

Research article

The effect of in silico targeting *Pseudomonas aeruginosa* patatin-like protein D, for immunogenic administration

Alireza Salimi Chirani^{a,*}, Robabeh Majidzadeh^a, Ramin Pouriran^b, Mohsen Heidary^c,
 Mohammad Javad Nasiri^a, Mehrdad Gholami^c, Mehdi Goudarzi^a, Vahid Fallah Omrani^{d,e}

^a Department of Microbiology, School of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran

^b School of medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran

^c Department of Microbiology, School of Medicine, Iran University of Medical Sciences, Tehran, Iran

^d Department of Biotechnology, School of Advanced Technologies in Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran

^e Student Research Committee, Shahid Beheshti University of Medical Sciences, Tehran, Iran

ARTICLE INFO

Article history:

Received 4 March 2017

Received in revised form 31 January 2018

Accepted 3 February 2018

Available online 5 February 2018

Keyword:

Patatin-like protein D

Pseudomonas aeruginosa

Type Vd secretion system

Immunoinformatic

Immunogenic

ABSTRACT

The vaccine candidates that have been introduced for immunization against *Pseudomonas aeruginosa* (*P. aeruginosa*) strains are quite diverse. In fact, there has been no proper antigen to act as an effective immunogenic substance against this ubiquitous pathogen in the market as yet. The complications caused by this bacterium due to the rapid development of multiple drug resistant strains have led to clinical problems worldwide. *P. aeruginosa* encodes many specific virulence elements that could be used as appropriate vaccine candidates. Type Vd secretion system, also known as patatin-like protein D, is a novel *P. aeruginosa* auto-transporter system. It is known that cellular or humoral immune responses could be elevated by chimeric proteins carrying epitopes. It has been recognized that in silico tools are essential for the evaluation of new chimeric antigens.

In this study, we have considered the patatin-like protein D (PlpD) molecule from *P. aeruginosa* and predicted some immunogenic properties of this strong cytotoxic phospholipase A2 with the use of in-depth computational and immunoinformatics assessment methods. The novelty of our in silico study is the modeling and assessment of both humoral and cellular immune potential against the PlpD molecule. The molecule was considered by multiple sequence alignment and homology valuation. The extremely conserved regions in the PlpD were predicted. The allergenic and physicochemical property predictions on the PlpD state that the molecule is a non-allergic and stable molecule. High-resolution secondary and tertiary conformations were created. Indeed, the B-cell and T-cell epitope mapping on the chimeric target protein confirmed that the engineered protein contained a tremendous number of both B-cell and T-cell corresponding epitopes. This investigation magnificently attained the chimeric molecule as being a potent lipolytic enzyme composed of numerous B-cell and T-cell restricted epitopes and could induce both humoral and cellular immune responses. The results indicated that this molecule has therapeutic potential against several potent pathogenic *P. aeruginosa* strains.

© 2018 Published by Elsevier Ltd.

1. Introduction

Pseudomonas aeruginosa represents an opportunistic human pathogen in hospitals. This pathogen is the causative agent for severe, acute, and chronic infections such as highly regarded ventilator-associated pneumonia and chronic obstructive pulmonary disease (COPD) (Gellatly and Hancock, 2013). Even, with the improvement in antimicrobial drugs, the morbidity and mortality

rates for multidrug-resistant (MDR) *Pseudomonas aeruginosa* (*P. aeruginosa*) strains have been increasing alarmingly (Waters and Smyth, 2015; Vitkauskienė et al., 2010). The pathogenesis of *P. aeruginosa* strictly rests on cell associated (type IV pili, flagellum, secretion systems, lipopolysaccharides, and alginate) and secreted (rhamnolipids, pyocyanin, proteases, exotoxins, and phospholipases) virulence factors. The secretion systems in *P. aeruginosa* that are quite vital virulence factors assist the pathogen to survive within the hosts to manipulate the cell response and escape the immune system. Recently, a novel bi-functional type V (named Vd or patatin-like protein (PlpD) auto-transporter system) secretion system (Salacha et al., 2010) was discovered in *P. aeruginosa*. The

* Corresponding author.

E-mail address: Alireza.salimichirani@sbmu.ac.ir (A.S. Chirani).

PlpD molecule works as a lipolytic enzyme and serves as an invasive compound against host cells (Gellatly and Hancock, 2013; Ben Haj Khalifa et al., 2011). Advantages in computational immunology represent a great role for formulating new hypotheses related to vaccination. The combination of computer science and experimental immunology helps scientists to deal with enormous amounts of information (Tomar and De, 2010; Tomar and De, 2014). In this study, by using in-depth immunoinformatics analysis and in silico modeling, we confirmed that the *P. aeruginosa* PlpD apparatus could act as a humoral immune system activator against several strains of *P. aeruginosa*. To the best of our knowledge, this report is the first in silico study and immunoinformatics analysis of the *P. aeruginosa* type Vd secretion system as a potent phospholipase A2 cytotoxic molecule.

2. Materials and methods

2.1. Sequences and databases

The nucleotide sequence of PlpD from *P. aeruginosa* PAO1 with the accession number NC_002516.2 was retrieved in FASTA format from the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov>). The PlpD corresponding amino acid sequence with the accession number NP_252029.1 was obtained from the NCBI protein database (<https://www.ncbi.nlm.nih.gov/protein/>).

Multiple sequence alignment (MSA) of the whole selected protein sequence of the PlpD was performed against 12 medically important *P. aeruginosa* strains: PAO1, PAK, PA14, PA1, PA7, PA103, BL13, E2, LESB58, NCGM₂-S₁, M18, and SCV20265. The sequence accession number for all the 12 mentioned strains are NP_252029.1 (PAO1), EOT19396.1 (PAK), EOT15094.1 (PA14), AHA17132.1 (PA1), ABR85663.1 (PA7), EYU06765.1 (PA103), ERY35050.1 (BL13), EKA56079.1 (E2), CAW26453.1 (LESB58), BAK91293.1 (NCGM₂-S₁), AEO74113.1 (M18), and AHC75823.1 (SCV20265).

The corresponding amino acid sequence of the PlpD of *P. aeruginosa* PAO1 was subjected to blastp against the non-redundant protein sequence (nr) database in the NCBI server (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). The obtained sequences with a similarity >90%, coverage >90%, and E-value <10⁻⁴, were selected and aligned with MSA tools from the European Bioinformatics Institute server (Thompson et al., 1994).

2.2. Predictions of antigenic, allergenic and physicochemical properties of PlpD

The prediction of allergenic property has become significantly pellucid, due to the consumption of modified proteins therapeutics, or bio-pharmaceuticals, etc. Allpred is an online tool that allows the prediction of allergens based on cross and merging diverse methods of IgE epitope monitoring, SVMs, MEME/MAST, and blast search. The presence of 2890 allergen-representative peptides (ARPs) in the server makes Allpred the top choice compared to the other servers (Saha and Raghava, 2006a). In order to evaluate several basic physicochemical properties, the primary structure of PlpD protein in FASTA format was subjected to ProtParam analysis (<http://web.expasy.org/protparam/>).

