axis denote the sequence position and antigenic propensity, respectively. The threshold is 1.048. The regions above the threshold value are antigenic, shown in yellow. **c** P30 protein (Uniprot ID: P75330); The X-axis and Y-axis denote the sequence position and antigenic propensity, respectively. The threshold is 1.029. The regions above the threshold value are antigenic, shown in yellow. (Color figure online)
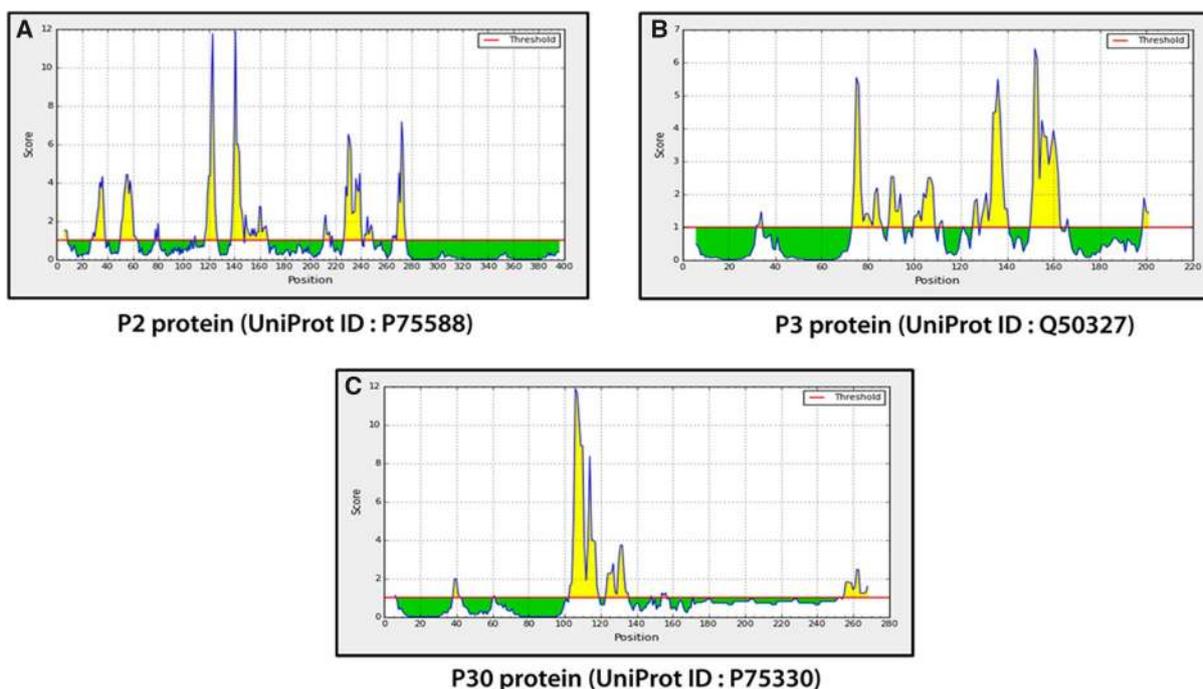
Fig. 17 Emini surface accessibility prediction of the best antigenic proteins. **a** P2 protein (Uniprot ID: P75588), **b** P3 protein (Uniprot ID: Q50327), **c** P30 protein (Uniprot ID: P75330); The X-axis and Y-axis denote the sequence position and surface probability, respectively. The threshold is 1.0. The regions above the threshold value are antigenic, shown in yellow. (Color figure online)

P2 protein (UniProt ID : P75588)

P3 protein (UniProt ID : Q50327)

