 doi:10.20944/preprints202102.0579.v1

Quest for novel preventive and therapeutic options against multidrug-resistant *Pseudomonas aeruginosa*

Sidra Irum, Saadia Andleeb*, Amjad Ali, Muhammad Ibrahim Rashid, Mahnoor Majid
Department of Industrial Biotechnology, Atta-ur-Rahman School of Applied Biosciences (ASAB), National University of Sciences & Technology (NUST), Islamabad, Pakistan.

*Corresponding Author
Dr. Saadia Andleeb
Associate Professor
Atta-ur- Rahman School of Applied Biosciences
National University of Sciences and Technology
Sector H-12, Islamabad, 44000, Pakistan
Phone: +92-51-90856133
Email address: saadia.andleeb@asab.nust.edu.pk
Saadiamarwat@yahoo.com

Abstract

Pseudomonas 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. Here, we designed a chimeric vaccine against *P. aeruginosa* using a subtractive proteomics approach and identified nine unique enzymes as novel drug targets in PAO1 proteome. A total of five unique proteins were selected as potential vaccine candidates based on essentiality, extracellular localization, virulence, antigenicity, pathway association, protein-protein interaction analysis, 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 proteins were further analyzed using a reverse vaccinology approach to identify immunogenic and antigenic T cell and B cell epitopes. The best scoring epitopes qualifying for all set criteria were then further subjected to the construction of a polypeptide multi-epitope vaccine construct with cholera toxin B (CtxB) subunit as an adjuvant.

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-hydroxymethyldihydropteridine 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-hydroxylauroyl] 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.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 February 2021 doi:10.20944/preprints202102.0579.v1

1. 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 [1]. Four factors make it one of the most unique and widespread bacterial pathogen: the wide range of hosts it infects (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 [2,3]. This pathogen 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 [4].

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 [5, 6, 5, 7].

The most serious clinical challenge associated with this bacterium is its resistance potential against the antibacterial chemotherapeutic agents, as emphasized by WHO [8,1]. Infections caused by this notorious pathogen are often difficult to eradicate despite intense antibiotic combination treatments [3].

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* [9]. Several vaccine candidates have been tested [10, 11], however, no candidate has been licensed yet for clinical trials [12]. Trivial approaches based on injecting live or attenuated pathogens have chances to provoke infections and might lead to death after primary and secondary shots [13]. "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 [14]. 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 [15]. 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 [5]. The identification of novel immunogenic candidates can potentially contribute to the development of effective vaccines. The prospective candidate proteins should be screened for surface exposure, high immunogenic potential, high abundance, least number of transmembrane helices and involvement in virulence [16]. *P. aeruginosa* possesses abundant outer membrane proteins [17], of which the function and expression of a large number are still unknown.

The basic theme of our study was the rigorous screening of *Pseudomonas aeruginosa* strain PAO1 proteome for the identification of novel epitopes which are capable of binding MHC (major histocompatibility complex class I & II) molecules for antigen presentation inside the host. The relatively large genome size of *P. aeruginosa* contributes to its versatility and variation in response to host immune responses and environmental factors. One of the challenges in developing vaccines against this pathogen is that it employs multiple pathways to cause and increase the severity of infections, and hence a single vaccine candidate that targets only one or two pathways do not provide effective protection [6]. The plasticity of microbial genomes is another hurdle in the selection of conserved antigenic candidates [18, 19]. Therefore, reverse vaccinology and subtractive proteomics approach were used to identify novel immunogenic peptides.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 February 2021 doi:10.20944/preprints202102.0579.v1

Genome projects and the development of bioinformatics tools have resulted in a paradigm shift in drug development research [20]. Conventional drug development strategies based on hunches and trial and error approaches are costly time consuming and inefficient [21-23]. 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 [24, 25]. The "druggable genome" indicates the genes which hold therapeutic potential and can serve as targets for drugs [20]. 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 [26]. 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** detailing all the screening steps for genome analysis. Subsequently, a non-human homologous exoproteome was selected and screened for B-cell and T-cell binding affinity. Highly conserved epitopes with the ability to bind to maximum MHC-I and MHC-II class molecules were selected followed by their linkage to design a multi-epitopic vaccine chimera that could be subjected to in-vivo testing. The cytoplasm localized proteins involved in essential metabolic pathways were subjected to druggability assessment

