



Quest for Novel Preventive and Therapeutic Options Against Multidrug-Resistant *Pseudomonas aeruginosa*

Sidra Irum¹ · Saadia Andleeb¹ · Amjad Ali¹ · Muhammad Ibrahim Rashid^{1,2} · Mahnoor Majid¹

Accepted: 9 July 2021 / Published online: 9 August 2021
© The Author(s), under exclusive licence to Springer Nature B.V. 2021

Abstract

Pseudomonas aeruginosa (*P. aeruginosa*) is a critical healthcare challenge due to its ability to cause persistent infections and the acquisition of antibiotic resistance mechanisms. Lack of preventive vaccines and rampant drug resistance phenomenon has rendered patients vulnerable. As new antimicrobials are in the preclinical stages of development, mining for the unexploited drug targets is also crucial. In the present study, we designed a B- and T-cell multi-epitope vaccine against *P. aeruginosa* using a subtractive proteomics and immunoinformatics approach. A total of five proteins were shortlisted based on essentiality, extracellular localization, virulence, antigenicity, pathway association, hydrophilicity, and low molecular weight. These include two outer membrane porins; *OprF* (P13794) and *OprD* (P32722), a protein activator precursor *pra* (G3XDA9), a probable outer membrane protein precursor PA1288 (Q9I456), and a conserved hypothetical protein PA4874 (Q9HUT9). These shortlisted proteins were further analyzed to identify immunogenic and antigenic B- and T-cell epitopes. The best scoring epitopes were then further subjected to the construction of a polypeptide multi-epitope vaccine and joined with cholera toxin B subunit adjuvant. The final chimeric construct was docked with TLR4 and confirmed by normal mode simulation studies. The designed B- and T-cell multi-epitope vaccine candidate is predicted immunogenic in nature and has shown strong interactions with TLR-4. Immune simulation predicted high-level production of B- and T-cell population and maximal expression was ensured in *E. coli* strain K12. The identified drug targets qualifying the screening criteria were: UDP-2-acetamido-2-deoxy-d-glucuronic acid 3-dehydrogenase WbpB (G3XD23), aspartate semialdehyde dehydrogenase (Q51344), 2-amino-4-hydroxy-6-hydroxymethylhydropteridine pyrophosphokinase (Q9HV71), 3-deoxy-D-manno-octulosonic-acid transferase (Q9HUH7), glycyl-tRNA synthetase alpha chain (Q9I7B7), riboflavin kinase/FAD synthase (Q9HVM3), aconitate hydratase 2 (Q9I2V5), probable glycosyltransferase WbpH (G3XD85) and UDP-3-O-[3-hydroxy-lauroyl] glucosamine N-acyltransferase (Q9HXY6). For druggability and pocketome analysis crystal and homology structures of these proteins were retrieved and developed. A sequence-based search was performed in different databases (ChEMBL, Drug Bank, PubChem and *Pseudomonas* database) for the availability of reported ligands and tested drugs for the screened targets. These predicted targets may provide a basis for the development of reliable antibacterial preventive and therapeutic options against *P. aeruginosa*.

Keywords *Pseudomonas aeruginosa* · Reverse vaccinology · Subtractive proteomics · Vaccine candidates · Chimeric vaccine · Druggable targets

Introduction

World Health Organization has classified *P. aeruginosa*, as a top priority critical pathogen against which there is a dire need to develop novel antimicrobial therapeutics (WHO 2021). The wide range of hosts (i.e. humans, plants and animals), ubiquitous environmental distribution, complex genome (carrying an enormous arsenal of virulence and resistance genes), and intrinsic antibiotic resistance potential make it one of the most unique and widespread pathogen

✉ Saadia Andleeb
saadia.andleeb@asab.nust.edu.pk;
Saadiamarwat@yahoo.com

¹ Department of Industrial Biotechnology, Atta-ur-Rahman School of Applied Biosciences (ASAB), National University of Science & Technology (NUST), Sector H-12, Islamabad 44000, Pakistan

² Institute of Basic Medical Sciences, Khyber Medical University, Peshawar, Pakistan

worldwide (Wang 2017; Lister et al. 2009). *P. aeruginosa* is capable of causing both acute as well as chronic infections and is reported to harbor resistance to almost all available classes of antibiotics through chromosomal mutations and horizontal gene transfer (Rizvi 2015).

The arsenal of virulence factors comes in handy at the sites of infection where it inflicts serious injuries to the host causing different types of cytotoxic effects leading to apoptosis, necrosis, ultimately resulting in immune evasion and immune modulation, hence paving the way for infection establishment (Rashid 2017; Gellatly and Hancock 2013; Yang 2017). Hence, the infections caused by this notorious pathogen are often difficult to eradicate despite intense antibiotic combination treatments (Lister et al. 2009).

Lack of effective treatment options has fueled many studies for exploring preventive approaches. Enormous efforts have been put in the past few decades to develop effective vaccine candidates against *P. aeruginosa* (Priebe and Goldberg 2014). Several vaccine candidates have been tested (François 2012; Westritschnig 2014), however, no candidate has been licensed yet for clinical trials (Vincent 2014). *P. aeruginosa* possesses abundant outer membrane proteins of which the function and expression of a large number is still unknown (Montor 2009). The plasticity of *P. aeruginosa* genome is another hurdle in the selection of conserved antigenic vaccine candidates (Mosquera-Rendón 2016; Pennington 1975). The identification of conserved novel immunogenic candidates can potentially contribute to the development of effective vaccines.

“Reverse vaccinology (RV)” is a much better option to prepare successful vaccines. This approach allows the identification of complete potential protein’s repertoire of an organism using genomic and proteomic information (Delany et al. 2013). RV has been found cost-effective with high accuracy compared to the traditional vaccinology approaches and has been successfully applied in designing vaccines for serogroup *B. meningococcal* infection (Giuliani 2006). This strategy includes a detailed analysis of the genome/proteome of pathogen i.e. essential and regulatory genes of the targeted pathogen. Screening of virulence-associated proteins involved in establishment and progression of infections, and prioritized protein’s interactions with host proteome (Rashid 2017). The prospective candidate proteins are screened for surface exposure, immunogenic potential, abundance, number of transmembrane helices and involvement in virulence (Brunham et al. 1993). *P. aeruginosa* possesses abundant outer membrane proteins (Montor 2009), of which the function and expression of a large number are still unknown.

The basic theme of our study was the rigorous screening of *P. aeruginosa* reference strain PAO1 proteome for the identification of non-human homologous exoproteins which

were subsequently screened for highly conserved novel epitopes with the ability to bind with maximum number of MHC-I and MHC-II class molecules. The epitopes were then joined together, and designed chimera was further tested for stability and ability to bind to TLR4. Finally immune simulation was performed and the multi-epitope vaccine coding gene was in silico cloned in pET28a vector for expression and purification of the designed vaccine for in vivo studies.

Genome projects and the development of bioinformatics tools have resulted in a paradigm shift in drug development research (Hopkins and Groom 2002). Conventional drug development strategies based on hunches and trial and error approaches are costly time consuming and inefficient (Adams and Brantner 2006; Craik 2013; Hedgecoe and Martin 2003). Drug Development in the age of genomics is more focused on the assessment of genes and their functions concerning particular diseases or conditions of interest. This approach works more efficiently and in a more guided manner limiting the scope of hunches and trials. The availability of a huge amount of genomic data has provided an opportunity to study these genes of interest and to use computational tools for determining their therapeutic potential before going into the lab (Lenz et al. 2000; Roses 2000). The “druggable genome” indicates the genes which hold therapeutic potential and can serve as targets for drugs (Hopkins and Groom 2002). These drug targets are involved in metabolic pathways of the organism, and one way or the other may contribute to pathophysiology. While druggability of a gene/protein is a more topological phenomenon that denotes the ability of the protein to have binding pockets for small drug-like molecules in its three-dimensional structure (Keller et al. 2006). These binding pockets provide the drug molecules with a place to bind and influence the activity of the target. A schematic overview of our adopted strategy is represented in Fig. 1.

Materials and Methods

Retrieval of *P. aeruginosa* Proteome and Removal of Paralogous Sequences

Primary genomic and proteomic data of *P. aeruginosa* reference strain PAO1 (RefSeq NC_002516.2, Proteome; UP000002438) was retrieved from NCBI GenBank (Bilofsky and Christian 1988) and UniProt database (Bairoch and Apweiler 2000). The pan-proteome of PAO1 was subjected to a stepwise identification of vital virulent proteins (Rashid 2017). The complete proteome was checked for redundancy through CD-HIT (Li and Godzik 2006). A schematic overview of our adopted strategy is represented in Fig. 1.

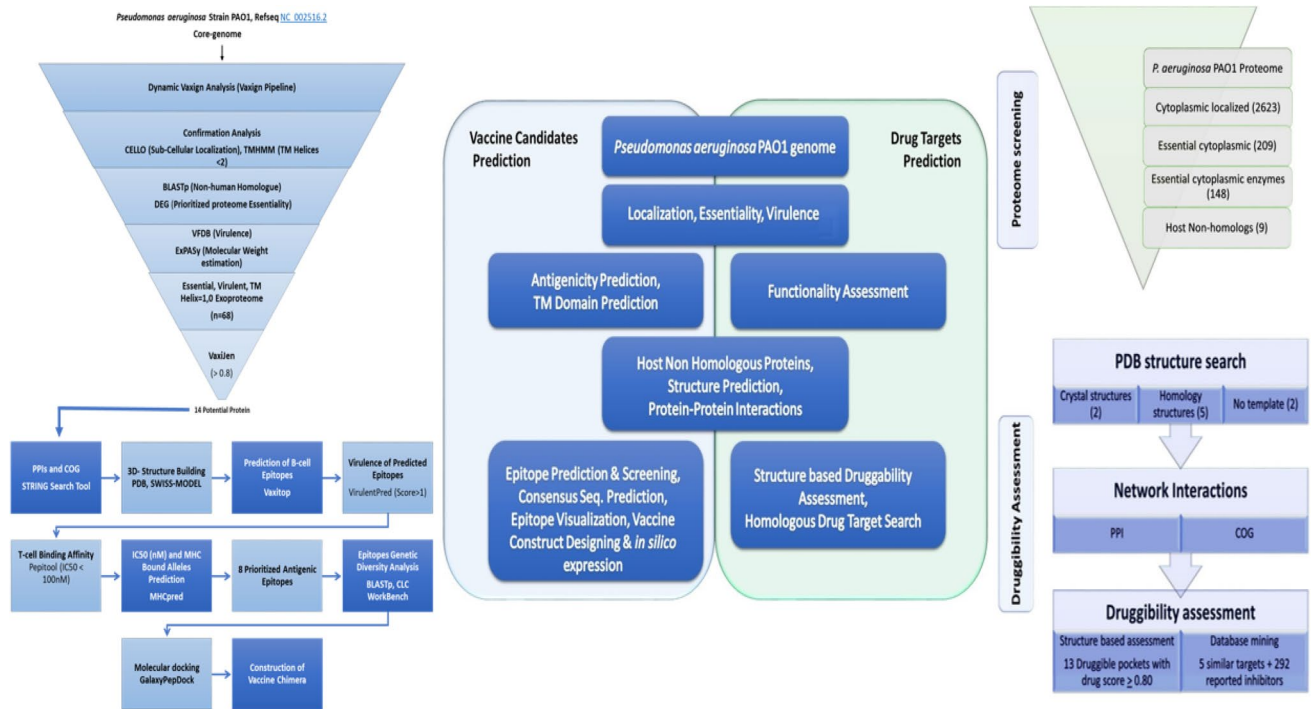


Fig. 1 The schematic workflow overview: The *Pseudomonas aeruginosa* reference strain PAO1 was subjected to reductive genome screening for the prediction of potential vaccine candidates and identification of novel drug targets

Subcellular Localization and Host Non-homologous Proteins Prediction by Vaxign Pipeline

Subcellular localization prediction is an important step for the selection of suitable vaccine and drug target candidates out of an infectious agent proteome. Proteome data was screened using the Vaxign pipeline (<http://www.violinet.org/vaxign/>) (Xiang and He 2009). Dynamic Vaxign analysis was performed instead of using the Vaxign query program due to the unavailability of *P. aeruginosa* genome in the list of pre-computed genome groups. Vaxign includes PSORTb for precise bacterial protein localization prediction. Due to the limitations of PSORTb in sorting proteins of multiple localization, CELLO v2.5 was also used in parallel (Yu et al. 2004).