P30 protein (UniProt ID : P75330)
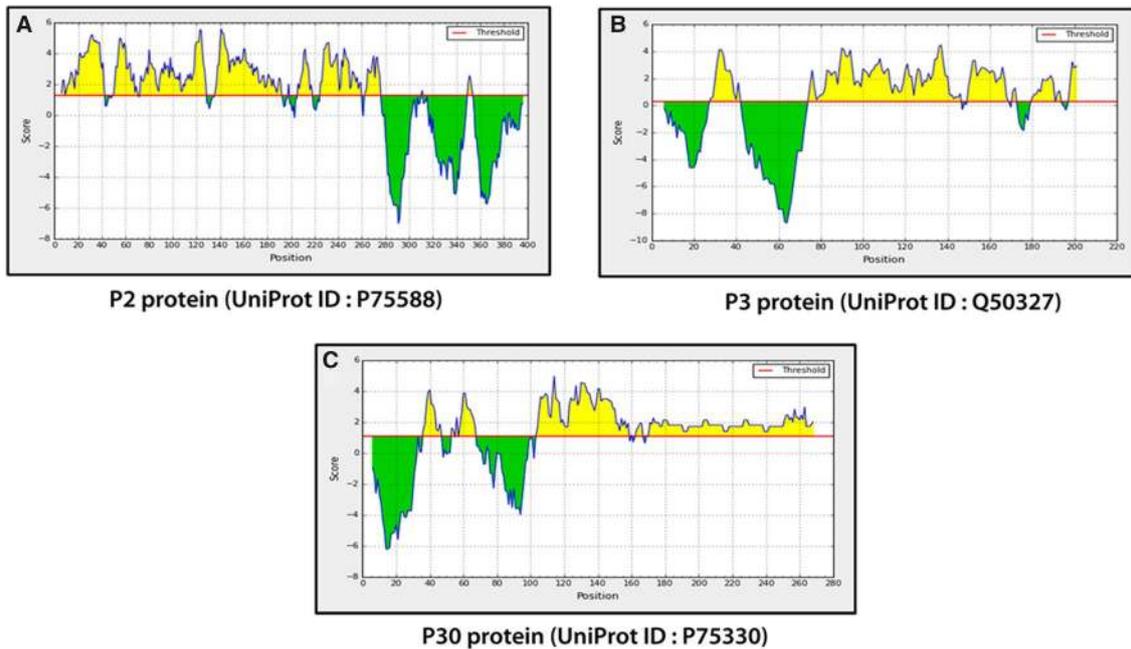
Fig. 18 Parker hydrophilicity prediction of the best antigenic proteins. **a** P2 protein (Uniprot ID: P75588); The X-axis and Y-axis denote the position and score, respectively. The threshold is 1.297. The regions above the threshold value, having beta turns in the protein which shown in yellow color. **b** P3 protein (Uniprot ID: Q50327); The X-axis and Y-axis denote the position and score, respectively. The threshold is 0.317. The regions above the threshold value, having beta turns in the protein which shown in yellow color. **c** P30 protein (Uniprot ID: P75330); The X-axis and Y-axis denote the position and score, respectively. The threshold is 1.121. The regions above the threshold value, having beta turns in the protein which shown in yellow color. (Color figure online)

P2 protein (UniProt ID : P75588)

P3 protein (UniProt ID : Q50327)