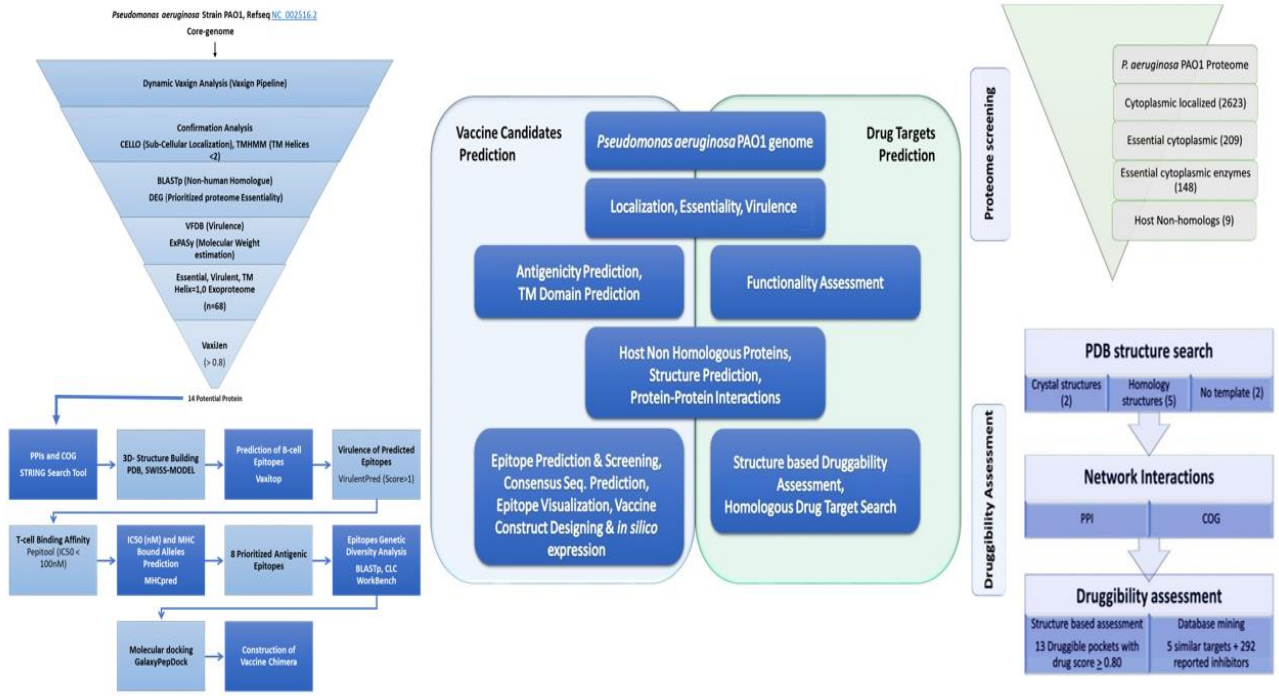

Fig. 1. The schematic workflow overview: The *Pseudomonas aeruginosa* reference strain PAO1 was subjected to reductive genome screening for the prediction of potential vaccine candidates and identification of novel drug targets.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 February 2021 doi:10.20944/preprints202102.0579.v1

2. Materials & Methods

2.1. Acquisition of non-paralogous proteins sequences

Primary genomic and proteomic data of *P. aeruginosa* reference strain PAO1 (RefSeq NC_002516.2, Proteome; UP000002438) was retrieved from NCBI GenBank [27] and UniProt [28]. The pan-proteome of PAO1 was subjected to a stepwise identification of vital virulent proteins [5]. The complete proteome was checked for redundancy through CD-HIT (http://weizhongli-lab.org/cdhit_suite/cgi-bin/index.cgi?cmd¼cd-hit) [29].

2.2. 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/) [30]. 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 [31].
The topology of secreted and surface-exposed proteins was verified using TMHMM v2.0 [32]. Proteins with more than one transmembrane helices were removed from the shortlisted pool. The Vaxign pipeline calculated adhesins using SPAAN [33] 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 [34] 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 [35] 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 [36]. Apart from the human homologs, orthologs search across Pseudomonas genus was conducted using organism-specific BLASTp search and “The Pseudomonas genome Database” [37].