The topology of secreted and surface-exposed proteins was verified using TMHMM v2.0 (Tusnady and Simon 2001). Proteins with more than one transmembrane helices were removed from the shortlisted pool. The Vaxign pipeline calculated adhesins using SPAAN (Sachdeva 2004) with up to 89% sensitivity and 100% specificity. These cell-surface components are also considered good vaccine targets.

Host homologous proteins could elicit autoimmune responses and thus removal of such homologs is essential. From a pharmacokinetic perspective, the ideal vaccine or drug target must be host non-homolog. Host homology/similarity analysis was done via Vaxign customized OrthoMCL

(Li et al. 2003) to calculate homology between screened/filtered proteins and protein of specific host (e.g., human). In addition to this, the Basic Local Alignment Search Tool for proteins (BLASTp) available at NCBI server (Altschul 1990) was used to filter out proteins with similarity to human proteins. Human homologs were removed from the prioritized protein pool. The cut-off criteria were set at 0.005 E-value and <40% sequence identity with at least 50% query coverage. Proteins having sequence similarity of $\geq 40\%$ were removed from the data set (Calabriso 2016). Apart from the human homologs, orthologs search across *Pseudomonas* genus was conducted using organism-specific BLASTp search and “The *Pseudomonas* genome Database” (Winsor et al. 2011).

Screening of Pathogen Essential, Virulent and Antigenic Proteins

Database of Essential Genes (DEG) version 10.4 (Luo 2014) was used to get information about the essentiality status of the shortlisted proteins as a potential vaccine candidate. Selected protein sequences were subjected to DEG limited BLAST run.

Proteins involved in metabolism can also serve as ideal drug targets as inhibiting such proteins could greatly hinder the pathogen’s survival. For this purpose, a separate DEG run was employed for screening essential proteins as a

potential druggable target. The screened pool was subjected to KEGG pathway analysis for shortlisting proteins involved in the metabolism (Weigt 2018). Enzymes were selected from the pool and were again subjected to host homology evaluation.

The pool of essential exoproteome and secretome was evaluated for their potential role in virulence by using the Virulence Factor Database (VFDB) (Chen et al. 2012) and SVM-based tool, Virulentpred (Garg and Gupta 2008). Antigenicity scores of the filtered proteins were estimated by Vaxijen v2.0 (Doytchinova and Flower 2007) (threshold = 0.4). Proteins with high antigenic scores were prioritized as good vaccine candidates.

Molecular Weight Estimation of Selected Vaccine Candidate Proteins

The ease with which a protein could be expressed, extracted and purified is important in the selection of potential proteins to be used for the development of peptide vaccines. For the analysis of the molecular weights of the proteins, their accession numbers were given to ExPASy (Bioinformatics resource portal) compute pI/Mw online tool (Gasteiger 2005). Proteins having a molecular weight of ≤ 110 kDa were selected while the rest were excluded (Table 1).

Prioritized Protein's 3-D Model Building and Druggability Assessment

The presence of a 3D protein structure was checked in Protein Data Bank (PDB) (Berman et al. 2006) for all proteins. SWISS-MODEL (Arnold 2006) and RaptorX (Källberg 2012) were used for homology modeling of proteins lacking crystal structures. For each protein, a suitable structural template was selected and 3-D models were generated.

Druggability was analyzed for each of the selected proteins for predicting the quality and number of drug binding pockets using proteins' tertiary structures. DoGSiteScorer

online server (<https://proteins.plus/>) was used for this purpose (Volkamer 2012). DoGSiteScorer uses a grid-based method with a Gaussian filter for the estimation of prospective binding pockets. This server uses descriptive values of volume, hydrophobicity and shape (enclosure) for determining the score ranged between 0 and 1 (1 as highest druggability potential) for the respective binding pocket. The retrieved crystal structures and predicted homology structures were submitted to the server and resultant scores and details of the predictions were saved. Drug binding pockets with estimated drug score > 0.50 were selected. In case of multiple pockets having a higher score than the cut off top five pockets were selected.

Protein–Protein Interactions (PPIs) and Clusters of Orthologous Group of Proteins (COGs) Analysis

Proteins interact with each other and constitute a complex network, where different proteins contribute synergistically as well as antagonistically towards common functions/pathways. STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) database (Szklarczyk 2011) was used which provides information about PPIs and COGs based on sequence identity. A chimeric vaccine has the potential to affect the expression of all other interacting proteins in the bacterial cell and thus reduces the pathogenesis. The intra specie protein–protein interactomes were compiled in a datasheet for analysis.

Antigenic Epitopes Prediction

Vaxitop online server was used to predict continuous B-cell Epitopes from selected proteins using default threshold values (He et al. 2010). The predicted peptides were reexamined for topology via TMHMM v2.0 (Tusnady and Simon 2001), and antigenicity scores using Vaxijen (Doytchinova and Flower 2007) (threshold value > 0.4). The peptides with antigenicity scores ≥ 0.8 were considered more antigenic

Table 1 Characteristics of finalized vaccine candidate proteins

Protein accession	Gene name	Protein product	Localization		Adhesins probability	Antigenicity score	Essentiality	Virulence
			PSORTb	CELLO				
P13794	oprF	Outer membrane porin F	Outer Membrane	Outer Membrane	0.553	0.8306	✓	✓
P32722	oprD	Porin D	Outer Membrane	Outer Membrane	0.564	0.8658	✓	✓
Q9I456	PA1288	Probable outer membrane protein	Outer Membrane	Outer Membrane	0.641	0.8067	✓	✓
G3XDA9	pra	Protein activator	Extracellular	Extracellular	0.692	0.8073	✓	✓
Q9HUT9	PA4874	Uncharacterized protein	Periplasmic	Periplasmic	0.636	0.858	✓	✓

and thus selected. The screening of epitopes was performed based on the least P-value and highest antigenicity score. The surface-exposed B-cell epitopes with the best antigenicity scores were evaluated for their affinity for T-cells using Tepitool server (Paul et al. 2016). Epitopes were studied for their interactions with MHC class molecules including DRB1*0101 as it is the commonest bound allele. MHCpred was used for calculation of antigenicity and IC₅₀ value (Guan 2003). Extracted prioritized peptide out of respective virulent protein candidates was again subjected to VirulentPred (Garg and Gupta 2008) (threshold value ≥ 0.5). The topology of the selected epitopes was visualized using Pepitope server (Mayrose 2007).

Evaluation of Genetic Diversity

Epitopic sequences were checked for conservation among *P. aeruginosa* taxonomic group (Taxid: 136841) and other *Pseudomonas* spp. via sequence alignment. Multiple sequence alignment was performed using CLC workbench V7 and consensus sequences were obtained in case of sequence variation for each of the respective finalized epitopes. The immunogenic potential of both 100% conserved and derived consensus peptide sequence was again determined using Vaxijen server (Doytchinova and Flower 2007).

Designing of Polypeptide Vaccine Construct

Assembly of the B-Cell Epitopes for Designing of the Vaccine

The vaccine sequence was built by adding the adjuvant at the amino-terminus followed by the top B cell epitopes for the selected proteins. The adjuvant selected was the Cholera toxin subunit B (CTB) linked with the first B cell epitope via an EAAAK linker while the rest of the epitopes were linked together by incorporating the GPGPG linkers. At the carboxy-terminal, a six His Tag was attached for protein identification and purification in later stages (Fig. 2).

Prediction of the Allergenicity and Solubility

To predict the allergenicity of the vaccine sequence, AllgPred v.2.0 (Saha and Raghava 2006) and AllerTOP v.2.0 (Dimitrov 2014) servers were utilized. These servers work on specific algorithms to forecast the allergenicity of the amino acid sequence. VaxiJen v2.0 server was employed for the prediction of the antigenicity of the vaccine protein whereas, the solubility was assessed through the Proso II (Smialowski et al. 2012) and Protein-sol servers (Hebditch 2017).

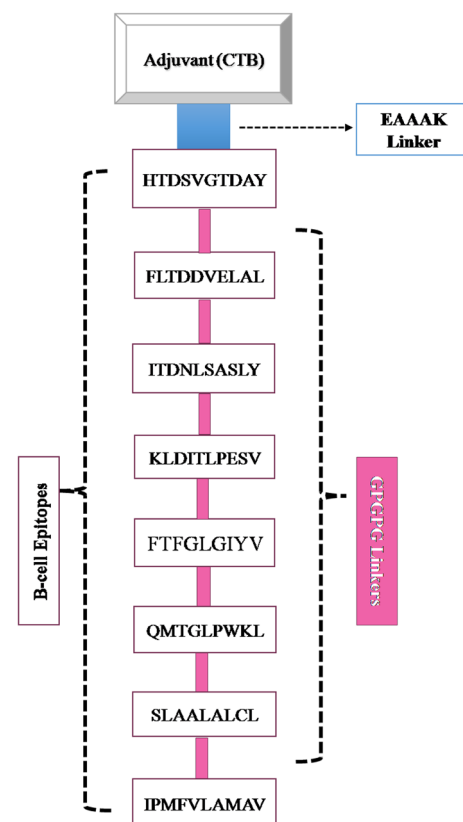


Fig. 2 Representation of the designed vaccine model. The amino terminus starts with the adjuvant attached to the first B-cell epitope with an EAAAK linker. The rest of the epitopes are conjoined together via GPGPG linkers. At the c-terminus, a 6×His Tag is attached for identification purpose

Evaluation of the Physicochemical Properties

For the functional characterization of the protein, ExPASy ProtParam (Gasteiger 2005) tool was used which explored the vaccine sequence on a number of physicochemical parameters such as its molecular weight, pI, aliphatic index, GRAVY value, instability index and estimated half-lives. The PSIPRED v3.3 was utilized for the investigation of the alpha helix, beta-sheet and coil structure of the vaccine (Buchan and Jones 2019).

Tertiary Structure Prediction and Validation

For the 3D modeling of the vaccine protein, the amino acid sequence was subjected to I-TASSER server that builds a 3D structure on the basis of the degree of similarity between the target protein and the available template structures from the Protein Data Bank (Yang and Zhang 2015). To validate the generated 3D structure, its refinement was carried out to enhance the stability of the predicted model using GalaxyRefine (Heo et al. 2013). The refined protein was then

validated through the Ramachandran plot assessment via RAMPAGE server (Lovell 2003).

Molecular Docking of the Vaccine Protein with TLR4

HADDOCK 2.2 web server was employed for predicting the binding affinity and favorable interactions between the designed vaccine and the human TLR4 (Vries et al. 2010). For the docking analysis, TLR4 3D model was retrieved from the PDB and active and passive residues were predicted for both the complexes via CPORT⁹⁴. iMODs⁹⁵ server was used for the prediction of the normal mode analysis. This server is a dynamic tool that simulates the protein model with reference to its internal normal modes and aids in determining the flexibility and mobility of the complexes via a number of factors including deformability, eigenvalues, B-factors, and covariance. The structural dynamics of the protein–protein complex were thus investigated.

Codon Optimization

JCat⁹⁶ server is a codon optimization tool that was utilized for the reverse translation of the vaccine sequence. *E. coli* (K12) was selected as the appropriate host for the codon optimization as it differs significantly from the human in terms of codon usage. Parameters such as Rho-independent transcription termination, prokaryote ribosome binding site and cleavage sites of restriction enzyme BglIII and ApaI were not selected for the final optimization of the vaccine sequence.

Analysis of the Immune Capacity of the Designed Vaccine

The peptide vaccine was subjected to C-ImmSim web server for the simulations of immune responses. The server incorporates the position-specific matrix for the generation of immune connections. The default parameters were employed for this procedure with three subsequent doses of injections⁹⁷.