P30 protein (UniProt ID : P75330)
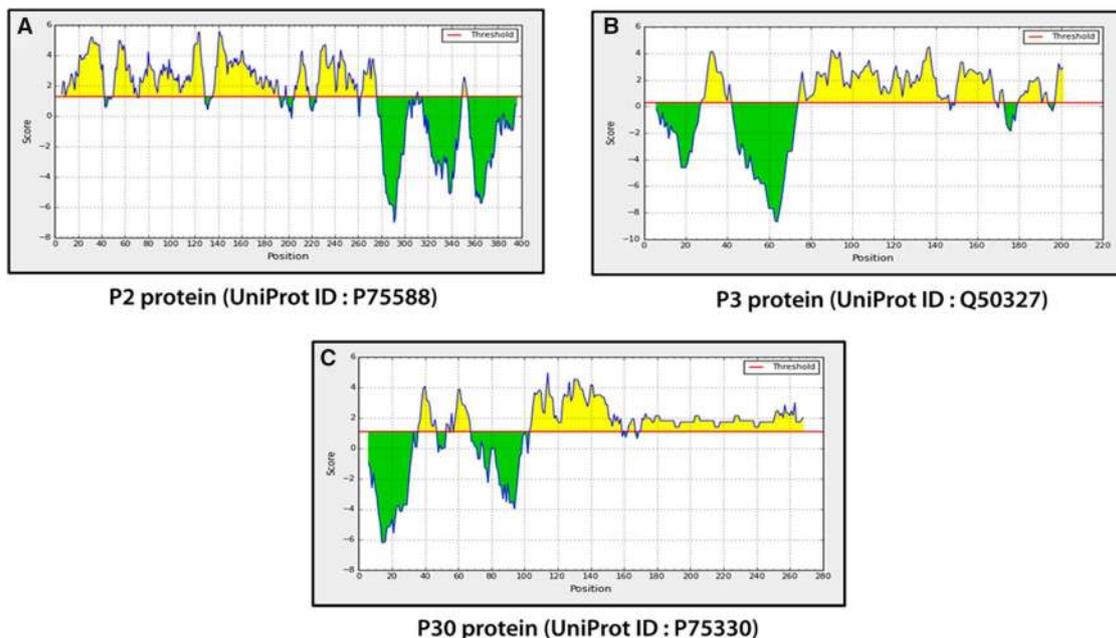
Fig. 19 Chou and Fasman beta-turn prediction of the best antigenic proteins. **a** P2 protein (Uniprot ID: P75588); The X-axis and Y-axis denote the position and score, respectively. The threshold is 0.984. The regions above the threshold value, having beta turns in the protein which shown in yellow color. **b** P3 protein (Uniprot ID: Q50327); The X-axis and Y-axis denote the position and score, respectively. The threshold is 0.826. The regions above the threshold value, having beta turns in the protein which shown in yellow color. **c** P30 protein (Uniprot ID: P75330); The X-axis and Y-axis denote the position and score, respectively. The threshold is 0.740. The regions above the threshold value, having beta turns in the protein which shown in yellow color. (Color figure online)
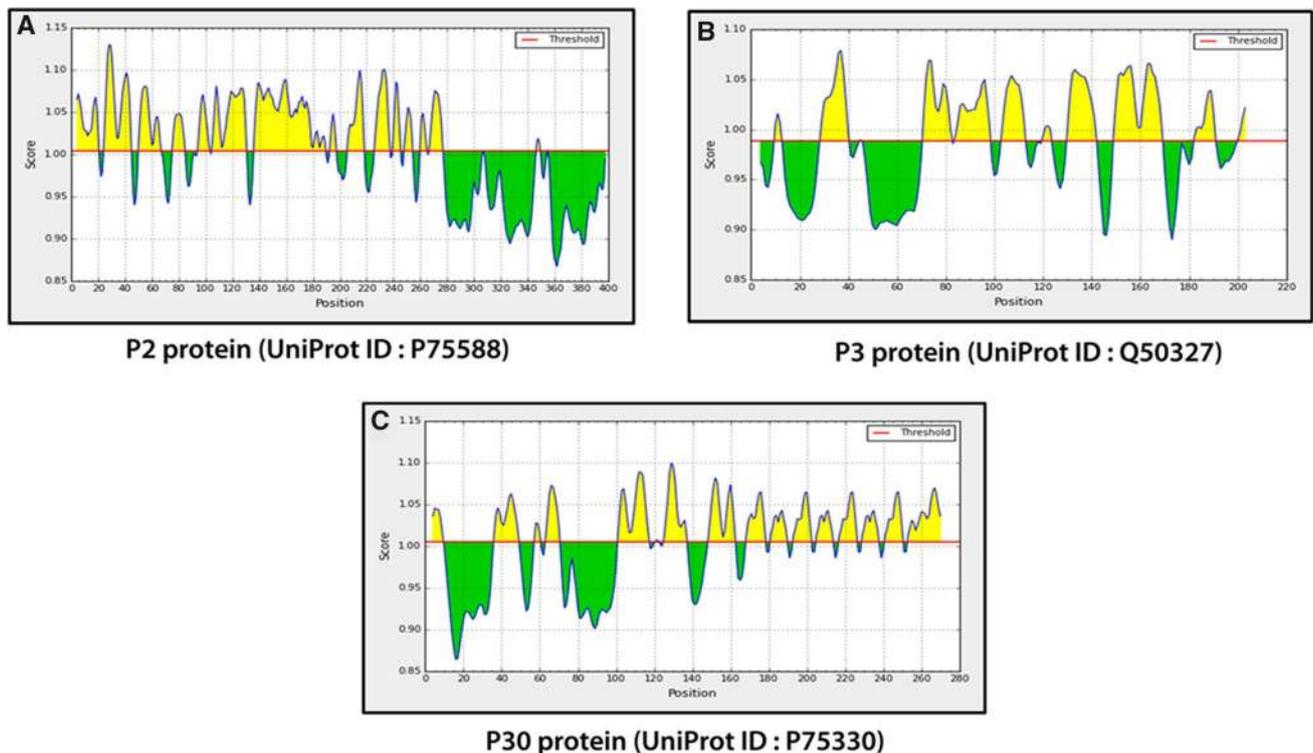
Fig. 20 Karplus and Schulz flexibility prediction of the best antigenic proteins. **a** P2 protein (Uniprot ID: P75588); The X-axis and Y-axis denote the position and score, respectively. The threshold is 1.004. The regions above the threshold value, flexible regions of the protein shown in yellow color. **b** P3 protein (Uniprot ID: Q50327); The X-axis and Y-axis denote the position and score, respectively. The threshold is 0.989. The regions above the threshold value, flexible regions of the protein shown in yellow color. **c** P30 protein (Uniprot ID: P75330); The X-axis and Y-axis denote the position and score, respectively. The threshold is 1.005. The regions above the threshold value, flexible regions of the protein shown in yellow color. (Color figure online)