2.3. Expectations of secondary, tertiary conformation, refinement and validation for PlpD

The PlpD secondary structures such as alpha helix and beta sheet were studied by the MINNOU server (Cao et al., 2006). In the next step, the turn structures were detected by using the AlphaPred, BetaTPred2, and GammaPred servers (Kaur and Raghava, 2004). The prediction of 3D structure for the PlpD was

done by the I-Tasser and RaptorX structure prediction servers. The I-Tasser service (iterative threading assembly refinement) is the top choice server that uses a hierarchical method for protein structure prediction (Roy et al., 2010). On the other hand, RaptorX is an excellent server for predicting the 3D structures for protein sequences without close homologs in the Protein Data Bank (PDB) (Källberg et al., 2012). The RaptorX server assigns P-value for the relative global quality, GDT (global distance test) and uGDT (un-normalized GDT) for the absolute global quality, and RMSD for the absolute local quality of each residue in the model as confidence scores (C-score) that indicate the quality of a predicted 3D model. The predicted 3D structures were corrected by minimizing the atomic-level energy for which the ModRefiner server was used (Xu and Zhang, 2011). By using the ProSA server, the predicted models were validated (Wiederstein and Sippl, 2007). Parameters such as hydrophilicity, flexibility, accessibility, turns, exposed surface, polarity, and antigenic propensity of the polypeptide chains were correlated with the location of continuous epitopes.

2.4. Model stability assessment by ramachandran plot

In a polypeptide, the main chain N-C α and C α -C bonds are relatively free to rotate. Ramachandran diagram ($\{\phi, \psi\}$ plot) is the simplest and most sensitive tool that is used for validating and evaluating the quality of a computational protein model in the absence of experimental data (Kleywegt and Jones, 1996). This method visualizes the conformations of the dihedral angles ψ against ϕ in amino acid residues for protein structure.

2.5. B-cell epitope mapping

B-cell epitope mapping was used to find the protective B-cell epitopes, which played a significant role in the vaccine enhancement process. The primary sequence of PlpD was subjected to the BepiPred server for predicting the flexible length linear B-cell epitopes. The BepiPred server predicts the location of linear B-cell epitopes by using a combination of a hidden Markov model and a transpicuous scale method (Larsen et al., 2006a). In order to confirm the accuracy of the predicted antigenic patches, the results were subjected to two other B-cell epitope mapping web services ABCpred (Saha and Raghava, 2006b) and Vaxijen v2.0 (Doytch-inoova and Flower, 2007) servers, which have their own specific algorithms. The cut-off scores of >0.5 and >0.75 were selected for ABCpred and Vaxijen, respectively. These cut-off values are the measures of prediction accuracy. Based on the ABCpred and Vaxijen methods, if the output value was greater than 0.5 and 0.75, respectively, then that peptide was predicted as a B-cell epitope, otherwise as a non-epitope.

Discontinuous epitopes were assumed to form the majority of strong antibody binding epitopes on approximately the entire proteins. The conformational B-cell epitopes in the full length PlpD were predicted by the CBTOPe (Ansari and Raghava, 2010) and DiscoTope (Kringelum et al., 2012) servers. For the CBTOPe-based prediction, we used the default server SVM threshold (0.0) to reconfirm the predictions. The PlpD primary structure was used by CBTOPe with the overall accuracy of 85% perspicuously. Discontinuous B-cell epitopes were predicted by DiscoTope web service in which a well-organized 3D PlpD refined model was used. The results of the DiscoTope server were interpreted on the basis of the propensity and DiscoTope scores. The propensity score indicates the probability/tendency of a certain amino acid of being part of an epitope. On the other hand, DiscoTope score is intended by merging the contact numbers with the propensity score. When the output value is greater than the threshold value, it is predicted as B-cell epitope. The default threshold (−3.7) was selected at DiscoTope. At this threshold, the sensitivity and specificity are 0.47

and 0.75, respectively. A sensitivity of 0.47 means that 47% of the epitope residues were predicted as part of epitopes and a specificity of 0.75 reveals that 25% of the non-epitope residues were predicted as part of epitopes.

2.6. Prediction of human HLA restricted antigenic determinant

The PlpD was subjected to ProPred server to predict MHC class II restricted peptides for 51 alleles (Singh and Raghava, 2001). In addition, MHC class I binding peptides for 47 alleles were detected by Pro-PredI server (Singh and Raghava, 2003). Fifty-percent maximal inhibitory concentration (IC₅₀) has been established as a degree of the efficiency of a molecule to inhibit 50% of a biochemical function or the dosage of a molecule that displaces 50% of a competitive ligand. The IC₅₀ for the predicted epitopes from both servers were calculated by using the MHCpred server (Guan et al., 2003). The epitopes with IC₅₀ value of < 1 000 nM for allele DRB1*0101 (DR1) were recognized as sufficient enough epitope and most frequently encountered MHC class II alleles in diverse populations.

3. Results

3.1. Sequence analysis and multiple sequence alignment

In this investigation, we gathered the whole sequence of the bi-functional lipase, PlpD, from *P. aeruginosa* PAO1, which could be secreted into the extracellular medium. The blast of the complete PlpD sequence among the 12 selected *P. aeruginosa* strains revealed that the PlpD molecule could be expressed by the mentioned strains. The multiple alignment of the complete PlpD sequence showed that 697 (95.74%) out of 728 residues were conserved

among the selected *P. aeruginosa* strains. The multiple RssA and surface antigen VNR domain sequence alignment revealed that these domains were conserved 100% in the entire selected organisms except in PA7. The amino acid sequences of these domains in PA7 strain were fairly different from the other strains. Indeed, this analysis on the Bac surface antigen showed that 283 (97.25%) amino acids out of 291 residues were conserved among the entire preferred strains (Fig. 1). On the other hand, the multiple alignment analysis revealed several amino acid changes in the whole PlpD protein sequence in the PA7 strain as compared to all the selected strains. These amino acid changes include V7L, L8P, L14F, S15F, A16C, T19A, A21T, I40V, I53V, V65I, Q88K, P95S, T124S, N151T, K156R, I214V, D221E, S243A, Y282F, V303L, K315Q, L317R, D323A, T361A, V507I, S536G, T620A, E685Q, and I697F as compared to the other strains included in this study.

3.2. Physico-chemical parameters of PlpD

The analysis of the physical and chemical parameters for the PlpD protein revealed that the total number of amino acids in PlpD was 709 and its molecular weight was calculated as 78827.08 Da. The total number of negatively charged residues (Asp + Glu) was 97, and the total number of positively charged residues (Arg + Lys) was 82. The theoretical pI of the protein was 5.21. Furthermore, the estimated half-life for the PlpD was predicted to be 1 h in mammalian reticulocytes (in vitro), 30 min in yeast (in vivo) and >10 h in *Escherichia coli* (in vivo) environment. The instability index for the PlpD was estimated as 25.39, which categorized the protein as a stable molecule. The extinction coefficient value at 280 nm, aliphatic index, and grand average hydropathy (GRAVY) for the protein was predicted as 72200 M⁻¹ cm⁻¹, 91.99, and -0.305, respectively.