Results

Proteome Subtraction

PAO1 proteome comprises of a total 5562 proteins. Initial filtration using the CD-HIT online server resulted in 5548 proteins with the removal of 14 paralog sequences from the pan-proteome. Elimination of paralogous sequences resulted in a non-redundant set of proteins which was used for further analysis. The resultant proteome was subjected

to dynamic Vaxign analysis which estimated 169 periplasmic, 169 outer-membranous, 65 extracellular and 1277 proteins localized on the cytoplasmic membrane. A total of 4628 proteins were having < 2 number of helices (Supplementary Table S1). Proteins with more number of helices get anchored onto the cell surface and hence are not accessible to the host immune system. Moreover, 1863 proteins were found essential and 1032 as virulent. The decision tree (DT) method was used to filter the proteome pool based on the predefined threshold values thus excluding the undesired proteins.

For druggable protein target identification, 2623 cytoplasmic proteins were selected. These selected proteins were subjected to essentiality checks via DEG and the number of selected proteins was reduced to 209. Using KEGG output 148 essential cytoplasmic enzymes were obtained. The essential and virulent exoproteins were further used for analysis.

Following localization and virulence prediction for selection of putative vaccine candidates' rigorous screening of the resultant protein pool involved; removal of 1224 human homologous proteins (identified by Vaxign pipeline), selection of surface-exposed proteins of < 2 transmembrane helices and > 0.5 adhesins probability. This yielded a hand full of 68 initially prioritized proteins (Supplementary Table S2, Fig. S1 (a, b)). Based on prioritizing parameters particularly antigenicity scores estimation by Vaxijen v2.0 online server (threshold value = 0.8), 14 proteins were selected for further analysis (Supplementary Table S3). The relative position of the genes encoding the 14 selected proteins on the PAO1 circular genome is shown in Supplementary Fig. S2.

In addition to this, organism-specific (*Homo sapiens*) BLASTp search of selected proteins revealed 10 proteins with no significant homology with human proteins (Supplementary Table S4). OprF (P13794) showed a 35% identity with only 20% query coverage predicting it non-homologous to humans. Three proteins PA0833 (Q9I5A7), PA1471 (Q9I3P1) and OprG (Q9HWW1) were removed from the pool because of their higher identity, query coverage scores and probability of evoking autoimmune responses in the host.

Apart from homologs search, orthologous sequence search using BLASTp and PATRIC v3.5.23 identified putative conserved domains for all proteins in other genera except for proteins PA1956 (Q9I2E8), pra (G3XDA9) and PA0724 (Q9I5K4) (Supplementary Tables S5, S6).

For druggable targets' search, the obtained 148 (KEGG output) essential cytoplasmic enzymes were subjected to host (human) homolog exclusion using BLAST against *Homo sapiens* proteomic data. Host homolog exclusion reduced the selection pool to 9 proteins (Supplementary Table S7).

Molecular Weight Estimation

Identified putative vaccine candidate proteins were further analyzed for their respective molecular weights. Only one protein PA4541 (Q9HVN6) had a molecular weight of 139 KDa, higher than the cutoff ≤ 110 kDa and hence was removed from the list of prioritized proteins (Supplementary Table S8). No such exclusion was performed for potential druggable targets.

Based on antigenicity scores estimation by VaxiJen v2.0 online server (threshold value = 0.8) and molecular weights (cutoff ≤ 110 kDa), 5 proteins were shortlisted as putative vaccine candidates suitable for screening and prediction of best immunogenic epitopes (Table 1).

Three-D Structure Prediction and Refinement

Prioritized proteins were checked for crystalline structure availability in Protein Data Bank. Only two potential vaccine candidate proteins (OprF, and OprD) have had available structures in both organism-specific and general PDB search. Remaining three proteins, PA1288, pra and PA4874 had partially resolved structures depending on the homology search

in PDB. Structures of these proteins were modeled using SWISS-MODEL and RaptorX. Only the RaptorX models were used for the antigenic region's visualization (Fig. 3a). For putative druggable targets, the initial organism-specific search (limited to *P. aeruginosa*) yielded two crystal structures for the proteins with 100% sequence identity and (PDB ID: 5BNT_C) and lpxD (PDB ID: 3PMO_A). Homology structures were predicted for proteins wbpB, folk, glyQ, ribF, acnB using the Swiss-Model server. Two proteins waaA, wbpH did not have any crystalline structure and also lacked any suitable template for homology prediction; hence they were not included in the druggability assessment study. The 3D structures presented an improved prospect for comprehension of the relationship of the studied proteins within the biological system (Fig. 3b).

PPIs and COGs Analysis

STRING network prediction of five screened proteins revealed intra-species interaction between different proteins. **OprF (OMP)** is a member of the cell outer-membrane complex with involvement in several biological processes, transport of organic substances, protein and ion transport.

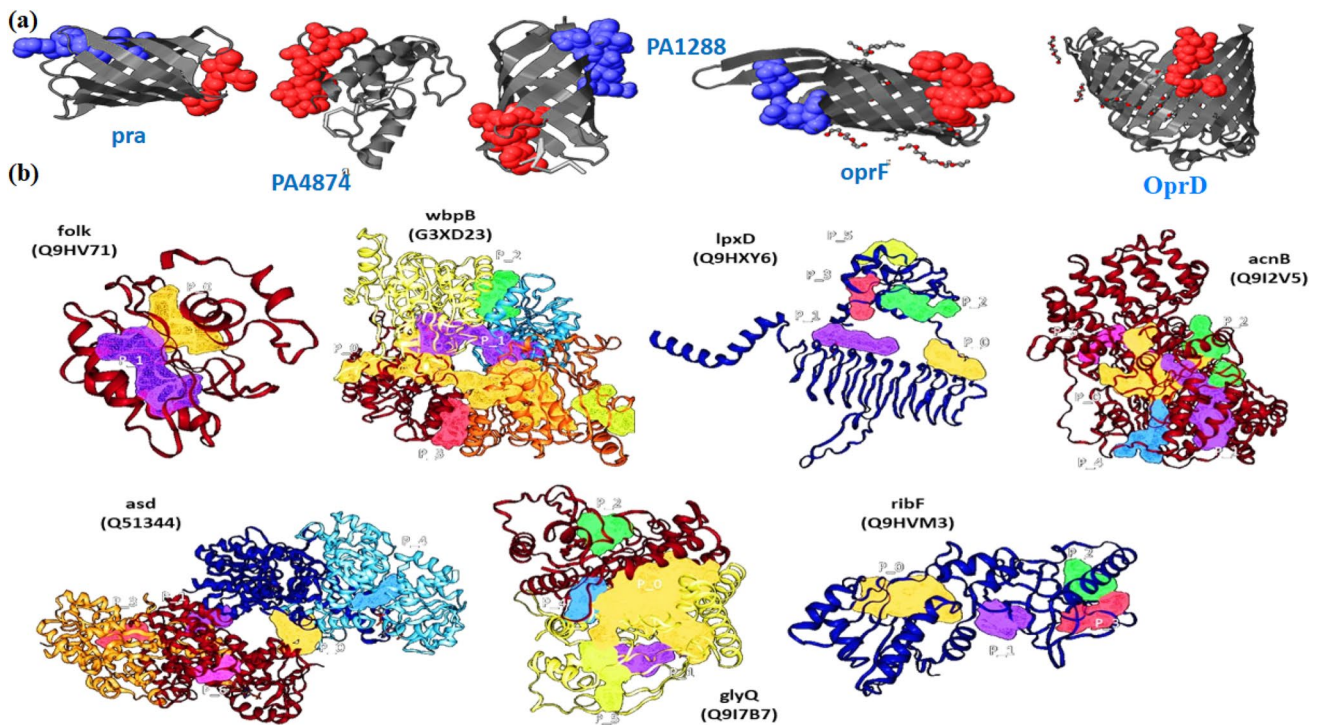


Fig. 3 Three-dimensional structure of the predicted vaccine candidate proteins along with the topological view of corresponding antigenic peptides (on top) and druggable target proteins. **a** illustrates the RaptorX modeled three-dimensional structures of proteins with their corresponding surface-exposed antigenic epitopes. Grey color is indicating the protein structure while the respective epitopes are shown

in cartoon form. Proteins with more than one antigenic epitope are shown in pink and blue colors. The images were visualized by using the Pepitope server (URL: <http://pepitope.tau.ac.il/>). Best scoring drug-binding pockets **b**: Pockets with a drug score of > 0.5 were selected for each of the screened druggable proteins

Being a member of OM-channel superfamily other predicted functional porin partners include OprD, OprP, OprB, OprE, OprG, OprI, OprH and tolB. There is also a functional link of OprF protein with RNA polymerase sigma factor SigX.

Direct interaction of many multidrug efflux outer membrane porins, transcriptional regulators and operon repressors is also observed with **OprD protein** (Porins with serine protease activity and specificity for basic amino acids). This includes OprJ; channel-forming component of multidrug efflux pump, OprN, OprM; multidrug ABC transporter and component of MexAB-OprM efflux system, nfxB; confer resistance to quinolones, may negatively regulate genes associated with cell permeability to drugs and mexR; MDR operon repressor. OprD also has involvement in KEGG pathways beta-lactam resistance, transmembrane transport activity and lipid binding. No significant functional enrichment was found for **pra** (protein activator of *P. aeruginosa*) although it was found to have indirect interaction with a range of virulence factors of *P. aeruginosa* including phzH, lasB, lap, and napD. **PA1288** and **PA4874**, being probable outer membrane protein precursors and hypothetical proteins respectively do not have any significant functional role. Individual interaction analysis of these proteins revealed interaction of PA1288 with a long chain fatty-acid-CoA ligase fadD2 protein, eraR protein, acyl-CoA dehydrogenase PA0508 and various hypothetical proteins of unknown functions. The predicted functional partners of PA4874 include sporulation protein SpoVR which is directly involved in spore formation, flagellar hook protein flgK, a unique metal-binding pterin Molybdopterin-binding protein PA3441, and a few hypothetical proteins.

A chimeric vaccine against these five proteins might affect all interacting proteins which will further enhance the efficacy of the vaccine by resulting in the reduction in the pathogenesis of *P. aeruginosa* infections. All protein network interactions are shown in Supplementary Fig. S3(a).

During COG analysis the protein OprF was sorted to COG2885—outer membrane protein and related peptidoglycan associated proteins whereas PA1288 lies in COG2067—predicted long-chain fatty acid transport protein. OprD and PA4874 were sorted into a non-supervised orthologous group. Outer membrane porin precursor OprD was sorted to NOG08180 and conserved hypothetical protein PA4874 to NOG79339. PA4874 has a probable link with phosphate starvation-inducible protein PsiF. Protein activator pra could not be sorted to any orthologous group.

The generated interaction and network map for putative druggable targets had about 19 nodes, 39 edges, with average node degree about 4.11 Supplementary Fig. S3 (b). The average clustering coefficient was 0.677 and PPI enrichment p-value was 3.55×10^{-08} . Proteins wbpB, and, folK, waaA, glyQ, ribF, acnB, wbpH, and lpxD are involved in multiple pathways and perform biological functions ranging

from amino acids, lipid and carbohydrate biosynthesis, LPS biosynthesis, protein translation, cell wall organization and energy generation. Protein **WbpB** and **WbpH** are involved in Biological Processes (GO): O antigen biosynthetic process (GO: 0009243), cell wall organization (GO: 0071555), lipopolysaccharide biosynthetic process (GO: 0009103), protein oligomerization (GO: 0051259) and protein complex biogenesis (GO: 0070271). Protein **asd** is primarily involved in amino acid biosynthesis as threonine biosynthetic process (GO: 0009088), aspartate family amino acid biosynthetic process (GO: 0009067), alpha-amino acid biosynthetic process (GO: 1901607), lysine biosynthetic process via diaminopimelate (GO: 0009089), and branched-chain amino acid biosynthetic process (GO: 0009082). **folK** plays its role in tetrahydrofolate biosynthetic process (GO: 0046654), cellular modified amino acid biosynthetic process (GO: 0042398), and coenzyme biosynthetic process (GO: 0009108). Enzymes **waaA** and **lpxD** both are collectively involved in lipid A biosynthetic process (GO: 0009245), liposaccharide metabolic process (GO: 1903509), carbohydrate biosynthetic process (GO: 0016051), lipid biosynthetic process (GO: 0008610), cellular lipid metabolic process (GO: 0044255), carbohydrate derivative biosynthetic process (GO: 1901137). **glyQ** is involved processes; nucleobase-containing compound metabolic process (GO: 0006139), tRNA aminoacylation for protein translation (GO: 0006418), and had aminoacyl-tRNA ligase activity (GO: 0004812). **ribF** is involved in Riboflavin metabolism (00740). **acnB** is involved in tricarboxylic acid cycle (GO: 0006099), aerobic respiration (GO: 0009060), propionate metabolic process (GO: 0019541) along with mRNA 3'-UTR binding potential (GO: 0003730), and aconitate hydratase activity (GO: 0003994).