were derived showed antigenic score of 0.6274, 0.6902 and 0.7167 respectively and all these proteins were classified as non-allergens according to the prediction of Allerhunter. A valuable insight was achieved from secondary structure analysis of proteins from SOPMA. The abundance of alpha helix and coiled region in SOPMA generated results implies greater stability and conservancy of proteins.

The three dimensional structure for the epitopes were predicted and validated to analyze the stereochemical quality of predicted models. The RAMPAGE analysis for the potent T cell epitopes showed the percentage of residues in the favorable region. Thus, all the T cell epitopes generated from the membrane and cytadherence proteins which showed percentage of residues in the favoured region greater than 75% in RAMPAGE analysis were subjected to docking with HLA DP, DQ and DR by CLUSPRO. Epitopes WIHGLILLF, VILLFLLLF and LLAWMLVLF showed binding energy values of -915.8 kcal/mol,-982.1 kcal/mol and-893 kcal/mol with HLA DP, DQ and DR respectively. Hence the significant

interaction between HLA and predicted T cell epitopes were confirmed. Proteins from the most potent T cell epitopes were generated were subjected to docking with the existing drugs which belongs to macrolide classes and found a lesser binding affinity with HLA which further confirms that the therapeutic potential of novel T cell epitopes derived from the membrane associated proteins of *M. pneumoniae*.

MD simulation studies were performed to analyze the stability of bound protein-peptide complex. Higher energy was observed for the HLA DP and HLA DR protein-peptide complexes with respect to the native protein, which implies a stable interaction between bound HLA and peptide complex.