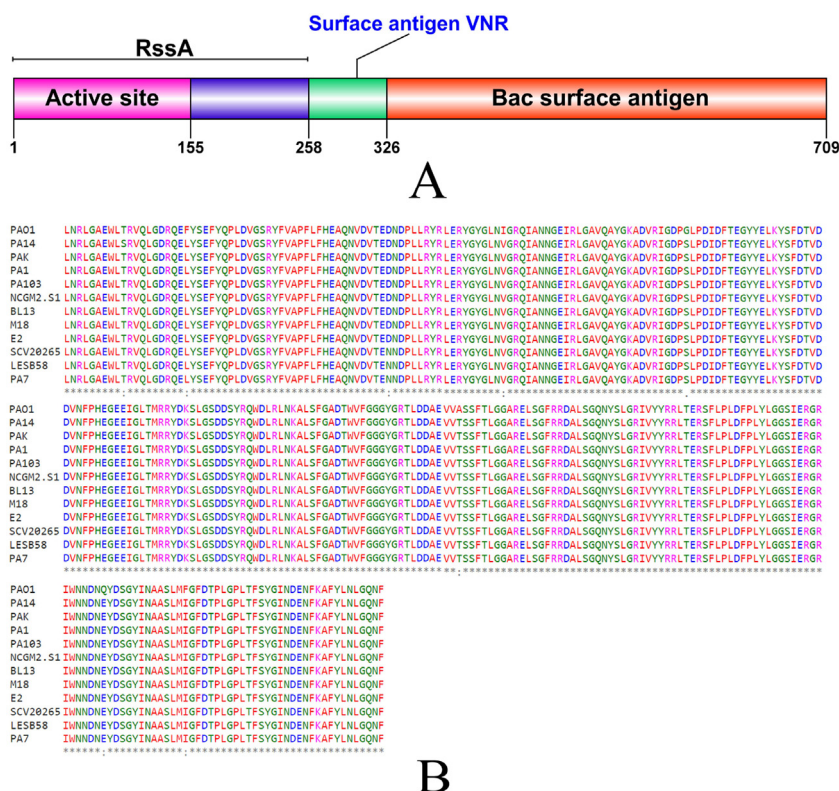


Fig. 1. The graphic representation of *P. aeruginosa* PAO1 PlpD consists modified RssA subunit integrated with two other T5dSS apparatuses components, surface antigen VNR and Bac surface antigen (A). The Multiple Sequence alignment of Bac surface antigen fragment of PlpD between some medically important *P. aeruginosa* strains; PAO1, PA14, PAK, PA1, PA103, NCGM2-S1, BL13, M18, E2, SCV20265, and LESB58. The result showed several highly conserved amino acids in the full length Bac surface antigen (B).

3.3. IgE epitope monitoring

The prediction of allergenic properties confirmed that the PlpD sequence was a non-allergenic protein. The prediction by mapping of the IgE epitope did not recognize the proven IgE epitope in the full length PlpD. Indeed, the specified scores reported by the amino acid composition (threshold value = -0.4) and dipeptide composition based methods (threshold value = -0.2) were 1.0110357 and -1.3612064 , respectively.

3.4. Prediction of PlpD tertiary and secondary structures and its refinement

Previous studies have confirmed the correlation between the location of antigenic sites and the presence of conformational turns as an environmentally exposed area in immunogenic proteins (Salacha et al., 2010; Ben Haj Khalifa et al., 2011; Tomar and De, 2010). The MINNOU server defined each residue into alpha helix, beta sheet, and random coil structures. This server predicted 13 alpha helices, 27 beta sheets, and 42 random coil fragments in the PlpD protein (Fig. 2). In addition, our prediction for finding out the turns clarified 15 alpha turns, 38 beta turns, and 47 gamma turns in the antigen. The three-dimensional structure of the PlpD from its primary structure was predicted by I-TASSER (Fig. 3A) and RaptorX (Fig. 3B) web services. Based on multiple-threading alignments, the sequence homology comparison, and comparative modeling methods, five 3D model structures for the PlpD were predicted by the I-TASSER server. Considering the quality of a prediction, in this study, we selected the top confidence score (C-score) of the predicted models defined as a confidence and/or the quality of the predicted models that was quantitatively measured by the I-TASSER server. Typically, the first model had a higher C-score and a better quality in most of the predicted cases (Tomar and De, 2014). In our study, the best quality in most of the predicted cases achieved -0.55 in C-score, which according to the C-score range ($[-5, 2]$) could be a good confidence for the best quality of predicted model. The estimated templates modeling (TM) score and the root-mean-square deviation (RMSD) for the top model were 0.64 ± 0.13 and 9.4 ± 4.6 Å, respectively. Based on the RaptorX server prediction, 709 (100%) residues were divided into two domains. The first domain extended from amino acid 1 to 297, while the second predicted domain extended from 298 to 709. The P-value, uGDT(GDT), uSeqId(SeqId), and the score of domain 1 were $6.44e^{-11}$, 251 (84), 246 (83) and 268, respectively. These

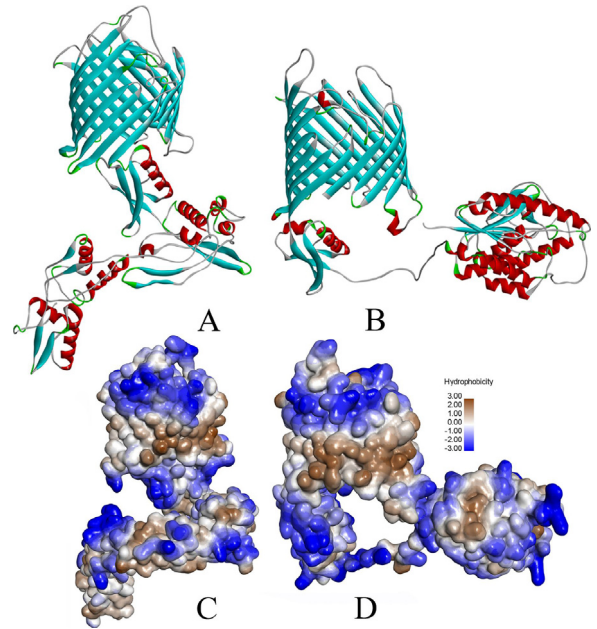


Fig. 3. Three-dimensional structure for PlpD from its primary structure predicted by I-TASSER (A) and RaptorX (B) servers which refined by modrefiner service. The 3D hydrophobicity modeling of PlpD by using I-TASSER (C) and RaptorX (D) after refinement process. The brown color represents the most hydrophobic region, while the blue color indicates antibody-accessible and high score hydrophilic residues. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

measures for domain 2 were, respectively, $1.95e^{-14244}$ (59), 6 (16), and 261. The refinement result from the I-TASSER predicted model showed several structural changes. The RMSD and TM-score for the refined I-TASSER model were 1.123 and 0.9854, respectively. The RMSD for the refined RaptorX predicted model was 0.900 and its TM-score was indeed predicted to be 0.9908. The corresponding predicted 3D hydrophobicity modeling for the I-TASSER and the RaptorX are shown in Fig. 3C and D, respectively.