Screening of B-Cell and T-Cell Epitopes

Vaxitop online server predicted enormous B-cell epitopes from selected proteins out of which a total of 54 best scoring B-cell epitopes based on least p-value, highest antigenicity score and topologically exposed on the surface of protein were shortlisted (Supplementary Table S9). A list of tested MHC-I and MHC-II alleles are provided in Supplementary Table S10. Selected peptides were subjected to IC₅₀ value analysis for DRB1*0101 allele, several bound alleles by epitopes and T-cell binding affinity (prediction by tepitool server, IC₅₀ value cut off < 100 nM). Virulence prediction of screened peptides revealed all peptides as virulent with good scores. Principally the targeted peptides should elicit both B-cell and T-cell mediated immunity. Hence selection of epitopes was performed based on the highest binding affinity and recognition by both B-cell and T-cell. Finally, 8 peptides with the lowest IC₅₀ values (cut off < 100 nM), highest antigenicity scores, having a binding affinity with a

maximum number of alleles of both classes were selected as shown in Table 2.

These predicted epitopes fulfill all attributes of a potent peptide capable of provoking an immune response. To check the topology of all epitopes as surface-exposed/outside and not folded/hidden within the globular proteins these were visualized on their respective proteins using Pepitope online server (Fig. 3a).

Genetic Diversity Analysis

Potential sequence variations in other strains are a serious issue as these variations can potentially result in a lack of immunogenicity of the predicted vaccine candidate. For evaluating such probable sequence divergence, sequences of potential epitopes were aligned with members of *Pseudomonas aeruginosa* taxonomic group (Taxid: 136841). Multiple sequence alignment of screened epitopes revealed the requirement of only one consensus sequence for a 9-mer pra antigenic epitope QMTGLPWKL at position number 8 from lysine to threonine (Supplementary Fig. S4). Immunogenicity potential of both predicted epitopes and sequence alignment derived consensus peptides was determined using Vaxijen v2.0 and revealed no significant difference between epitope predicted by pipeline and derived consensus peptide.

Alternate Drug Targets Assessment

Structure-Based Druggability Assessment

Out of 9 screened potential targets, 7 having the 3d structure were subjected to druggability assessment using DoGSiteScorer. The online server revealed druggable pockets for each target. The pockets with drug scores > 0.50 were selected as

shown in Fig. 3b. Best pockets were selected based on the generated drug score by the server to be ≥ 0.80 . The list of best pockets is shown in Table 3. The highest scoring pocket was P_2 detected in acnB (Q9I2V5) with a drug score of 0.87. While the second and third high scoring pockets were P_1 glyQ (Q9I7B7) and P_3 wbpB (G3XD23) with drug scores about 0.85 and 0.84 respectively.

Sequence-Based Database Mining for Reported Inhibitors Against Selected Targets

Chemical and drug databases i.e. PubChem, ChEMBL, DrugBank were searched using keyword and sequences (as database allowed) for identification of any potentially reported inhibitor for our screened targets. PubChem and

Table 3 Best scoring drug binding pockets

Protein	Name	Volume (Å ³)	Surface (Å ²)	Drug score	Simple score
Q9I2V5	P_2	693.74	796.85	0.87	0.4
Q9I7B7	P_1	669.89	690.94	0.85	0.4
G3XD23	P_3	706.51	884.98	0.84	0.45
Q9I2V5	P_4	498.1	578.96	0.84	0.24
Q9I7B7	P_2	614.78	638.43	0.84	0.35
G3XD23	P_2	711.81	810.21	0.83	0.45
Q9HVM3	P_3	343.36	732.47	0.82	0.21
G3XD23	P_0	3435.73	3823.06	0.81	0.63
Q9HVM3	P_0	1271.36	1712.96	0.81	0.65
Q9I2V5	P_0	1923.9	1907.15	0.81	0.56
Q9I7B7	P_0	3312.32	2928.38	0.81	0.59
G3XD23	P_1	1954.12	1955.06	0.8	0.63
Q9I2V5	P_1	1193.96	1104.93	0.8	0.55

Table 2 List of prioritized antigenic and virulent epitopes in potential vaccine candidate proteins

Protein accession	Gene name	Protein length (aa)	Localization	B-cell epitope	IC50 value (nm)	MHC I allele count	MHC II allele count	P value	Antigenicity score
P13794	<i>oprF</i>	350	Outer membrane	HTDSVGT-DAY	65.53	4	0	0.000385	1.9872
		350	Outer membrane	FLTDDVELAL	2.09	18	0	0.0042	0.4731
P32722	<i>oprD</i>	443	Outer membrane	ITDNLSASLY	4.73	15	0	0.00000231	0.7422
Q9I456	PA1288	424	Outer membrane	KLDITLPESV	5.28	6	3	0.00166	1.2296
		424	Outer membrane	FTFGLGIYV	55.34	27	1	0.00277	2.1776
G3XDA9	<i>pra</i>	162	Extracellular	QMTGLPWKL	9.23	12	0	0.0125	2.2073
		162	Extracellular	SLAALALCL	3.85	12	0	0.0149	0.9236
Q9HUT9	PA4874	138	Periplasmic	IPMFV-LAMAV	5.06	14	23	0.000385	1.6836

DrugBank search failed to reveal any similar target while ChEMBL revealed 43 different targets with about 1708 inhibitors for our 9 selected screened targets with a variable degree of % identity ranging from 22.55 to 51.3%. The selection was made by excluding all the inhibitors with % identity < 40%. Any similar targets originating in Homo sapiens were also excluded for avoiding any potential side effects. The most identical one was against folK (Q9HV71) reported in *E. coli* (ChEMBL3217379) 2-amino-4-hydroxy-6-hydroxymethyl-dihydropteridine pyrophosphokinase (P26281) having 74 reported compounds and 51.3% identity. The most similar reported targets are described in Table 4.

Designing of Polypeptide Vaccine Construct

Prediction of the Allergenicity and Solubility

The protein sequence of the vaccine was found to be non-allergenic in nature as predicted by the AllgPred server. The antigenicity of the designed vaccine was estimated to be 0.9193 exhibiting a potential immunogenic response. Upon solubility analysis, Proso II indicated the construct as

soluble molecule. Moreover, the designed vaccine is soluble as per its charge distribution, hydrophobicity and stability.

Physicochemical Properties of the Designed Vaccine

ExPASy server was employed for determining various physicochemical properties of the designed vaccine. The molecular weight of the designed sequence is 26.5 kDa. The pI of the vaccine was predicted to be 6.35 which indicates its acidic nature. The aliphatic index was recorded to be 83.98 along with the GRAVY value of – 0.092. The vaccine was forecasted to have an estimated half-life of one hour within the mammalian reticulocytes in vitro and greater than ten hours in *E. coli* in vivo. All these parameters show the thermostable nature of the designed vaccine.

Multi-epitope Vaccine Structure Prediction and Validation

PSIPRED v3.3 categorized the secondary structure of the designed vaccine as; 28% alpha-helix, 18% β sheets and, 52% random coil (Fig. 4a). I-TASSER was used for predicting the tertiary structure and first phase of refinement was carried out by the ModRefiner which was preceded

Table 4 List of reported inhibitors available for similar targets searched based on sequence homology

Query Sequence	ChEMBL ID	Preferred name	UniProt accession	Target type	Organism	Compounds	% Identity	BLAST score	E-value
Q9HV71	CHEMBL3217379	2-amino-4-hydroxy-6-hydroxymethyl-dihydropteridine pyrophosphokinase	P26281	Single protein	<i>Escherichia coli</i> (strain K12)	74	51.3	164	3.00E–41
Q9I7B7	CHEMBL3086	Acetyl-Coenzyme A carboxylase	Q5SWU9	Single protein	<i>Mus musculus</i>	2	42.5	28.9	1.80E+00
Q9I7B7	CHEMBL2397	Acetyl-CoA carboxylase 1	P11497	Single protein	<i>Rattus norvegicus</i>	150	42.5	28.9	1.90E+00
G3XD85	CHEMBL3774301	Cysteine proteinase 1	Q01957	Single protein	<i>Entamoeba histolytica</i>	3	42.11	28.9	2.20E+00
G3XD85	CHEMBL3774301	Cysteine proteinase 1	Q01957	SINGLE PROTEIN	<i>Entamoeba histolytica</i>	3	42.11	28.9	2.20E+00
Q9HV71	CHEMBL3559681	2-amino-4-hydroxy-6-hydroxymethyl-dihydropteridine pyrophosphokinase	X5EH84	SINGLE PROTEIN	<i>Staphylococcus aureus</i>	60	41.55	117	5.00E–27

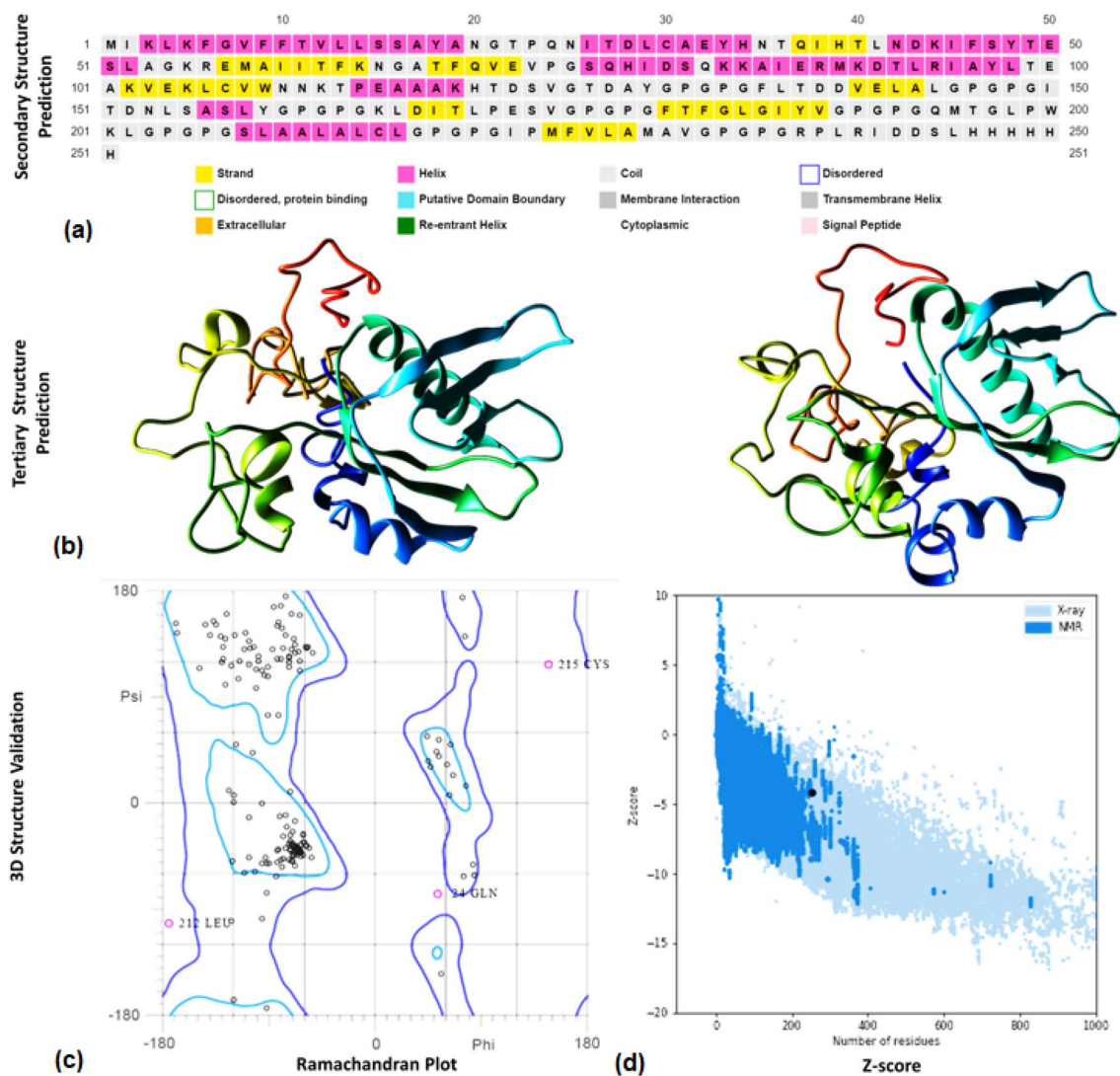


Fig. 4 Proposed multi-epitope vaccine structure and its validation, **a** elements of secondary structure **b** the 3D model of the designed vaccine obtained through I-TASSER based on homology modelling and refined structure of vaccine protein obtained via Galaxy Refine after

its refinement on ModRefiner **c** Ramachandran plot showing 86.7% residues (inside blue circle) of the protein in the favored region **d** ProSA-Web analysis showing the z-score value of -4.13

by the GalaxyRefine treatment for the effective enhancement of the tertiary structure of the vaccine (Fig. 4b). The Ramachandran plot predicted 86.7% residues of the protein in the favored region whereas 8.4% residues in the allowed and 4.8% residues in the outlier area (Fig. 4c). The quality of the predicted model was assessed via ProSA-Web server, which reported the Z-score of -4.13 (Fig. 4d).