Conclusion

In conclusion, the present study pointed out the application of immunoinformatics to predict epitope binding HLA alleles of pathogens as a potential strategy to hasten
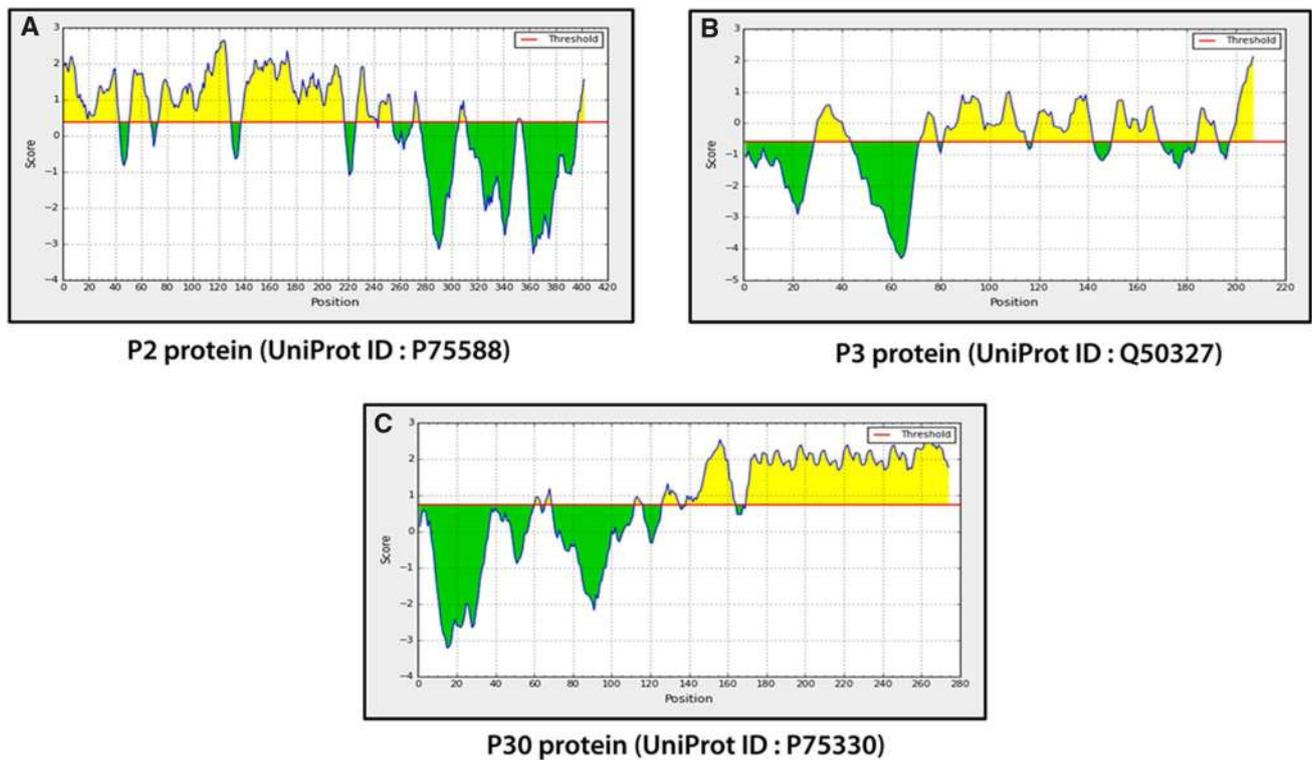
Fig. 21 Bepipred linear epitope prediction of the best antigenic proteins. **a** P2 protein (Uniprot ID: P75588); The X-axis and Y-axis denote the position and score, respectively. The threshold is 0.376. The regions having beta turns are shown in yellow color. The highest peak region depicts the most potent B-cell epitope. **b** P3 protein (Uniprot ID: Q50327); The X-axis and Y-axis denote the position and score, respectively. The threshold is 0.578. The regions having beta turns are shown in yellow color. The highest peak region depicts the most potent B-cell epitope. **c** P30 protein (Uniprot ID: P75330); The X-axis and Y-axis denote the position and score, respectively. The threshold is 0.740. The regions having beta turns are shown in yellow color. The highest peak region depicts the most potent B-cell epitope

Table 7 Ligand information for docking analysis and protein–ligand interaction

S. no	Drug	Molecular formula	Molecular weight	Function
1	Azithromycin	$C_{38}H_{72}N_2O_{12}$	748.9845	Inhibits RNA dependent protein synthesis in bacterial cells
2	Claithromycin	$C_{38}H_{69}NO_{13}$	747.9534	Blocks translocation of amino-acyl transfer RNA and polypeptide synthesis
3	Doxycycline	$C_{22}H_{24}N_2O_8$	444.4346	Blocks the binding of aminoacyl tRNA to the mRNA and inhibits bacterial protein synthesis
4	Telithromycin	$C_{43}H_{65}N_5O_{10}$	812.018	Binds with 50S ribosomal subunit
5	Roxithromycin	$C_{41}H_{76}N_2O_{15}$	837.0465	Roxithromycin binds to the subunit 50S of the bacterial ribosome, and thus inhibits the translocation of peptides
6	Tetracycline	$C_{22}H_{24}N_2O_8$	444.4346	Binds to the bacterial 30S ribosomal subunit and blocks incoming aminoacyl tRNA from binding to the ribosome acceptor site

vaccine development. Based on this approach, the T cell epitopes and B cell epitopes in the membrane associated and cytadherence proteins of *M. pneumoniae* were screened out and studied to be used as promising vaccine candidates. However, the T cell and B cell stimulation prospective of the selected epitopes need to be tested using extensive laboratory based techniques along with this computational study for their effective practice as vaccines against *M. pneumoniae*. Nonetheless, these findings not only provide novel and valuable epitope candidates, but also prompt the way to develop the vaccines against Mycoplasmal Pneumonia and its associated clinical manifestations.