3.5. Validation of predicted PlpD structures

According to the ProSA-web server, the z-score indicates the overall model quality and measures the deviation of the total energy of the structure with respect to an energy distribution derived from

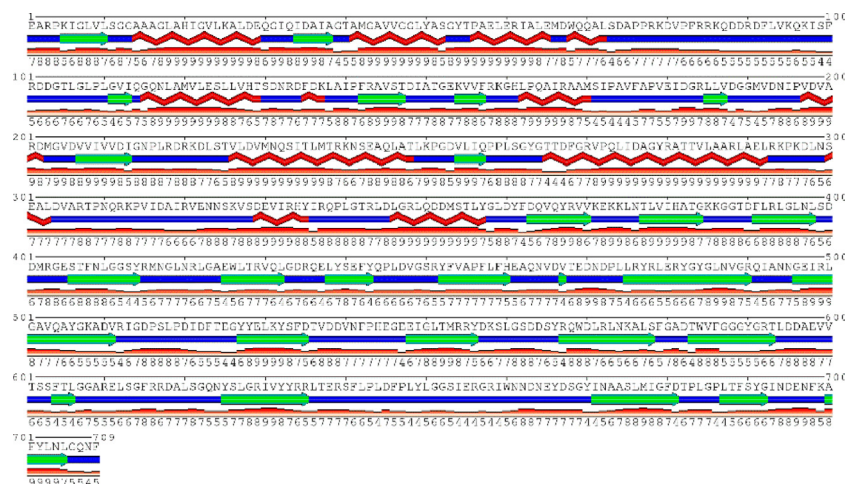


Fig. 2. Analysis of PlpD protein secondary structure by MINNOU server. The blue line shows random coil. The green and red colors represent beta sheet and alpha helix constructions respectively. The orange columns and numbers show value of confidence of prediction (minimum value is 1, maximum value is 10) for each amino acid. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

random conformations. The ProSA-web server z-score values for the predicted PlpD 3D model by the I-TASSER server and its refined structure were -5.73 (Supplementary Fig. 2A) and -5.93 (Supplementary Fig. 2C), respectively. However, the z-scores for the RaptorX modeling before and after refinements were -7.91 (Supplementary Fig. 2B) and -8.57 (Supplementary Fig. 2D), respectively. The ProSA-web server plot designed by groups of structures from different sources (X-ray, NMR) and are distinguished by different colors (Thompson et al., 1994). Based on our data, the location of the z-score for both the refined structures of I-TASSER (-5.93) and RaptorX (-8.57) are in the range of native conformations, but RaptorX achieved better z-scores value, which indicated better overall model quality as compared to the I-TASSER result. The outcomes for the Ramachandran plot from the predicted models showed that 510 (72.1%) residues from I-TASSER initial model were in the favored region (A, B and L) and 135 (19.1%) amino acid residues were in the allowed region (a, b, l, and p). As a matter of fact, only 62 (8.8%) residues of the PlpD primary I-TASSER model were in the outlier region. However, the refined I-TASSER model showed that 613 (86.7%) residues were in the favored region and 77 (10.9%) residues were in the allowed region, whereas the number of residues in the outlier region was 17 (2.4%). The 17 residues in the refined I-TASSER model that covered the outlier region were Ala15, Met66, Pro82, Arg85, Gly104, Ser130, Val199, Arg201, Asp202, Leu223, Leu236, Pro258, Pro339, Asp360, Gly433, Gly580, and Pro647 (Fig. 4A). In order to compare the statistical data, the initial and the refined models from the RaptorX web server were evaluated by the RAMPAGE server. The results revealed that 668 (94.5%) residues were in the favored region, 31 (4.4%) were in the allowed region, and only 8 (1.1%) residues were detected in the outlier region. On the other hand, the Ramachandran plot of the RaptorX refined model detected 675 (95.5%) amino acids in the favored region and 30 (4.2%) in the allowed region, and in contrast, only 2 (0.3%) residues (Ala619 and Thr666) were in the outlier region (Fig. 4B).

3.6. Prediction of B-cell epitopes

In this study, the numerous flexible length linear B-cell epitopes in the PlpD molecule were predicted by using the BepiPred 1.0

Server and on the basis of a random forest algorithm trained on the epitopes annotated from antibody-antigen protein structures. The largest predicted epitope fragment included 20 amino acids in length. From several results obtained by the BepiPred server, only five epitope fragments achieved the highest (above 0.6 thresholds) scores from the Vaxijen server (Table 1). In another evaluation, the discontinued B cell epitope was monitored by using the CBTOPE server. Each amino acid was designed as a probability value based on a scale ranging from 0 to 9 for detecting antigenic properties. A higher value indicates the elevated probability of the residue as an antibody epitope. The CBTOPE server results are presented in Table 2.

The results obtained from the DiscoTope server were based on the propensity and the DiscoTope scores. The propensity score indicates the probability/tendency of being part of an epitope for that particular residue. The propensity is represented by amino acid epitope log-odds ratios, which are calculated on a set of antigens. The DiscoTope score is calculated by combining the contact numbers with the propensity score. A score above the threshold value (-3.7) indicates positive predictions and that below the threshold value indicates negative predictions. As per

Table 1

The flexible length epitopes predicted by BepiPred were evaluated by Vaxijen server based on Auto Cross Covariance (ACC) transformation of protein sequences into uniform vectors of the principal amino acid properties.

No	Position	Epitope	Length	Vaxijen Score
1	51–58	ASGYTPAE	8	0.4389
2	70–89	QALSDAPPRKDVFPFRKQDD	20	0.6412
3	215–222	NPLRDRKD	8	−0.3179
4	259–270	PLSGYGTDFGR	12	−0.0722
5	293–300	RKPKDLNS	8	1.1999
6	303–314	LDVARTPNQRKP	12	0.5475
7	464–473	NVDVTEENDP	10	0.4895
8	510–524	DVRIGDPSLPDIDFT	15	0.7641
9	539–548	DVNFPEEGEE	10	0.3693
10	558–567	KSLGSDDSYR	10	1.0941
11	589–599	GYGRTLDAAEV	11	0.6980
12	659–671	IWNNDNEYDSGYI	13	−0.2832