Molecular Docking and Normal Mode Analysis

To analyze the binding affinity and pattern of the vaccine with the host TLR4, molecular docking was carried out by using the HADDOCK 2.2 server (Fig. 5). For the analysis

of the flexibility of the docked vaccine with TLR4, normal mode simulation gave suitable results. iMODs predicted the mobility of the complex by the orientation of arrows that were directed towards each other. The B-factor value was equivalent to RMS. The Eigenvalue of the complex was estimated as $7.045871e^{-05}$. The covariance of the vaccine-TLR4 complex was exhibited by different colors mainly red, white and blue, whereas the variance was found to be inversely linked with the assessed Eigenvalue. The elasticity of the vaccine was analyzed through the generation of an elastic network model where the light and dark shades of each dot represented the varying stiffness of the model.

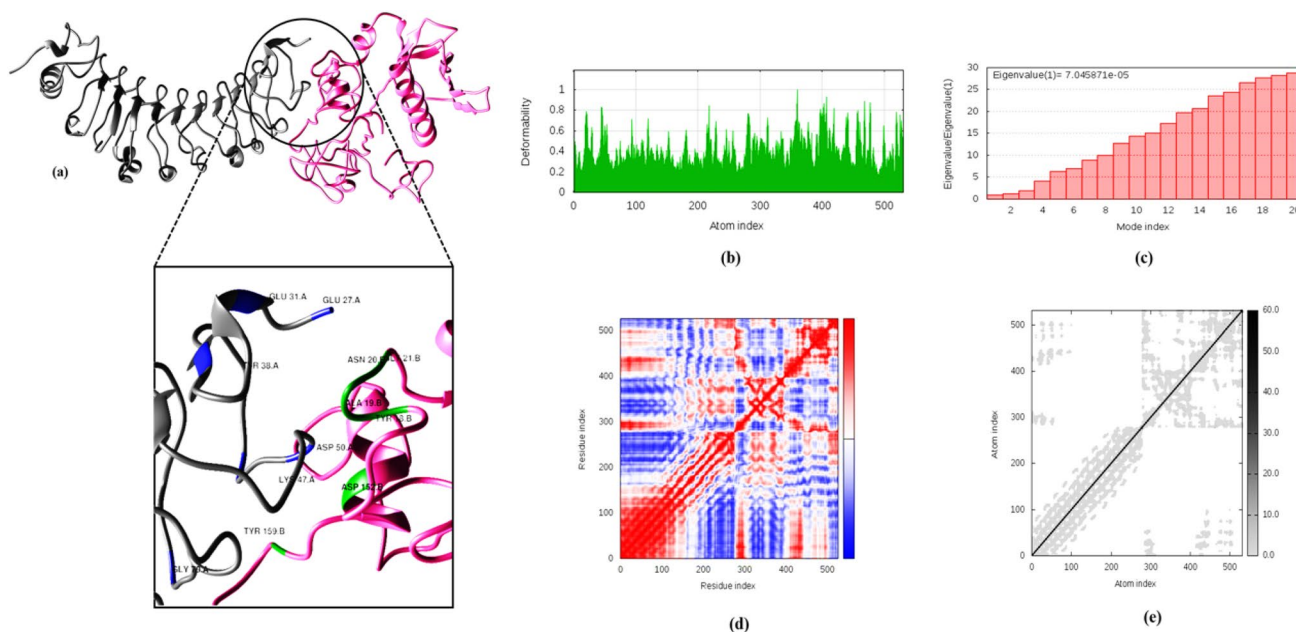


Fig. 5 Molecular docking of designed vaccine with TLR4 and Normal mode analysis. **a** The docking complex shows the binding of the protein with the immune receptor TLR4. The grey molecule represents the vaccine with its active amino acid residues colored as blue whereas the pink molecule represents the TLR4 with green colored

active amino acid residues. **b** The B-factor value of the protein model. **c** The eigenvalue of the protein model. **d** The co-variance of the model generated depicted via three colors: blue, red and white. **e** The elasticity model of the designed vaccine representing its flexibility

Codon Adaptation

For the expression of the designed vaccine in an appropriate host, reverse translation of its sequence was performed. JCat server predicted the codon adaptation index of 0.933 value along with a GC content of 55.77% (Fig. 6). The construct does not carry sites for BglIII and ApaI that indicates its safety for cloning.

Immune Response Against Designed Vaccine

C-ImmSim was employed for the simulation of immune responses of the finalized vaccine construct. The results indicated a consistent rise in the production of primary and secondary immune responses. The antigens along with the IgG antibodies were also found to increase in concentration (Fig. 7).

Discussion

Being one of the top three globally most prevalent nosocomial infectious agents and equipped with virulence and resistance, *P. aeruginosa*'s infections require additional resources and time for patients to recover (Pires, et al. 2018; Moradali et al. 2017; Juan et al. 2017). Countless studies have reported the extent of antibiotic resistance due

to multiple factors ranging from over-prescription, self-medication, and abuse of antibiotics on farms, all ultimately affecting the patients' survival and recovery (Xipell 2017; McLean 2019; Pang 2019; Frieri et al. 2017; Qiao 2018; Karkman 2018). While the lack of vaccines for superbugs was not even an area of interest for academia as well as an industry until recently (Rappuoli et al. 2017). The fact that no preventive vaccine is available for protecting the susceptible patients was the main reason for us to look into the pathogen's genome/proteome for finding solutions to this challenge (Rashid 2020; Baker et al. 2019; Priebe 2003). In this study, we attempted to answer the two key questions associated with any nosocomial infection; prevention and cure in the face of superbugs rampant in healthcare setups. To meet our two objectives effective prevention and cure, we devised an in silico subtractive genomics strategy intending to identify potential vaccine candidates and alternative drug targets against *P. aeruginosa*.

Since the dawn of genomics and the availability of publicly accessible genomic data and analysis tools, many different studies have been done to expedite the vaccine and drug development against noticeable pathogens (Sette and Rappuoli 2010). The most prominent one was the case of the meningococcal group B vaccine which is currently available in markets (Giuliani 2006). Unfortunately, the situation is quite adverse for *P. aeruginosa* in terms of both prevention and treatment (Priebe 2003).

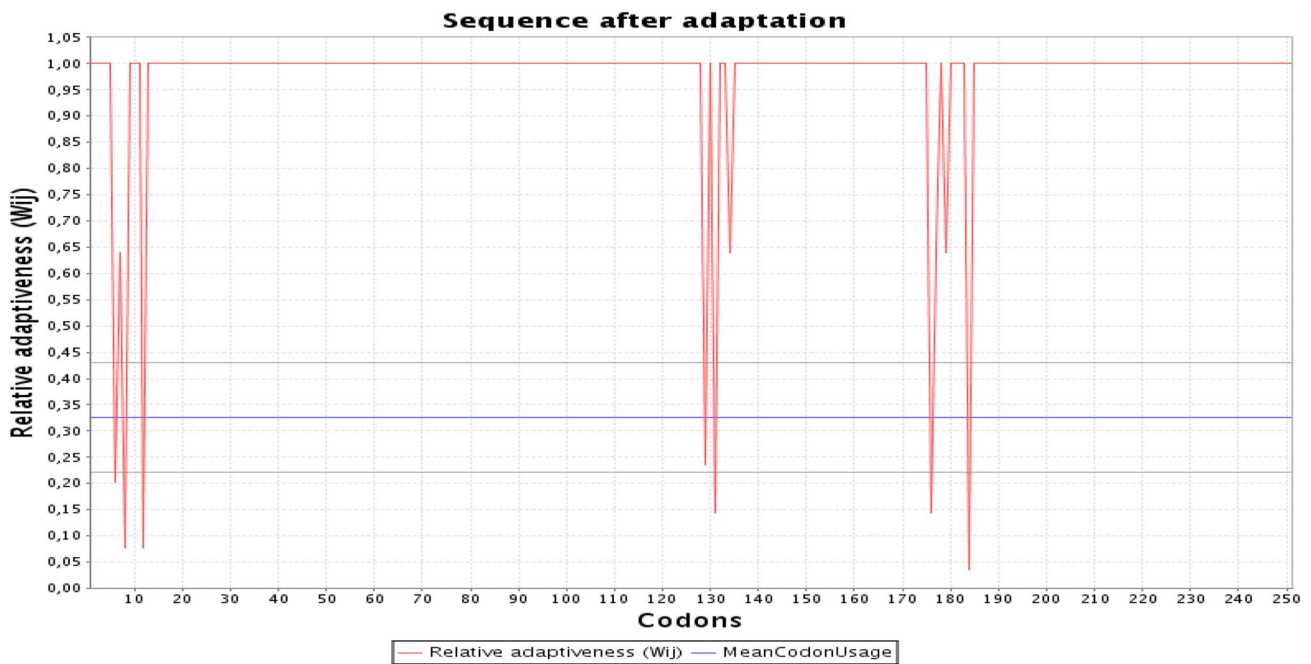


Fig. 6 The codon adaptation index (CAI). The CAI graph indicates that codon optimization of the vaccine with 0.933% and the GC content of the optimized codons (55.77%)

In our previous study, we have presented an overview of all the vaccine development studies and reports (Rashid 2017). In that study, we primarily focused on the essential proteins involved in the virulence and pathogenicity against the host (i.e. a component of lipopolysaccharide, T3SS, proteins involved in adhesion, mobility, and penetration) as it has been done for almost a century now, yet with an updated methodology. In the current study, we have taken into consideration the resistance phenomenon and endeavored to not only focus on the essential proteins that are involved in virulence but also play their due role in antibiotic resistance as well. Two sets of essential proteins were scrutinized based on their computational analysis for preventive and therapeutic potentials.

Dalsass et al. have recently reported a comparative analysis of reverse vaccinology tools and have discussed the advantages and disadvantages of the two methodologies (Dalsass 2019). Reductive screening (about genomic/proteomic data) can be performed using two different algorithms, Decision Tree (DT) or filtering method and Support Vector Machine (SVM) or classification method. In the DT method the proteome is filtered based on the predetermined threshold values and proteins are excluded from the pool thus enriching the desirable proteins, while in SVM, all the desired parameters have assigned quantitative values and based on the accumulative score, proteins are selected with the most desirable features.