Table 8 Docking score for protein–ligand interaction

S. no	Protein	Drug	Docking score
1	P2 (P75588)	Claithromycin	− 101.1
		Azithromycin	− 106.6
		Doxycycline	− 121.7
		Telithromycin	− 114
		Tetracycline	− 117.7
2	P3 (Q50327)	Doxycycline	− 127.6
		Tetracycline	− 111.1
		Telithromycin	− 110.4
		Azithromycin	− 110.3
		Claithromycin	− 103.4
3	P30 (P75330)	Doxycycline	− 141.2
		Azithromycin	− 131.3
		Clarithromycin	− 105.5
		Telithromycin	− 123.1
		Tetracycline	− 149.8

References

1. Atkinson TP, Balish MF, Waites KB (2008) Epidemiology, clinical manifestations, pathogenesis and laboratory detection of *Mycoplasma pneumoniae* infections. FEMS Microbiol Rev 32(6):956–973. https://doi.org/10.1111/j.1574-6976.2008.00129.x
2. Liu G, Talkington DF, Fields BS, Levine OS, Yang Y, Tondella MLC (2005) Chlamydia pneumoniae and *Mycoplasma pneumoniae* in young children from China with community-acquired pneumonia. Diagn Microbiol Infect Dis 52(1):7–14. https://doi.org/10.1016/j.diagmicrobio.2005.01.005
3. Vervloet LA, Marguet C, Camargos PAM (2007) Infection by *Mycoplasma pneumoniae* and its importance as an etiological agent in childhood community-acquired pneumonias. Braz J Infect Dis 11(5):507–514
4. Waites KB, Balish MF, Atkinson TP (2008) New insights into the pathogenesis and detection of *Mycoplasma pneumoniae* infections. Future Microbiol 3(6):635–648. https://doi.org/10.2217/17460913.3.6.635
5. Razin S, Yogev D, Naot Y (1998) Molecular Biology and Pathogenicity of Mycoplasmas. Microbiol Mol Biol Rev 62(4):1094–1156
6. Waites KB, Talkington DF (2004) *Mycoplasma pneumoniae* and its role as a human pathogen. Clin Microbiol Rev 17(4):697–728. https://doi.org/10.1128/CMR.17.4.697-728.2004 table of contents.
7. Baernstein HD, Trevisani E, Axtell S, Quilligan JJ (1965) *Mycoplasma pneumoniae* (Eaton atypicalpneumonia agent) in children's respiratory infections. J Pediatr 66(5):829–837. https://doi.org/10.1016/S0022-3476(65)80057-9
8. Wilson MH, Collier AM (1976) Ultrastructural study of *Mycoplasma pneumoniae* in organ culture. J Bacteriol 125(1):332–339
9. Fonsêca MM, da Zaha A, Caffarena ER, Vasconcelos ATR (2012) Structure-based functional inference of hypothetical proteins from Mycoplasma hyopneumoniae. J Mol Model 18(5):1917–1925. https://doi.org/10.1007/s00894-011-1212-3
10. Himmelreich R, Hilbert H, Plagens H, Pirkl E, Li BC, Herrmann R (1996) Complete sequence analysis of the genome of the bacterium *Mycoplasma pneumoniae*. Nucleic Acids Res 24(22):4420–4449
11. Dorigo-Zetsma JW, Wilbrink B, van der Nat H, Bartelds AI, Heijnen ML, Dankert J (2001) Results of molecular detection of *Mycoplasma pneumoniae* among patients with acute respiratory infection and in their household contacts reveals children as human reservoirs. J Infect Dis 183(4):675–678. https://doi.org/10.1086/318529
12. Krivan HC, Olson LD, Barile MF, Ginsburg V, Roberts DD (1989) Adhesion of *Mycoplasma pneumoniae* to sulfated glycolipids and inhibition by dextran sulfate. J Biol Chem 264(16):9283–9288
13. Roberts DD, Olson LD, Barile MF, Ginsburg V, Krivan HC (1989) Sialic acid-dependent adhesion of *Mycoplasma pneumoniae* to purified glycoproteins. J Biol Chem 264(16):9289–9293
14. Krause DC (1996) *Mycoplasma pneumoniae* cytadherence: unravelling the tie that binds. Mol Microbiol 20(2):247–253
15. Krause DC, Leith DK, Wilson RM, Baseman JB (1982) Identification of *Mycoplasma pneumoniae* proteins associated with hemadsorption and virulence. Infect Immun 35(3):809–817
16. Bradley JS, Byington CL, Shah SS, Alverson B, Carter ER, Harrison C et al. (2011). The management of community-acquired pneumonia in infants and children older than 3 months of age: clinical practice guidelines by the Pediatric Infectious Diseases Society and the Infectious Diseases Society of America. Clin Infect Dis 53(7), e25–76. https://doi.org/10.1093/cid/cir531
17. Okada T, Morozumi M, Tajima T, Hasegawa M, Sakata H, Ohnari S et al (2012) Rapid effectiveness of minocycline or doxycycline against macrolide-resistant *Mycoplasma pneumoniae* infection in a 2011 outbreak among Japanese children. Clin Infect Dis 55(12):1642–1649. https://doi.org/10.1093/cid/cis784
18. Eshaghi A, Memari N, Tang P, Olsha R, Farrell DJ, Low DE et al (2013) Macrolide-resistant *Mycoplasma pneumoniae* in humans, Ontario, Canada, 2010–2011. Emerg Infect Dis, 19(9). https://doi.org/10.3201/eid1909.121466
19. Peuchant O, Ménard A, Renaudin H, Morozumi M, Ubukata K, Bébéar CM, Pereyre S (2009) Increased macrolide resistance of *Mycoplasma pneumoniae* in France directly detected in clinical specimens by real-time PCR and melting curve analysis. J Antimicrob Chemother 64(1):52–58. https://doi.org/10.1093/jac/dkp160
20. Linchevski I, Klement E, Klmenet E, Nir-Paz R (2009) *Mycoplasma pneumoniae* vaccine protective efficacy and adverse reactions–systematic review and meta-analysis. Vaccine 27(18):2437–2446. https://doi.org/10.1016/j.vaccine.2009.01.135
21. Yang X, Yu X (2009) An introduction to epitope prediction methods and software. Rev Med Virol 19(2):77–96. https://doi.org/10.1002/rmv.602
22. Farhadi T, Nezafat N, Ghasemi Y, Karimi Z, Hemmati S, Erfani N (2015) Designing of complex multi-epitope peptide vaccine based on omps of *Klebsiella pneumoniae*: an in silico approach. Int J Pept Res Ther 21(3):325–341. https://doi.org/10.1007/s10989-015-9461-0
23. Oyarzún P, Kobe B (2016) Recombinant and epitope-based vaccines on the road to the market and implications for vaccine design and production. Hum Vaccines Immunother 12(3):763–767. https://doi.org/10.1080/21645515.2015.1094595