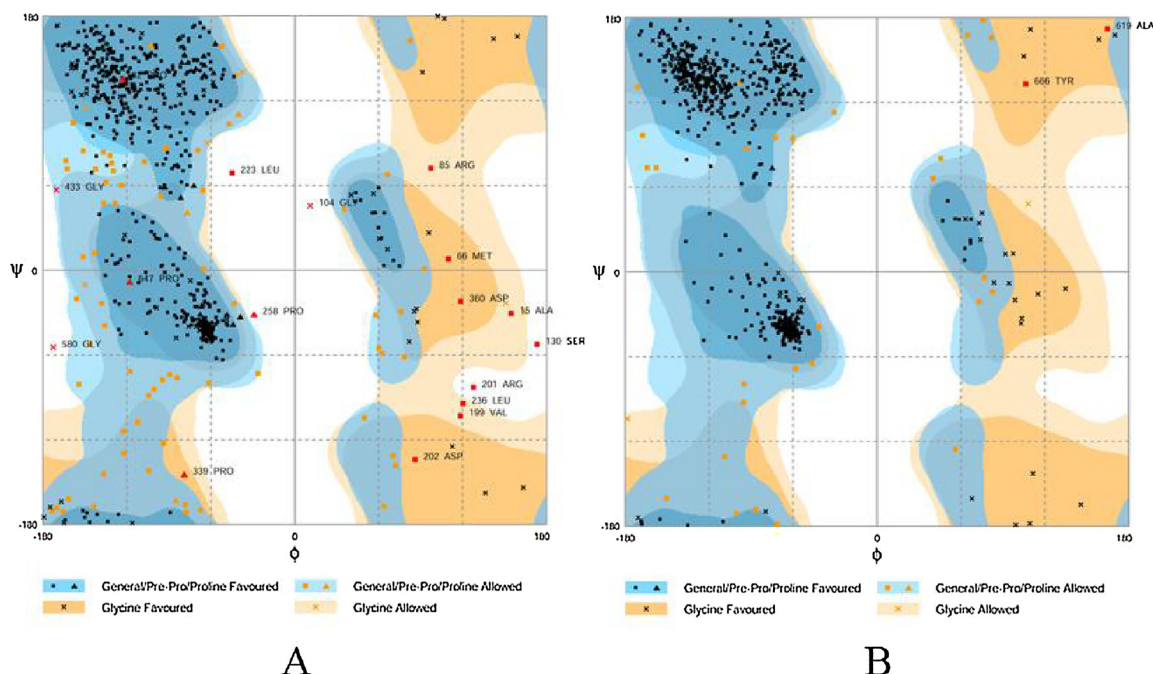


Fig. 4. Validation and desirable protein stability for the refined structure of rPlpD protein predicted by I-TASSER (A) and RaptorX (B) servers by using RAMPAGE server.

Table 2

Conformational B-cell epitopes predicted from full length PlpD by using CBTOPE server.

Amino acid	Position	Probability scale	Amino acid	Position	Probability scale
AR	2–3	4	LQ	349–350	4
G	7	4	FDQ	362–364	4
VLS	9–11	4	RV	368–369	4
IGVL	21–24	4	LVI	378–380	4
GY	53–54	4	H	381	5
T	55	5	A	382	4
PAEL	56–59	4	L	392	4
IA	62–62	4	LNLS	405–409	4
LSDAP	71–75	4	FYQP	451–455	4
D	80	4	E	469	4
D	102	4	LR	476–477	4
L	110	4	AV	502–503	4
NRD	132–134	4	PDI	519–521	4
V	254	4	FTE	523–525	4
IQPP	256–259	4	T	536	4
GYGT	262–265	4	EEIG	547–550	4
G	278	4	YDKS	556–559	4
R	280	4	SYRQWDLRLN	565–574	4
AA	286–287	4	S	578	4
P	295	4	RDALS	617–621	4
NSEALDV	299–305	4	GQ	622–623	5
VID	315–317	4	NYS	624–626	4
IRVEN	319–323	4	G	628	4
SKVSDEV	325–331	4	IVY	630–632	4
I	332	5	T	637	4
R	333	6	I	654	4
H	334	5	RIWNNNDNEYDSGYI	658–671	4
Y	335	4	A	673	4

these facts, the server identified 173 potential conformational B-Cell epitope residues in the PlpD molecule (Kringelum et al., 2012). The result showed that 22 single amino acids achieved top propensity and DiscoTope scores (Table 3).

Table 3

The predicted results of conformational B-cell epitopes by using Discotope-2.0 Server.

Amino acid	Position Contact	Propensity score	DiscoTope score	Amino acid	Position Contact	Propensity score	DiscoTope score
Pro76	7	1.614	0.623	Gln350	3	0.450	0.053
Pro77	3	2.844	2.172	Ser565	6	1.209	0.380
Arg78	0	2.118	1.874	Asp596	7	1.751	0.745
Lys79	4	2.697	1.927	Trp660	13	1.921	0.205
Asp80	1	3.315	2.819	Asn661	6	2.529	1.548
Val81	4	2.700	1.929	Asn662	18	2.827	0.432
Arg85	5	1.679	0.911	Asp663	14	1.832	0.011
Asn310	2	1.445	1.048	Asn664	12	2.306	0.661
Gln311	10	1.668	0.326	Glu665	0	2.044	1.809
Arg312	9	1.814	0.570	Tyr666	4	2.445	1.704
Lys313	8	1.741	0.621	Ser668	2	1.724	1.296

3.7. Identification of T-cell-restricted epitopes

Recognizing T-cell restricted antigenic epitopes is quite essential for understanding the pathogenesis of diseases and their etiologies particularly for designing vaccine candidates. The ProPred-I server identified 21 MHC Class-I restricted epitopes, whereas ProPred detected 35 valuable epitopes for MHC class-II. Some of the predicted MHC Class-I and MHC Class-II restricted epitopes with their IC_{50} values are listed in Supplementary Table. Moreover, the comparative analysis between the predicted B-cell epitopes and the identified T-cell antigenic determinants showed that several antigenic segments were expected to provoke strong humoral and cellular immune responses reciprocally.

4. Discussion

At present, a majority of infection types caused by *P. aeruginosa* is a clinical problem globally. The attempt to find an efficient and effective vaccine against this pathogen is not a new concept. In this study, the PlpD molecule was chosen as a suitable candidate for targeting and inducing effective immune responses against *P. aeruginosa*. In the past decade, the exceptional roles of lipases were confirmed such as in cellular functions, including membrane maintenance, cellular turnover, and inflammatory response generation (Salacha et al., 2010; Burke and Dennis, 2009). A new family of bacterial lipolytic enzymes, PlpD, has been proposed recently (Salacha et al., 2010; Banerji and Flieger, 2004; da Mata Madeira et al., 2016). The lipolytic passenger domain of PlpD is cleaved auto-catalytically after completion of auto-transport and could be secreted into the extracellular medium (Salacha et al., 2010). Former studies have indicated that the outer membrane proteins and the surface-expressed secretion system apparatus of *P. aeruginosa* could be protective in a number of acute and chronic infectious diseases in mouse models (Priebe and Goldberg, 2014). By using blastp tool in this study, we identified the PlpD homologue sequences in the numerous strains of *P. aeruginosa* (Salacha et al., 2010). The homology study and the alignment results showed that the PA7 associated sequence was virtually heterogeneous among the studied strains. These facts assume that the selected sequence could be useful in stimulating strong immune responses against most of the *P. aeruginosa* strains. Based on a previous investigation, it was confirmed that the full-length heterologous PlpD molecules are produced in *E. coli* BL21 (DE3) (da Mata Madeira et al., 2016). In *P. aeruginosa*, lipases, alone or synergistically with other virulence factors, can cause cell death, severe tissue destruction, and necrosis in the human host (Tielen et al., 2013). The majority of solvent-accessible surface area preferentially recognized by antibodies (antigenic epitopes) can assist in the development of vaccine components against pathogens and immuno-diagnostic reagents (Liang et al., 2009). These regions are properly scattered in PlpD the sequence. The

calculated isoelectric point (pI) being lower than 7 (pI=5.21) indicates that PlpD is acidic in nature.