We adopted a comprehensive screening strategy employing both SVM and DT algorithms. These two algorithms have their respective benefits and shortcomings hence it was difficult to rely on one screening strategy while overlooking the advantages of the other. To minimize the probability of off-target interactions for our prospective vaccine and drug candidates we adopted a more cautious DT approach for primary screening where the essentiality, localization, functionality and host homology were assessed, and only proteins with desirable qualities were selected. The obtained proteins were essential in their respective metabolic or physiological pathways of the pathogen and lacked any homology with humans. The screened prospective vaccine candidates were either secreted or surface-exposed, involved in pathogenicity and resistance while the potential drug target proteins were enzymes localized within the cytoplasm and involved in essential metabolic pathways. Secretome and exoproteome are considered as appealing vaccine candidates because of their potential role in virulence, adherence, invasion, survival, and proliferation into the host cell (Ranganathan and Garg 2009). In the first round, the Vaxign pipeline screened PAO1 proteome for the identification of the best vaccine candidate proteins. PSORTb is an integral part of the Vaxign pipeline for subcellular localization prediction of proteins which emphasize specificity over sensitivity and do not detect lipoprotein motifs. Hence to overcome such limitations another subcellular localization prediction tool,

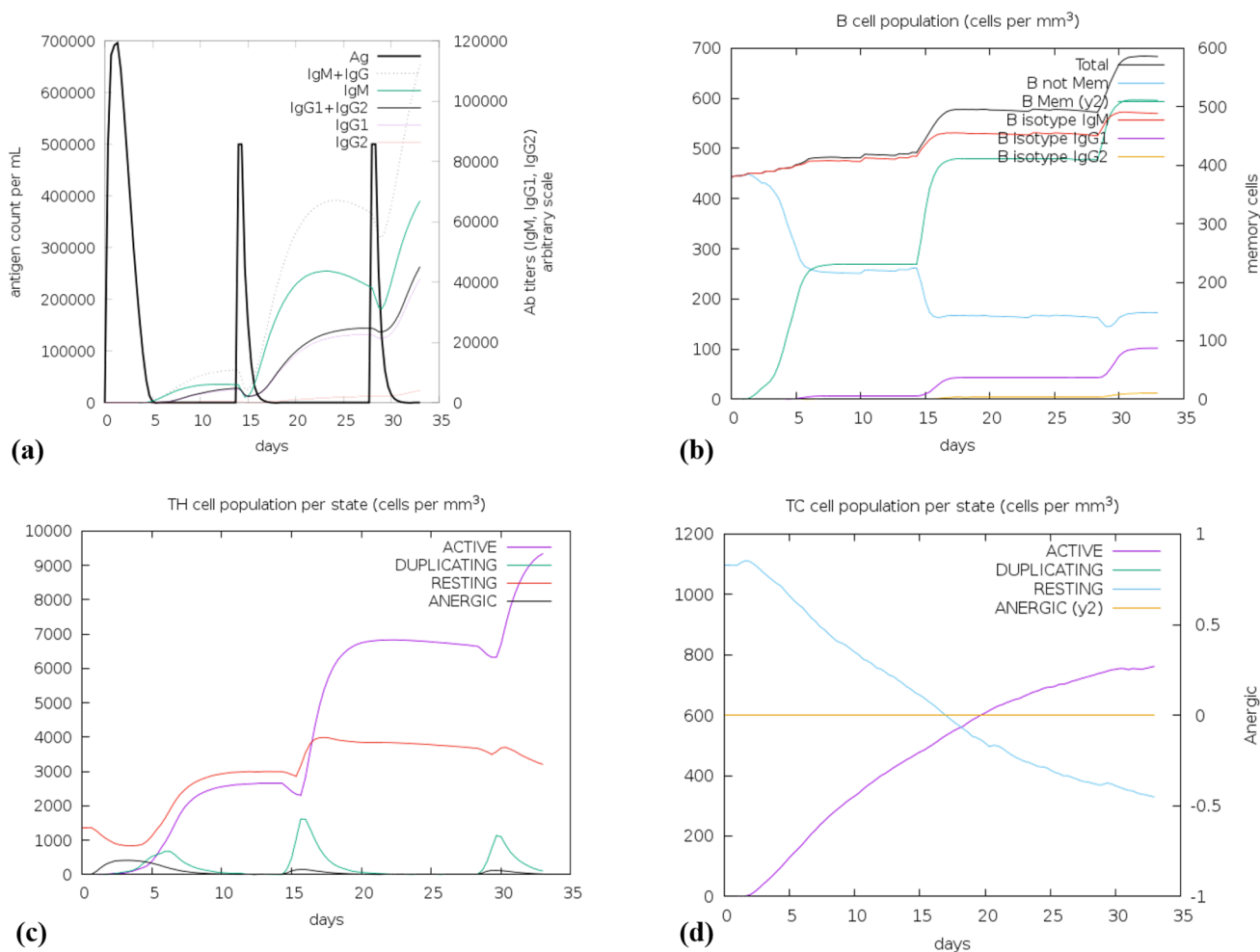


Fig. 7 The in silico immune simulation of the proposed vaccine. **a** Antibodies production, **b** B-cell production, **c** Th cell production, and **d** T-cell production

CELLO v2.5 was synergistically used. Inclusion of complementing results provided by both servers resulted in 7.2% of PAO1 exoproteome. This initially yielded a total of 68 proteins for further analysis.

Although the screened putative protein does not always have to be a virulence factor, selecting proteins with a crucial role in bacterial survival and pathogenesis is extremely useful, because antibodies elicited with such antigens may hinder their functions and neutralize conditions before infections, i.e. use of adhesins as protective antigens may be of great significance through reduction of attachment and colonization by pathogen thus providing herd immunity (Delany et al. 2013). For this reason, proteins directly or indirectly involved in virulence were prioritized. Similarly, the selection of non-human homologs is a prerequisite for therapeutic target identification to avoid any kind of autoimmune response.

Selection of essential, virulent and antigenic exoproteins revealed five outer membrane porin proteins (OprF, OprD,

OprG, OprO, and OprE), six uncharacterized/hypothetical outer-membrane proteins, a protein activator (pra) of *P. aeruginosa*, an adhesion protein (CupB₅) and a probable coat protein A of bacteriophage Pf1 (PA0724). The majority of essential genes and respective proteins play a fundamental role in cellular functions but interestingly a bulk of hypothetical genes of unknown functions are also considered essential and their essentiality is growth-condition dependent (Juhas 2015). Multiple B-cell and T-cell epitopes prediction servers were used to get the best immunogenic epitopes. The selected potential peptide must be capable of binding efficiently with a higher number of MHC class I and class II alleles and thus ensure to cover a major population (Foged et al. 2012). After rigorous screening, our proposed pipeline finally yielded 8 surface-exposed epitopes belonging to five proteins. These prioritized epitopes include components of outer membrane porin proteins (OprF and OprD), protein activator of *P. aeruginosa* (pra), a probable extracellular/outer-membrane protein (PA1288) and uncharacterized/

hypothetical secretory protein components (PA4874). Although the remaining five proteins yielded conserved antigenic B-cell epitopes those were not capable of efficient recognition and binding to cytotoxic T-cell at $IC_{50} < 100$ nM, hence were removed from the list of best epitopes. The best target proteins and their respective epitopic components were reviewed in the literature to get further insight into the prioritized candidates.

Outer membrane proteins (Opr's) of *P. aeruginosa* form porins which are essential structural components of the bacterial cell surface. Components of these porins are attractive for peptide vaccines due to their presence on the cell surface and conserved antigenic domains across wild-type strains of *P. aeruginosa* (Sorichter 2009). In a small scale, clinical trial these candidates have been found promising in eliciting mucosal immunity in the respiratory tract of cystic fibrosis (CF) patients (Fito-Boncompte 2011). Our pipeline screened and prioritized two outer membrane porin proteins (OprF and OprD) as putative candidates for developing efficacious vaccines. OprF porins are small water-filled channels and have a structural role in maintaining cell shape and growth in a low osmolarity medium. Fito-Boncompte et al., have also reported that OprF is a significant virulence factor of *P. aeruginosa* and is involved in biofilm formation in oxygen-deprived conditions. Interaction of OprF with gamma interferons has also resulted in the production of phenazine pyocyanin and lectin PA-IL, both of which are important virulence factors (Lücking 1997).

Several studies have reported elicitation of opsonizing, cross-reactive and protective antibodies in response to immunization with epitopes obtained from OprF and OprI in human and animal models (Knapp 1999). In another study to check the synergistic immunogenicity of OprF and OprI epitopes, the recombinant hybrid vaccines were found protective against *P. aeruginosa* systemic infections in animal models using both active and passive immunization strategies. Mansouri et al. conducted a phase I/III human clinical trial study in 2003 for OprF-OprI epitopes based vaccines (Mansouri 2003).

The loss or reduced OprD porins and decrease in the expression level of these porins in the outer membrane is associated with resistance to carbapenem antibiotics in *P. aeruginosa* infections. Reverse vaccinology based study by Zhaohui et al., on vaccine candidates identification for *Acinetobacter baumannii*, also found OprD superfamily outer membrane porin proteins: CAM88440.1 in AYE and CAM86576.1 in WP_000018327.1 as a potent vaccine target (Ni 2017). More analysis and investigation are required to know the potential of OprD porins as vaccine candidates. Not much is known about the exact functioning of protein activator (Pra) precursor yet. The PPIs revealed interactions with phenazine modifying protein phzH for the biosynthesis of pyocyanin (virulence factor of *P. aeruginosa*),

lap and lasB which are secreted aminopeptidase and assists pathogen in cleaving host elastin, IgG, collagen and several components of complement system proteins. napD protein of periplasmic nitrate reductase (NAP) complex also interacts directly with pra protein activator of both pathogenic and non-pathogenic strains. Hence based on the indirect involvement in virulence and high antigenic potential of pra protein, two epitopic regions QMTGLPWKL at amino acid position number 88–96 and SLAALALCL at position 9–17 have been identified as the best vaccine candidates. Multiple sequence alignment revealed only one residue substitution at position number 8 of the epitope (QMTGLPWKL) from lysine, a basic and polar amino acid with a positive side chain to (QMTGLPWTL) threonine, a hydroxyl-containing polar amino acid with a neutral side chain which did not significantly affect the antigenicity of peptide and thus could be considered as a potent vaccine candidate.

The proposed vaccine candidate's prediction pipeline also predicted two uncharacterized proteins (PA1288 and PA4874). These hypothetical proteins were selected based on their sequence profile, outer membrane localization, immunogenic potency and interactions with different proteins of significant functional assignments. Regardless of lack of functional information expression profiling by microarray of both proteins in response to lung surfactant in CF patients has been reported (Harmer et al. 2013). To the best of our knowledge, the immunogenic potential of these hypothetical proteins has not been explored yet. A few immunogenic hypothetical proteins have also been identified in one of our previous studies.

The best scoring epitopes were then further subjected to the construction of a polypeptide multi-epitope vaccine construct with cholera toxin B (CtxB) subunit as an adjuvant for oral administration. The selection of a suitable adjuvant is a sensitive decision that has a greater impact on the efficacy of any vaccine. Hence we selected CtxB for its stability and efficacy as demonstrated in many recombinant vaccine strategies (Bavandpour et al. 2020). The predicted epitopes were joined together through the use of two types of linkers i.e. EAAAK and GPGPG. These linkers were employed in the construct owing to their potential role in avoiding the immunogenicity that might arise due to the juxtaposition of the epitopes. Moreover, several studies have reported them to be efficient in designing a rational vaccine as they are flexible and frequently utilized (Majid and Andleeb 2019). The designed vaccine construct was evaluated for its chemical and structural stability, solubility and ability to interact effectively with TLR receptors, to ensure its efficacy. Based on the results, the vaccine was found to be antigenic, non-allergenic and thermostable. The molecular weight was calculated to be 26.5 kDa which is indicative of its stability. According to these parameters, the

designed vaccine is predicted to be a putative vaccine model that can be further validated *in-vivo*. To analyze the interactions of the vaccine with the immune receptors, molecular docking was performed which concluded favorable interactions. The immune simulations generated a consistent increase in the number of antigens and antibodies which depicts a harmonious increment in immune responses. Further, the immunoreactivity of the vaccine model was evaluated by its expression in *E. coli* (K12) (Ahmad et al. 2021), an appropriate host expression system. The optimization of the codon generated relevant CAI index (0.993) and GC content (55.7%) which are representatives of the suitable expression of the designed protein in the microorganism.