24. Bijker MS, Melief CJM, Offringa R, van der Burg SH (2007) Design and development of synthetic peptide vaccines: past, present and future. Expert Rev Vaccines 6(4):591–603. https://doi.org/10.1586/14760584.6.4.591

25. Baloria U, Akhoon BA, Gupta SK, Sharma S, Verma V (2012) In silico proteomic characterization of human epidermal growth factor receptor 2 (HER-2) for the mapping of high affinity antigenic determinants against breast cancer. Amino Acids 42(4):1349–1360. https://doi.org/10.1007/s00726-010-0830-x

26. Akhoon BA, Slathia PS, Sharma P, Gupta SK, Verma V (2011) In silico identification of novel protective VSG antigens expressed by Trypanosoma brucei and an effort for designing a highly immunogenic DNA vaccine using IL-12 as adjuvant. Microb Pathog 51(1–2):77–87. https://doi.org/10.1016/j.micpath.2011.01.011

27. Gupta SK, Smita S, Sarangi AN, Srivastava M, Akhoon BA, Rahman Q, Gupta SK (2010) In silico CD4 + T-cell epitope prediction and HLA distribution analysis for the potential proteins of Neisseria meningitidis Serogroup B–a clue for vaccine development. Vaccine 28(43):7092–7097. https://doi.org/10.1016/j.vaccine.2010.08.005

28. Gupta SK, Srivastava M, Akhoon BA, Gupta SK, Grabe N (2012) In silico accelerated identification of structurally conserved CD8 + and CD4 + T-cell epitopes in high-risk HPV types. Infect Genet Evol 12(7):1513–1518. https://doi.org/10.1016/j.meegid.2012.02.022

29. Kumar GS, Sarita S, Kumar GM, K, P. K., & K, S. P (2010) Definition of potential targets in *Mycoplasma pneumoniae* through subtractive genome analysis. J Antivir Antiretrovir 2(2). https://doi.org/10.4172/jaa.1000020

30. Doytchinova IA, Flower DR (2007) VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinform 8:4. https://doi.org/10.1186/1471-2105-8-4

31. Muh HC, Tong JC, Tammi MT (2009) AllerHunter: A SVM-pairwise system for assessment of allergenicity and allergic cross-reactivity in proteins. PLoS ONE, 4(6). https://doi.org/10.1371/journal.pone.0005861