The molar extinction coefficient for the PlpD molecule at 280 nm was as high as ($72200 \text{ M}^{-1} \text{ cm}^{-1}$), thereby representing a high concentration of amino acids containing aromatic side chains such as Phe, Trp, and Tyr. The high PlpD aliphatic index could be a positive factor in implying stability to the wide-range of temperature fluctuations. Antigen solubility is an important pre-requirement for biophysical studies on the epitope-based vaccine design. The low GRAVY value (-0.305) indicates a possible better interaction of PlpD with water. The ProtParam web service classified PlpD as being stable in the test tube (instability index <40). The long estimated half-life of PlpD (more than 10 h in vitro) and its high stability are attractive features for protein-based vaccine candidates. The PlpD protein sequence does not contain experimentally proven IgE epitope; nonetheless, the mapping of IgE epitope(s) did not predict PlpD protein as an allergen molecule either. PlpD contains a high percentage of random coils which could preferably be recognized by antibodies (Saad et al., 1988). The use of reliable three-dimensional atomic models of proteins is often crucial for vaccine design. Hence, we predicted the PlpD 3D structure by using two different methods, the close homologs method in the Protein Data Bank (PDB) and the alignment based comparison between the target sequence and one or multiple distantly related template proteins by using a novel nonlinear scoring function and a probabilistic-consistency algorithm. For the I-TASSER predicted structures, the RMSD values of less than 3 Å ($\text{\AA}$) indicated high accuracy and TM-score >0.5 showed a model of correct topology (Xu and Zhang, 2010). The confidence score or C-score (-5 to 2) is the quality indicator for the predicted model. Higher values signify a high confidence and vice-versa. These values interpret the precision of the predicted 3D PlpD structure. The RaptorX server uses different measures to evaluate the quality of a predicted 3D structure model. The P-value estimates the relative quality of a model. A smaller P-value indicates a higher quality of the model. The uGDT is the un-normalized global distance test score that measures the absolute model quality. For a protein with >100 residues, uGDT and GDT >50 is a proper indicator. The comparisons of the four ProSA-web validation results indicated that all the predicted features were in the range of native structures. The ProSA Z-scores clearly showed that the RaptorX refined configuration could be the closest structure (with the Z-score = -5.87) to the native state. The Ramachandran plot predictions showed desirable protein stability in each refined model with slight differences. To be effective, a *P. aeruginosa* vaccine candidate should be able to induce strong B-cell and humoral mediated immunity. For this reason, the identification of the protective B-cell antigenic determinant plays a key role in designing vaccines or antibody mediated immunotherapy and immuno-diagnostic tests (Casadevall and Pirofski, 2012). A previous study revealed that the use of bioinformatics approach to compute B-cell epitopes was correlated with the experimental method (Kringelum et al., 2012). In this experiment, the Bepipred server used for the prediction of flexible length continuous B-cell epitopes utilized a combination method by combining the hidden Markov model with one of the best propensity scale methods (Larsen et al., 2006b). The adaptation of our predicted secondary structure and the B-cell linear epitopes showed that those antigenic determinants were significantly enriched with irregular conformation and turn structures. A comparison of the surface accessibility predictions and the B-cell linear epitope sequences indicated that most of the predicted epitopes were surface unexposed sequences, capable of stimulating the production of specific antibodies. Both the conformational epitope sequences and the secondary structure prediction results showed that the structural epitopes derived from CBTOPE and DiscoTope were also

enriched with irregular structures and turn motifs. The CBTOPE prediction was done by the amino acid sequence of PlpD, whereas DiscoTope used the refined 3D structure of PlpD which seems to be more reliable. According to the data obtained from this experiment, the T-cell epitope mapping for PlpD was accomplished by using ProPred and ProPred-I. As of now, many types of computational and immunoinformatics T-Cell epitope mapping approaches have been developed. However, there is no consensus and comprehensive approach in this area (Lafuente and Reche, 2009). In this study, several T-cell epitopes from PlpD were predicted with strong binders ($\text{IC}_{50} \leq 50$) of class I and class II MHC which could trigger strong immune response. These data revealed that according to parameters such as the number of binding alleles and IC_{50} value, the class II MHC-restricted epitopes are higher than MHC Class-I binding epitopes. It is quite eminent to mention that the Class II MHC is loaded with extracellular proteins, such as wound or the blood associated extracellular pathogens (*P. aeruginosa*, *A. baumannii*, etc.) components and mainly interact with the T-helper cell (TCD4^{+}) immune cells. The result of this direct interaction may include localized inflammation and the recruitment of antigen presenting cells and macrophage or may lead to a full-force effective antibody mediated immune response due to the activation of B-cells against these extracellular pathogens. On the other hand, it is noteworthy to say that in the development of peptide-based immunotherapeutic and vaccines, finding or synthesis of peptide fragments that stimulate both MHC II and I mediated immunity is highly useful and recommended. Accordingly, our evaluations revealed some antigenic determinants of PlpD which could induce both B-cell and T-cell mediated immunity synergistically. It is noteworthy to mention that three of the selected epitopes were 100% conserved in all the selected strains, even in PA7 as a heterogenic strain for the PlpD molecule, and just "ALSDAPPRKDVFPFRKQDDR" and "WVFGGGYGRITLDDAEVVTSSFTLGGA" exclusively had variations in P95S (a non- polar to polar change) and T620A (a polar to non-polar change) in the two mentioned epitopes in the PA7 strain, respectively. Despite these 29-point changes in the whole sequence of the PlpD protein in PA7, it seems that these epitopes could be used effectively against all the *P. aeruginosa* strains. To sum-up, our study clearly revealed that the *P. aeruginosa* PlpD molecule significantly enriched the immunogenic peptide fragments, thereby qualifying for further in vitro and in vivo immunological investigations.

5. Conclusion

The inhibition of *P. aeruginosa* PlpD activity as a lipolytic enzyme by adaptive and cellular immune responses could have essential therapeutic potential against this ubiquitous opportunistic pathogen. In this study, according to computational analysis and some strategic immunoinformatics-based predictions, we concluded that the peptide sequences including "ALSDAPPRKDVFPFRKQDDR", "ITLMTRKNSAQATLTKPGD", "KEKKLNTLVIHATGKKGGTD", "DDMRGESTFNLGGSYRMNGL", and "WVFGGGYGRITLDDAEVVTSSFTLGGA" could serve as peptide-based immunogenic substances against *P. aeruginosa*. On the other hand, the stimulation of effective and specific immune responses against the PlpD molecule could interfere with the enzymatic function of this cytotoxic lipase. Further studies are recommended regarding the production of monoclonal antibody against *P. aeruginosa* Patatin-like protein D in the wounds of burned and cystic fibrosis patients.

Conflict of interest

The authors declare no conflicts of interest.