Antigenicity and druggability potentials of the selected proteins were evaluated in the second round of screening using the SVM approach. Our adopted strategy focused on essential enzymes involved in *P. aeruginosa* metabolism with the least homology in the host (human) genome. The proteome was screened and out of ~6000 genes, initially, 209 cytoplasmic essential genes were obtained. Among those 148 were enzymes that catalyze essential metabolic pathways. Host homologs exclusion reduced the number to 9. These 9 proteins were lpxD, asd, wbpB, FolK, glyQ, ribF, acnB, WaaA and wbpH. The earliest genome screening study reported about 300 essential protein targets for *P. aeruginosa* as short communication (Sakharkar et al. 2004). This approach adopted reductive screening of *P. aeruginosa* genome and focused on the essential protein which could be exploited as therapeutic targets. Latter these authors published their finding in 2007 (Perumal, et al. 2007). Our predicted 4 out of 9 potential drug targets were also reported in their findings. Our reported targets were lpxD, WaaA, glyQ and folK, while similar proteins were reported as lpxA, lpxB, lpxC, lpxK, WaaP, glyS and folK. Another study by the same group focused on Constraint-based flux balance analysis for predicting novel targets for *P. aeruginosa* (Perumal 2011). The authors developed a genome-scale metabolic network by assessing the gene deletion effects on flux balance and screened 41 unique targets. When compared with our predicted drug targets again 4 out of 9 were similar proteins lpxD, WaaA, ribF and folK, while similar proteins were reported as lpxA, lpxB, lpxC, and lpxK, WaaP, WaaG, ribC, ribD, ribH, folP and folB. Unlike previous studies, the target druggability was also evaluated dually. The binding pockets were predicted and best scoring pockets are presented. While database search revealed similar targets in other bacterial pathogen providing confidence to the predicted drug targets reported. Overlapping predicted protein candidates and the ones in the discussed vaccination studies demonstrate the validity of our proposed pipeline. Hence

this pipeline could also be employed for other pathogenic bacteria.

Conclusions

The availability of complete genomes/proteomes have facilitated many computational approaches. In this study, we report novel vaccine candidate proteins, stepwise prioritized using Reverse Vaccinology approach for rapid reduction of unwanted number of *P. aeruginosa* proteins. The targeted proteins play a fundamental role in bacterial survival, pathogenesis, and progression of the infection. Moreover, 8 antigenic, non-toxic and non-allergenic MHC-I, MHC-II and B-cell epitopes were selected and merged with Cholera toxin subunit B adjuvant to enhance the immune response of the final chimeric vaccine construct to improve its immunogenic efficacy. The stability of multi-epitope vaccine construct was assessed using normal mode simulation and molecular docking analysis that could be further tested *in-vivo*. In addition, the designed vaccine need to be experimentally validated to ensure its efficacy against multiple *P. aeruginosa* infections. In the second round, we found 9 essential enzymes with the best druggability score. These identified drug targets were: UDP-2-acetamido-2-deoxy-d-glucuronic acid 3-dehydrogenase WbpB (G3XD23), aspartate semi-aldehyde dehydrogenase (Q51344), 2-amino-4-hydroxy-6-hydroxymethylidihydropteridine pyrophosphokinase (Q9HV71), 3-deoxy-D-manno-octulosonic-acid transferase (Q9HUH7), glycyl-tRNA synthetase alpha chain (Q9I7B7), riboflavin kinase/FAD synthase (Q9HVM3), aconitate hydratase 2 (Q9I2V5), probable glycosyltransferase WbpH (G3XD85) and UDP-3-O-[3-hydroxy-lauroyl] glucosamine N-acyltransferase (Q9HXY6). Based on *in-silico* analyses, these reported vaccine and drug targets can help manage the antibiotic resistance phenomenon with special reference to hospital-acquired infections.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10989-021-10255-3>.

Author Contribution SI and SA designed the study and wrote the draft. MIR and MM performed analysis and helped in making figures. AA edited and revised the draft. SA reviewed and approved the final draft.

Funding The study is based on computational analysis for prediction of vaccine and drug candidates in PAO1 genome. Therefore, there is no funding or support involved from any company or organization.

Data Availability All data generated or analyzed during the study are included in the submitted manuscript. The sequences of the protein analyzed can be retrieved from UniProt database (uniprot.org) using their accession numbers.

Declarations

Conflict of interest There is no conflict of interest among the authors as well as with any organization in the world.

Research Involving Human and Animal Rights There is no involvement of human and/or animals trials in this study.

References

- Adams CP, Brantner VV (2006) Estimating the cost of new drug development: is it really \$802 million? *Health Aff* 25(2):420–428. <https://doi.org/10.1377/hlthaff.25.2.420>
- Ahmad S et al (2021) Immuno-informatics analysis of pakistan-based HCV subtype-3a for chimeric polypeptide vaccine design. *Vaccines* 9(3):293.
- Altschul SF et al (1990) Basic local alignment search tool. *J Mol Biol* 215(3):403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Arnold K et al (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics (oxford, England)* 22(2):195–201. <https://doi.org/10.1093/bioinformatics/bti770>
- Bairoch A, Apweiler R (2000) The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Res* 28(1):45–48. <https://doi.org/10.1093/nar/28.1.45>
- Baker SM, McLachlan JB, Morici LA (2019) Immunological considerations in the development of *Pseudomonas aeruginosa* vaccines. *Hum Vaccin Immunother*. <https://doi.org/10.1080/21645515.2019.1650999>
- Bavandpour AK, Bakhshi B, Najar-Peerayeh S (2020) The roles of mesoporous silica and carbon nanoparticles in antigen stability and intensity of immune response against recombinant subunit B of cholera toxin in a rabbit animal model. *Int J Pharm* 573:118868. <https://doi.org/10.1016/j.ijpharm.2019.118868>
- Berman H et al (2006) The worldwide Protein Data Bank (wwPDB): ensuring a single, uniform archive of PDB data. *Nucleic Acids Res* 35:D301–D303. <https://doi.org/10.1093/nar/gkl971>
- Bilofsky HS, Christian B (1988) The GenBank® genetic sequence data bank. *Nucleic Acids Res* 16(5):1861–1863. <https://doi.org/10.1093/nar/16.5.1861>
- Brunham RC, Plummer FA, Stephens RS (1993) Bacterial antigenic variation, host immune response, and pathogen-host coevolution. *Infect Immun* 61(6):2273. <https://doi.org/10.1128/iai.61.6.2273-2276.1993>
- Buchan DWA, Jones DT (2019) The PSIPRED protein analysis workbench: 20 years on. *Nucleic Acids Res* 47(W1):W402–W407. <https://doi.org/10.1093/nar/gkz297>
- Calabriso N et al (2016) Multiple anti-inflammatory and anti-atherosclerotic properties of red wine polyphenolic extracts: differential role of hydroxycinnamic acids, flavonols and stilbenes on endothelial inflammatory gene expression. *Eur J Nutr* 55(2):477–489. <https://doi.org/10.1007/s00394-015-0865-6>
- Chen L et al (2011) VFDB 2012 update: toward the genetic diversity and molecular evolution of bacterial virulence factors. *Nucleic Acids Res*. <https://doi.org/10.1093/nar/gkr989>
- Craik DJ et al (2013) The future of peptide-based drugs. *Chem Biol Drug Des* 81(1):136–147. <https://doi.org/10.1111/cbdd.12055>
- Dalsass M et al (2019) Comparison of open-source reverse vaccinology programs for bacterial vaccine antigen discovery. *Front Immunol* 10:113. <https://doi.org/10.3389/fimmu.2019.00113>
- de Vries SJ, van Dijk M, Bonvin AM (2010) The HADDOCK web server for data-driven biomolecular docking. *Nat Protoc* 5(5):883–897. <https://doi.org/10.1038/nprot.2010.32>
- Delany I, Rappuoli R, Seib KL (2013) Vaccines, reverse vaccinology, and bacterial pathogenesis. *Cold Spring Harb Perspect Med* 3(5):a012476. <https://doi.org/10.1101/cshperspect.a012476>
- Dimitrov I et al (2014) AllerTOP vol 2—a server for in silico prediction of allergens. *J Mol Model* 20(6):2278. <https://doi.org/10.1007/s00894-014-2278-5>
- Doytchinova IA, Flower DR (2007) VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinf* 8(1):4. <https://doi.org/10.1186/1471-2105-8-4>
- Fito-Boncompte L et al (2011) Full virulence of *Pseudomonas aeruginosa* requires OprF. *Infect Immun* 79(3):1176–1186. <https://doi.org/10.1128/IAI.00850-10>
- Foged C, Hansen J, Agger EM (2012) License to kill: formulation requirements for optimal priming of CD8+ CTL responses with particulate vaccine delivery systems. *Eur J Pharm Sci* 45(4):482–491. <https://doi.org/10.1016/j.ejps.2011.08.016>
- François B et al (2012) Safety and pharmacokinetics of an anti-PcrV PEGylated monoclonal antibody fragment in mechanically ventilated patients colonized with *Pseudomonas aeruginosa*: a randomized, double-blind, placebo-controlled trial. *Crit Care Med* 40(8):2320–2326. <https://doi.org/10.1097/CCM.0b013e31825334f6>
- Frieri M, Kumar K, Boutin A (2017) Antibiotic resistance. *J Infect Public Health* 10(4):369–378. <https://doi.org/10.1016/j.jiph.2016.08.007>
- Garg A, Gupta D (2008) VirulentPred: a SVM based prediction method for virulent proteins in bacterial pathogens. *BMC Bioinf* 9(1):62. <https://doi.org/10.1186/1471-2105-9-62>
- Gasteiger E et al (2005) Protein identification and analysis tools on the ExPASy server. Springer, Berlin
- Gellatly SL, Hancock RE (2013) *Pseudomonas aeruginosa*: new insights into pathogenesis and host defenses. *Pathogens and Disease* 67(3):159–173. <https://doi.org/10.1111/2049-632X.12033>
- Giuliani MM et al (2006) A universal vaccine for serogroup B meningococcus. *Proc Natl Acad Sci* 103(29):10834–10839. <https://doi.org/10.1073/pnas.0603940103>
- Guan P et al (2003) MHCpred: a server for quantitative prediction of peptide–MHC binding. *Nucleic Acids Res* 31(13):3621–3624. <https://doi.org/10.1093/nar/gkg510>
- Harmer C et al (2013) Modulation of gene expression by *Pseudomonas aeruginosa* during chronic infection in the adult cystic fibrosis lung. *Microbiology* 159:2354–2363. <https://doi.org/10.1099/mic.0.066985-0>
- He Y, Xiang Z, Mobley HL (2010) Vaxign: the first web-based vaccine design program for reverse vaccinology and applications for vaccine development. *J Biomed Biotechnol* 2010:297505. <https://doi.org/10.1155/2010/297505>
- Hebdtich M et al (2017) Protein-Sol: a web tool for predicting protein solubility from sequence. *Bioinformatics* 33(19):3098–3100. <https://doi.org/10.1093/bioinformatics/btx345>
- Hedgecoe A, Martin P (2003) The drugs don't work: expectations and the shaping of pharmacogenetics. *Soc Stud Sci* 33(3):327–364. <https://doi.org/10.1177/03063127030333002>
- Heo L, Park H, Seok C (2013) GalaxyRefine: protein structure refinement driven by side-chain repacking. *Nucleic Acids Res* 41(W1):W384–W388. <https://doi.org/10.1093/nar/gkt458>
- Hopkins AL, Groom CR (2002) The druggable genome. *Nat Rev Drug Discovery* 1(9):727. <https://doi.org/10.1038/nrd892>
- Juan C, Peña C, Oliver A (2017) Host and pathogen biomarkers for severe *Pseudomonas aeruginosa* infections. *J Infect Dis* 215:S44–S51. <https://doi.org/10.1093/infdis/jiw299>