32. Geourjon C, Deléage G (1995) SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. Comput Appl Biosci 11(6):681–684

33. Atanasova M, Patronov A, Dimitrov I, Flower DR, Doytchinova I (2013) EpiDOCK: a molecular docking-based tool for MHC class II binding prediction. Protein Eng Des Sel 26(10):631–634. https://doi.org/10.1093/protein/gzt018

34. Thévenet P, Shen Y, Maupetit J, Guyon F, Derreumaux P, Tufféry P (2012) PEP-FOLD: an updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides. Nucleic Acids Res 40:W288–W293. https://doi.org/10.1093/nar/gks419 (Web Server issue)

35. Comeau SR, Gatchell DW, Vajda S, Camacho CJ (2004) ClusPro: a fully automated algorithm for protein-protein docking. Nucleic Acids Res 32:W96–W99. https://doi.org/10.1093/nar/gkh354 (Web Server issue)

36. Berendsen HJC, van der Spoel D, van Drunen R (1995) GROMACS: a message-passing parallel molecular dynamics implementation. Comput Phys Commun 91(1–3):43–56. https://doi.org/10.1016/0010-4655(95)00042-E

37. Jorgensen WL, Maxwell DS, Tirado-Rives J (1996) Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. J Am Chem Soc 118(45):11225–11236. https://doi.org/10.1021/ja9621760

38. Kolaskar AS, Tongaonkar PC (1990) A semi-empirical method for prediction of antigenic determinants on protein antigens. FEBS Lett 276(1–2):172–174

39. Emini EA, Hughes JV, Perlow DS, Boger J (1985) Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. J Virol 55(3):836–839

40. Parker JM, Guo D, Hodges RS (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(19):5425–5432

41. Vihinen M, Torkkila E, Riikonen P (1994) Accuracy of protein flexibility predictions. Proteins 19(2):141–149. https://doi.org/10.1002/prot.340190207

42. Haste Andersen P, Nielsen M, Lund O (2006) Prediction of residues in discontinuous B-cell epitopes using protein 3D structures. Protein Sci 15(11):2558–2567. https://doi.org/10.1110/ps.062405906

43. Rini JM, Schulze-Gahmen U, Wilson IA (1992) Structural evidence for induced fit as a mechanism for antibody-antigen recognition. Science 255(5047):959–965

44. Chou PY, Fasman GD (1978) Prediction of the secondary structure of proteins from their amino acid sequence. Adv Enzymol Relat Areas Mol Biol 47:45–148

45. Roy A, Kucukural A, Zhang Y (2010) I-TASSER: a unified platform for automated protein structure and function prediction. Nat Protoc 5(4):725–738. https://doi.org/10.1038/nprot.2010.5

46. Seeliger D, de Groot BL (2010) Ligand docking and binding site analysis with PyMOL and Autodock/Vina. J Comput Aided Mol Des 24(5):417–422. https://doi.org/10.1007/s10822-010-9352-6

47. Terry FE, Moise L, Martin RF, Torres M, Pilotte N, Williams SA, De Groot AS (2015) Time for T? immunoinformatics addresses vaccine design for neglected tropical and emerging infectious diseases. Expert Rev Vaccines 14(1):21–35. https://doi.org/10.1586/14760584.2015.955478

48. Talkington DF, Schwartz SB, Besser RE, Waites KB (2001) Emerging from obscurity: understanding pulmonary and extrapulmonary syndromes, pathogenesis, and epidemiology of human *Mycoplasma pneumoniae* infections, 57–84. https://doi.org/10.1128/9781555816988.ch4

49. Chaudhry R, Ghosh A, Chandolia A (2016) Pathogenesis of *Mycoplasma pneumoniae*: an update. Indian Journal of Medical Microbiology 34(1):7. https://doi.org/10.4103/0255-0857.174112

50. Zimmerman RK, Santibanez TA, Fine MJ, Janosky JE, Nowalk MP, Bardella IJ et al (2003) Barriers and facilitators of pneumococcal vaccination among the elderly. Vaccine 21(13–14):1510–1517

51. Chakraborty S, Chakravorty R, Ahmed M, Rahman A, Waise TMZ, Hassan F et al (2010) A computational approach for identification of epitopes in dengue virus envelope protein: a step towards designing a universal dengue vaccine targeting endemic regions. In Silico Biol 10(5–6):235–246. https://doi.org/10.3233/ISB-2010-0435