Ethical issues

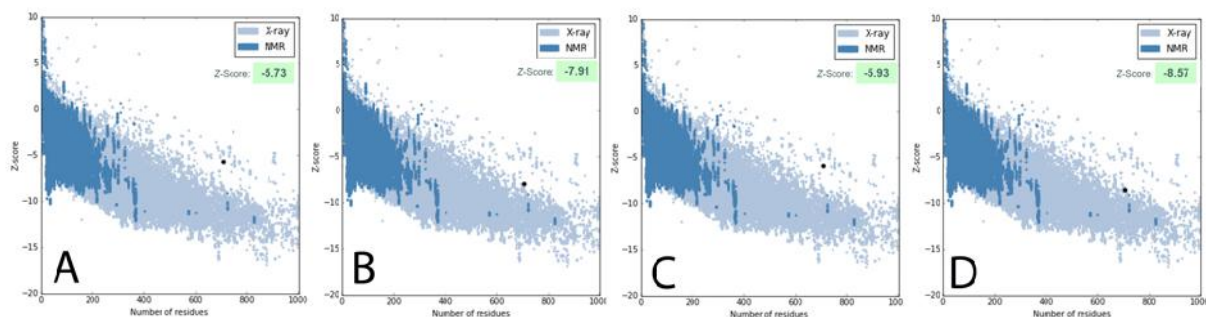
Not applicable.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.compbiolchem.2018.02.001>.

References

- Ansari, H.R., Raghava, G.P., 2010. Identification of conformational B-cell epitopes in an antigen from its primary sequence. *Immunom. Res.* 6, 6. doi:<http://dx.doi.org/10.1186/1745-7580-6-6>.
- Banerji, S., Flieger, A., 2004. Patatin-like proteins: a new family of lipolytic enzymes present in bacteria? *Microbiology (Reading England)* 150 (Pt. 3), 522–525. doi:<http://dx.doi.org/10.1099/mic.0.26957-0>.
- Ben Haj Khalifa, A., Moissenet, D., Vu Thien, H., Khedher, M., 2011. Virulence factors in *Pseudomonas aeruginosa*: mechanisms and modes of regulation. *Ann. Biol. Clin. (Paris)* 69 (4), 393–403. doi:<http://dx.doi.org/10.1684/abc.2011.0589>.
- Burke, J.E., Dennis, E.A., 2009. Phospholipase A2 structure/function, mechanism, and signaling. *J. Lipid Res.* S237–S242. doi:<http://dx.doi.org/10.1194/jlr.R800033-JLR200>.
- Cao, B., Porollo, A., Adamczak, R., Jarrell, M., Meller, J., 2006. Enhanced recognition of protein transmembrane domains with prediction-based structural profiles. *Bioinformatics (Oxford, England)* 22 (3), 303–309. doi:<http://dx.doi.org/10.1093/bioinformatics/bti784>.
- Casadevall, A., Pirofski, L.-a., 2012. A new synthesis for antibody-mediated immunity. *Nat. Immunol.* 13 (1), 21–28.
- da Mata Madeira, P.V., Zouhir, S., Basso, P., Neves, D., Laubier, A., Salacha, R., et al., 2016. Structural basis of lipid targeting and destruction by the type V secretion system of *pseudomonas aeruginosa*. *J. Mol. Biol.* 428 (9 Pt A), 1790–1803. doi:<http://dx.doi.org/10.1016/j.jmb.2016.03.012>.
- Doytchinova, I.A., Flower, D.R., 2007. Vaxijen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinf.* 8, 4. doi:<http://dx.doi.org/10.1186/1471-2105-8-4>.
- Gellatly, S.L., Hancock, R.E.W., 2013. *Pseudomonas aeruginosa*: new insights into pathogenesis and host defenses. *Pathog. Dis.* 67 (3), 159–173. doi:<http://dx.doi.org/10.1111/2049-632X.12033>.
- Guan, P., Doytchinova, I.A., Zygori, C., Flower, D.R., 2003. MHCpred: bringing a quantitative dimension to the online prediction of MHC binding. *Appl. Bioinformatics* 2 (1), 63–66.
- Källberg, M., Wang, H., Wang, S., Peng, J., Wang, Z., Lu, H., et al., 2012. Template-based protein structure modeling using the RaptorX web server. *Nat. Protoc.* 7 (8), 1511–1522.
- Kaur, H., Raghava, G.P., 2004. Prediction of alpha-turns in proteins using PSI-BLAST profiles and secondary structure information. *Proteins* 55 (1), 83–90. doi:<http://dx.doi.org/10.1002/prot.10569>.
- Kleywegt, G.J., Jones, T.A., 1996. Phi/Psi-chology: Ramachandran revisited. *Structure* 4 (12), 1395–1400. doi:[http://dx.doi.org/10.1016/S0969-2126\(96\)00147-5](http://dx.doi.org/10.1016/S0969-2126(96)00147-5).
- Kringelum, J.V., Lundegaard, C., Lund, O., Nielsen, M., 2012. Reliable B cell epitope predictions: impacts of method development and improved benchmarking. *PLoS Comput. Biol.* 8 (12), e1002829. doi:<http://dx.doi.org/10.1371/journal.pcbi.1002829>.
- Lafuente, E.M., Reche, P.A., 2009. Prediction of MHC-peptide binding: a systematic and comprehensive overview. *Curr. Pharm. Des.* 15 (28), 3209–3220.
- Larsen, J.E., Lund, O., Nielsen, M., 2006a. Improved method for predicting linear B-cell epitopes. *Immunom. Res.* 2, 2. doi:<http://dx.doi.org/10.1186/1745-7580-2-2>.
- Larsen, J.E.P., Lund, O., Nielsen, M., 2006b. Improved method for predicting linear B-cell epitopes. *Immunom. Res.* 2, 2-. doi:<http://dx.doi.org/10.1186/1745-7580-2-2>.
- Liang, S., Zheng, D., Zhang, C., Zacharias, M., 2009. Prediction of antigenic epitopes on protein surfaces by consensus scoring. *BMC Bioinf.* 10 (1), 1–10. doi:<http://dx.doi.org/10.1186/1471-2105-10-302>.
- Priebe, G.P., Goldberg, J.B., 2014. Vaccines for *Pseudomonas aeruginosa*: a long and winding road. *Expert Rev. Vaccines* 13 (4), 507–519. doi:<http://dx.doi.org/10.1586/14760584.2014.890053>.
- 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. doi:<http://dx.doi.org/10.1038/nprot.2010.5>.
- Saad, B., Corradin, G., Bosshard, H.R., 1988. Monoclonal antibody recognizes a conformational epitope in a random coil protein. *Eur. J. Biochem./FEBS* 178 (1), 219–224.
- Saha, S., Raghava, G.P., 2006a. AlgPred: prediction of allergenic proteins and mapping of IgE epitopes. *Nucleic Acids Res.* 34 (Web Server issue), W202–W209. doi:<http://dx.doi.org/10.1093/nar/gkl343>.
- Saha, S., Raghava, G.P., 2006b. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. *Proteins* 65 (1), 40–48. doi:<http://dx.doi.org/10.1002/prot.21078>.
- Salacha, R., Kovacic, F., Brochier-Armanet, C., Wilhelm, S., Tommassen, J., Filloux, A., et al., 2010. The *Pseudomonas aeruginosa* patatin-like protein PlpD is the archetype of a novel Type V secretion system. *Environ. Microbiol.* 12 (6), 1498–1512. doi:<http://dx.doi.org/10.1111/j.1462-2920.2010.02174.x>.
- Singh, H., Raghava, G.P., 2001. ProPred: prediction of HLA-DR binding sites. *Bioinformatics (Oxford, England)* 17 (12), 1236–1237.
- Singh, H., Raghava, G.P.S., 2003. ProPred1: prediction of promiscuous MHC Class-I binding sites. *Bioinformatics (Oxford, England)* 19 (8), 1009–1014. doi:<http://dx.doi.org/10.1093/bioinformatics/btg108>.
- Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22 (22), 4673–4680.
- Tielen, P., Kuhn, H., Rosenau, F., Jaeger, K.-E., Flemming, H.-C., Wingender, J., 2013. Interaction between extracellular lipase LipA and the polysaccharide alginate of *Pseudomonas aeruginosa*. *BMC Microbiol.* 13 (1), 1–12. doi:<http://dx.doi.org/10.1186/1471-2180-13-159>.
- Tomar, N., De, R.K., 2010. Immunoinformatics: an integrated scenario. *Immunology* 131 (2), 153–168. doi:<http://dx.doi.org/10.1111/j.1365-2567.2010.03330.x>.
- Tomar, N., De, R.K., 2014. Immunoinformatics: a brief review. *Methods Mol. Biol. (Clifton, NJ)* 1184, 23–55. doi:http://dx.doi.org/10.1007/978-1-4939-1115-8_3.
- Vitkauskienė, A., Skrodenienė, E., Dambrauskienė, A., Macas, A., Sakalauskas, R., 2010. *Pseudomonas aeruginosa* bacteremia: resistance to antibiotics, risk factors, and patient mortality. *Medicina (Kaunas, Lithuania)* 46 (7), 490–495.
- Waters, V., Smyth, A., 2015. Cystic fibrosis microbiology: advances in antimicrobial therapy. *J. Cyst. Fibros.* 14 (5), 551–560. doi:<http://dx.doi.org/10.1016/j.jcf.2015.02.005>.
- Wiederstein, M., Sippl, M.J., 2007. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* 35 (Web Server issue), W407–W410. doi:<http://dx.doi.org/10.1093/nar/gkm290>.
- Xu, J., Zhang, Y., 2010. How significant is a protein structure similarity with TM-score = 0.57? *Bioinformatics (Oxford England)* 26 (7), 889–895. doi:<http://dx.doi.org/10.1093/bioinformatics/btq066>.
- Xu, D., Zhang, Y., 2011. Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization. *Biophys. J.* 101 (10), 2525–2534. doi:<http://dx.doi.org/10.1016/j.bpj.2011.10.024>.