- Juhas M (2015) *Pseudomonas aeruginosa* essentials: an update on investigation of essential genes. *Microbiology* 161(11):2053–2060. <https://doi.org/10.1099/mic.0.000611>
- Källberg M et al (2012) Template-based protein structure modeling using the RaptorX web server. *Nat Protoc* 7(8):1511. <https://doi.org/10.1038/nprot.2012.085>
- Karkman A et al (2018) Antibiotic-resistance genes in waste water. *Trends Microbiol* 26(3):220–228. <https://doi.org/10.1016/j.tim.2017.09.005>
- Keller TH, Pichota A, Yin Z (2006) A practical view of ‘drugability.’ *Curr Opin Chem Biol* 10(4):357–361. <https://doi.org/10.1016/j.cbpa.2006.06.014>
- Knapp B et al (1999) A recombinant hybrid outer membrane protein for vaccination against *Pseudomonas aeruginosa*. *Vaccine* 17(13–14):1663–1666. [https://doi.org/10.1016/s0264-410x\(98\)00420-4](https://doi.org/10.1016/s0264-410x(98)00420-4)
- Lenz GR, Nash HM, Jindal S (2000) Chemical ligands, genomics and drug discovery. *Drug Discovery Today* 5(4):145–156. [https://doi.org/10.1016/s1359-6446\(00\)01468-9](https://doi.org/10.1016/s1359-6446(00)01468-9)
- Li W, Godzik A (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22(13):1658–1659. <https://doi.org/10.1093/bioinformatics/btl1158>
- Li L, Stoekert CJ, Roos DS (2003) OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res* 13(9):2178–2189. <https://doi.org/10.1101/gr.1224503>
- Lister PD, Wolter DJ, Hanson ND (2009) Antibacterial-resistant *Pseudomonas aeruginosa*: clinical impact and complex regulation of chromosomally encoded resistance mechanisms. *Clin Microbiol Rev* 22(4):582–610. <https://doi.org/10.1128/CMR.00040-09>
- Lovell SC et al (2003) Structure validation by C α geometry: phi, psi and C β deviation. *Proteins* 50(3):437–450. <https://doi.org/10.1002/prot.10286>
- Lücking H et al (1997) The *Pseudomonas aeruginosa* outer membrane protein I vaccine: immunogenicity and safe administration in man. *Behring Inst Mitt* 98:326–337 (PMID: 9382756)
- Luo H et al (2014) DEG 10, an update of the database of essential genes that includes both protein-coding genes and noncoding genomic elements. *Nucleic Acids Res* 42(D1):D574–D580. <https://doi.org/10.1093/nar/gkt1131>
- Majid M, Andleeb S (2019) Designing a multi-epitopic vaccine against the enterotoxigenic *Bacteroides fragilis* based on immunoinformatics approach. *Sci Rep* 9(1):1–15.
- Mansouri E et al (2003) Clinical study to assess the immunogenicity and safety of a recombinant *Pseudomonas aeruginosa* OprF-OprI vaccine in burn patients. *FEMS Immunol Med Microbiol* 37(2–3):161–166. [https://doi.org/10.1016/S0928-8244\(03\)00072-5](https://doi.org/10.1016/S0928-8244(03)00072-5)
- Mayrose I et al (2007) Pepitope: epitope mapping from affinity-selected peptides. *Bioinformatics* 23(23):3244–3246. <https://doi.org/10.1093/bioinformatics/btm493>
- McLean K et al (2019) Genomic analysis identifies novel *Pseudomonas aeruginosa* resistance genes under selection during inhaled aztreonam therapy in vivo. *Antimicrob Agents Chemother* 63(9):e00866–e919. <https://doi.org/10.1128/AAC.00866-19>
- Montor WR et al (2009) Genome-wide study of *Pseudomonas aeruginosa* outer membrane protein immunogenicity using self-assembling protein microarrays. *Infect Immun* 77(11):4877–4886. <https://doi.org/10.1128/IAI.00698-09>
- Moradali MF, Ghods S, Rehm BH (2017) *Pseudomonas aeruginosa* lifestyle: a paradigm for adaptation, survival, and persistence. *Front Cell Infect Microbiol* 7:39. <https://doi.org/10.3389/fcimb.2017.00039>
- Mosquera-Rendón J et al (2016) Pangenome-wide and molecular evolution analyses of the *Pseudomonas aeruginosa* species. *BMC Genomics* 17(1):45. <https://doi.org/10.1186/s12864-016-2364-4>
- Ni Z et al (2017) Antibiotic resistance determinant-focused *Acinetobacter baumannii* vaccine designed using reverse vaccinology. *Int J Mol Sci* 18(2):458. <https://doi.org/10.3390/ijms18020458>
- Pang Z et al (2019) Antibiotic resistance in *Pseudomonas aeruginosa*: mechanisms and alternative therapeutic strategies. *Biotechnol Adv* 37(1):177–192. <https://doi.org/10.1016/j.biotechadv.2018.11.013>
- Paul S et al (2016) TepiTool: a pipeline for computational prediction of T cell epitope candidates. *Curr Protoc Immunol* 114:18 19 1-18 19 24. <https://doi.org/10.1002/cpim.12>
- Pennington JE et al (1975) Use of a *Pseudomonas aeruginosa* vaccine in patients with acute leukemia and cystic fibrosis. *Am J Med* 58(5):629–636. [https://doi.org/10.1016/0002-9343\(75\)90498-2](https://doi.org/10.1016/0002-9343(75)90498-2)
- Perumal D et al (2007) Differential genome analyses of metabolic enzymes in *Pseudomonas aeruginosa* for drug target identification. *In Silico Biol* 7:453–465
- Perumal D et al (2011) Targeting multiple targets in *Pseudomonas aeruginosa* PAO1 using flux balance analysis of a reconstructed genome-scale metabolic network. *J Drug Target* 19(1):1–13. <https://doi.org/10.3109/10611861003649753>
- Pires DPP, et al (2018) Bacteriophage interaction with *P. aeruginosa* and *A. baumannii* biofilms and colonizing human airway epithelium. <https://doi.org/10.1186/s12941-020-00389-5>
- Priebe GP, Goldberg JB (2014) Vaccines for *Pseudomonas aeruginosa*: a long and winding road. *Expert Rev Vaccines* 13(4):507–519. <https://doi.org/10.1586/14760584.2014.890053>
- Priebe GP et al (2003) Protection against fatal *Pseudomonas aeruginosa* pneumonia in mice after nasal immunization with a live, attenuated *aroA* deletion mutant. *Infect Immun* 71(3):1453–1461. <https://doi.org/10.1128/IAI.71.3.1453-1461.2003>
- Qiao M et al (2018) Review of antibiotic resistance in China and its environment. *Environ Int* 110:160–172. <https://doi.org/10.1016/j.envint.2017.10.016>
- Ranganathan S, Garg G (2009) Secretome: clues into pathogen infection and clinical applications. *Genome Med* 1(11):113. <https://doi.org/10.1186/gm113>
- Rappuoli R, Bloom DE, Black S (2017) Deploy vaccines to fight superbugs. *Nature*. <https://doi.org/10.1038/d41586-017-08323-0>
- Rashid MI (2020) Study of pyocyanin induced virulence & pathogenicity of nosocomial infectious agent *Pseudomonas aeruginosa* and its prevention. *Pak J Pharmaceu Sci* 33(3):915–922
- Rashid MI et al (2017) Prediction of vaccine candidates against *Pseudomonas aeruginosa*: an integrated genomics and proteomics approach. *Genomics* 109(3):274–283. <https://doi.org/10.1016/j.ygeno.2017.05.001>
- Rizvi M et al (2015) Synergy of drug combinations in treating multidrug-resistant *Pseudomonas aeruginosa*. *Australas Med J* 8(1):1. <https://doi.org/10.4066/AMJ.2015.2096>
- Roses AD (2000) Pharmacogenetics and future drug development and delivery. *Lancet* 355(9212):1358–1361. [https://doi.org/10.1016/S0140-6736\(00\)02126-7](https://doi.org/10.1016/S0140-6736(00)02126-7)
- Sachdeva G et al (2004) SPAAN: a software program for prediction of adhesins and adhesin-like proteins using neural networks. *Bioinformatics* 21(4):483–491. <https://doi.org/10.1093/bioinformatics/bti028>
- Saha S, Raghava GP (2006) AlgPred: prediction of allergenic proteins and mapping of IgE epitopes. *Nucleic Acids Res* 34:W202–W209. <https://doi.org/10.1093/nar/gkl343>
- Sakharkar KR, Sakharkar MK, Chow VT (2004) A novel genomics approach for the identification of drug targets in pathogens, with special reference to *Pseudomonas aeruginosa*. *Silico Biology* 4(3):355–360
- Sette A, Rappuoli R (2010) Reverse vaccinology: developing vaccines in the era of genomics. *Immunity* 33(4):530–541. <https://doi.org/10.1016/j.immuni.2010.09.017>
- Smialowski P et al (2012) PROSO II—a new method for protein solubility prediction. *FEBS J* 279(12):2192–2200. <https://doi.org/10.1111/j.1742-4658.2012.08603.x>

- Sorichter S et al (2009) Immune responses in the airways by nasal vaccination with systemic boosting against *Pseudomonas aeruginosa* in chronic lung disease. *Vaccine* 27(21):2755–2759. <https://doi.org/10.1016/j.vaccine.2009.03.010>
- Szklarczyk D et al (2011) The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Res* 39(suppl 1):D561–D568. <https://doi.org/10.1093/nar/gkq973>
- Tusnady GE, Simon I (2001) The HMMTOP transmembrane topology prediction server. *Bioinformatics* 17(9):849–850. <https://doi.org/10.1093/bioinformatics/17.9.849>
- Vincent J-L (2014) Vaccine development and passive immunization for *Pseudomonas aeruginosa* in critically ill patients: a clinical update. *Future Microbiol* 9(4):457–463. <https://doi.org/10.2217/fmb.14.10>
- Volkamer A et al (2012) DoGSiteScorer: a web server for automatic binding site prediction, analysis and druggability assessment. *Bioinformatics* 28(15):2074–2075. <https://doi.org/10.1093/bioinformatics/bts310>
- Wang K et al (2017) The rapid in vivo evolution of *Pseudomonas aeruginosa* in ventilator-associated pneumonia patients leads to attenuated virulence. *Open Biol* 7(9):170029. <https://doi.org/10.1098/rsob.170029>
- WHO (2021) Global priority list of antibiotic research bacteria to guide research, discovery, and development of new antibiotics. World Health Organization. <https://www.combatamr.org.au/news-events/who-global-priority-pathogens-list-of-antibiotic-resistant-bacteria>
- Weigt S et al (2018) Gene expression profiling of bronchioalveolar lavage cells during aspergillus colonization of the lung allograft. *Transpl Clin Stud* 102(6):986–993. <https://doi.org/10.1097/TP.0000000000002058>
- Westritschnig K et al (2014) A randomized, placebo-controlled phase I study assessing the safety and immunogenicity of a *Pseudomonas aeruginosa* hybrid outer membrane protein OprF/I vaccine (IC43) in healthy volunteers. *Hum Vaccin Immunother* 10(1):170–183. <https://doi.org/10.4161/hv.26565>
- Winsor GL et al (2011) Pseudomonas Genome Database: improved comparative analysis and population genomics capability for *Pseudomonas genomes*. *Nucleic Acids Res* 39:D596–D600. <https://doi.org/10.1093/nar/gkq869>
- Xiang Z, He Y (2009) Vaxign: a web-based vaccine target design program for reverse vaccinology. *Proc Vaccinol* 1(1):23–29. <https://doi.org/10.1016/j.provac.2009.07.005>
- Xipell M et al (2017) Successful treatment of three severe MDR or XDR *Pseudomonas aeruginosa* infections with ceftolozane/tazobactam. *Future Microbiol* 12(14):1323–1326. <https://doi.org/10.2217/fmb-2017-0018>
- Yang J, Zhang Y (2015) I-TASSER server: new development for protein structure and function predictions. *Nucleic Acids Res* 43(W1):W174–W181. <https://doi.org/10.1093/nar/gkv342>
- Yang F et al (2017) Protective efficacy of the trivalent *Pseudomonas aeruginosa* vaccine candidate PcrV-OprI-Hcp1 in murine pneumonia and burn models. *Sci Rep* 7(1):3957. <https://doi.org/10.1038/s41598-017-04029-5>
- Yu CS, Lin CJ, Hwang JK (2004) Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n-peptide compositions. *Protein Sci* 13(5):1402–1406. <https://doi.org/10.1110/ps.03479604>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.