Supplementary Figure: The ProSA server Z-scores of all protein chains in PDB determined by x-ray crystallography (light blue) or nuclear magnetic resonance (NMR) spectroscopy (dark blue) with respect to their length are shown in the all plots. The Z-scores indicate overall model quality. The Figure 7A and 7C indicate Z-scores of initial ($Z = -5.73$) and its refined ($Z = -5.93$) models predicted by I-TASSER server. Preliminary, RaptorX server predicted model ($Z = -7.91$) and its refined structure ($Z = -8.57$) are shown in Figure 7B and 7D respectively. The Z-scores which highlighted as large dots are in the range of native conformations.

Supplementary Table: Common T-cell epitopes from PlpD predicted by Propred, which can produce T-cell mediated immunity, are represented along with their assorted parameters.

MHC Class-I binding regions	MHC Class-I binding alleles	Predicted IC50 Value (nM)	Confidence of IC50 Value prediction for (Max = 1)	MHC Class-II binding regions	MHC Class-II binding alleles	Predicted IC50 Value (nM)	Confidence of IC50 Value prediction for (Max = 1)
ARPKIGLVL	11	416.8	1.00	LVLSGGAAA	17	144.21	0.89
LSGGAAAGLA	11	29.11	1.00	LAHIGVLKA	19	58.34	1.00
AGLAHIGVL	17	356.45	0.89	IGVLKALDE	8	2.89	0.89
IQIDAIAGT	4	51.64	0.89	IQIDAIAGT	11	51.64	0.89
YTPAELERI	9	7.73	0.89	LEMDWQQAL	6	30.48	1.00
LEMDWQQAL	14	30.48	1.00	FLVKQKISFR	18	7.89	1.00
ISFRDDGTL	15	190.99	0.89	FRDDGTGLGL	14	9.84	1.00
DDGTGLGPL	6	23.88	1.00	MVLESLLVH	26	10.42	0.78
LGVIQGQNLAMVL	9	36.14	1.00	LESLLVHTS	29	56.62	0.89
SDNRDFDKL	7	28.44	0.89	LAIPFRAVS	7	8.02	1.00
IPAVFAPVE	2	92.90	1.00	FRAVSTDIA	11	91.62	0.89
IVVDIGNPL	17	4.23	1.00	IRAAMSIPA	24	57.94	0.89
NPLRDRKDL	19	77.09	0.78	IVVDIGNPL	17	4.23	1.00
DRKDLSTVL	6	6.50	1.00	ITLMTRKNS	30	8.63	0.89
TRKNSEAQL	3	23.93	0.78	VLIQPPLSG	36	1035.14	1.00
NSEAQLATL	7	28.31	0.89	YRATTVLAA	32	4.55	1.00
AGYRATTVL	14	187.50	1.00	IRVENNSKV	13	49.09	0.89
AELRKPKDL	10	158.85	0.78	IRHYIRQPL	21	29.17	1.00
IRQPLGTRL	9	14.09	1.00	LVIHATGKK	37	161.06	1.00
YRVVKEKKL	11	6.90	1.00	FLRLGLNLS	36	1.03	0.89
KGGTDFLRL	14	28.51	1.00	MRGESTFNL	17	135.52	0.89
LNRLGAEWL	6	56.75	0.89	YRMNGLNRL	40	2.64	0.89
AEWLTRVQL	12	32.73	1.00	YQPLDVGSR	15	16.03	0.89
IANNGEIRL	11	125.89	0.89	FLFHEAQNV	12	1.63	0.89
EIRLGAVQA	2	125.89	0.89	LLRYRLERY	15	47.10	0.78
DVRIGDPSL	5	272.27	0.78	YGLNVGRQI	6	80.72	0.78
TMRRYDKSL	6	105.93	0.56	IRLGAVQAY	13	35.97	0.89
WDLRLNKAL	10	71.45	0.89	YSFDTVDDV	5	5.18	0.78
FGGGYGRTL	12	941.89	1.00	LRLNKALSF	46	6.40	0.78
EVVTSSFTL	9	55.59	0.89	FRRDALSGQ	19	4.69	0.89
FTLGGAREL	14	11.80	0.78	YYRRLTERS	23	1.39	0.89
EYDSGYINA	3	52.00	0.89	YINAASLMI	24	1.50	0.89
SGYINAASL	16	30.90	1.00	IGFDTPLGP	8	19.54	1.00