2.3. Screening of pathogen essential, virulent and antigenic proteins

Database of Essential Genes (DEG) version 10.4 [38] 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 [39]. 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) [40] and SVM-based tool, Virulentpred [41]. Antigenicity scores of the filtered proteins were estimated by Vaxijen v2.0 [42]

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 February 2021 doi:10.20944/preprints202102.0579.v1

(threshold = 0.4). Proteins with high antigenic scores were prioritized as good vaccine candidates.

2.4. 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 [43]. Proteins having a molecular weight of **≤110 kDa** were selected while the rest were excluded **(Table 1).**

Table 1. Characteristics of finalized vaccine candidate proteins

Protein Accession	Gene Name	Protein Product	Localization		Adhesins Probability	Trans-membrane helices		Antigenicity score	Essentiality	Virulence	Molecular Weight (kDa)
			PSORTb	CELLO		HMMTOP	TMHMM				
P13794	oprF	Outer membrane porin F	Outer Membrane	Outer Membrane	0.553	0	0	0.8306	✓	✓	37.63
P32977	oprO	Porin O	Outer Membrane	Outer Membrane	0.542	1	1	0.8091	✓	✓	47.78
P32722	oprD	Porin D	Outer Membrane	Outer Membrane	0.564	0	0	0.8658	✓	✓	48.36
Q9I5A7	PA0833	Uncharacterized protein	Outer Membrane	Extracellular	0.544	0	0	0.8186	✓	✓	12.07
Q9I3P1	PA1471	Uncharacterized protein	Extracellular	Extracellular	0.547	1	1	0.8782	✓	✓	6.31
G3XDA5	oprE	Anaerobically-induced outer membrane porin OprE	Outer Membrane	Outer Membrane	0.549	1	1	0.8417	✓	✓	49.66
Q9I456	PA1288	Probable outer membrane protein	Outer Membrane	Outer Membrane	0.641	0	0	0.8067	✓	✓	45.56
G3XDA9	pra	Protein activator	Extracellular	Extracellular	0.692	0	0	0.8073	✓	✓	16.19
Q9I2E8	PA1956	Uncharacterized protein	Extracellular	Extracellular	0.666	0	0	0.8333	✓	✓	15.83
Q9HWW1	oprG	Outer membrane protein OprG	Outer Membrane	Outer Membrane	0.552	0	0	0.8837	✓	✓	25.19
Q9HWU6	cupB5	Adhesive protein CupB5	Extracellular	Extracellular	0.677	1	1	0.8326	✓	✓	100.43
Q9HUT9	PA4874	Uncharacterized protein	Periplasmic	Periplasmic	0.636	0	0	0.858	✓	✓	14.43
Q9I5K4	PA0724	Probable coat protein A of bacteriophage Pf1	Extracellular	Extracellular	0.666	1	1	1.0295	✓	✓	42.91

2.5. Prioritized protein's 3-D model building and druggability assessment

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 February 2021 doi:10.20944/preprints202102.0579.v1

The presence of a 3D protein structure was checked in Protein Data Bank (PDB) [44] for all proteins. SWISS-MODEL [45] and RaptorX [46] 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 [47]. 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.

2.6. 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 [48] 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.

2.7. Antigenic epitopes prediction

Vaxitop online server was used to predict continuous B-cell Epitopes from selected proteins using default threshold values [49]. The predicted peptides were reexamined for topology via TMHMM v2.0 [32], and antigenicity scores using Vaxijen [42] (threshold value > 0.4). The peptides with antigenicity scores $\geq$ 0.8 were considered more antigenic 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 [50]. 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 IC_{50} value [51]. Extracted prioritized peptide out of respective virulent protein candidates was again subjected to VirulentPred [41] (threshold value $\geq$ 0.5). The topology of the selected epitopes was visualized using Pepitope server [52].

2.8. 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 [42].

2.9. Designing of polypeptide vaccine construct

2.9.1. Assembly of the B-cell epitopes for designing of the vaccine

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 February 2021 doi:10.20944/preprints202102.0579.v1

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**).

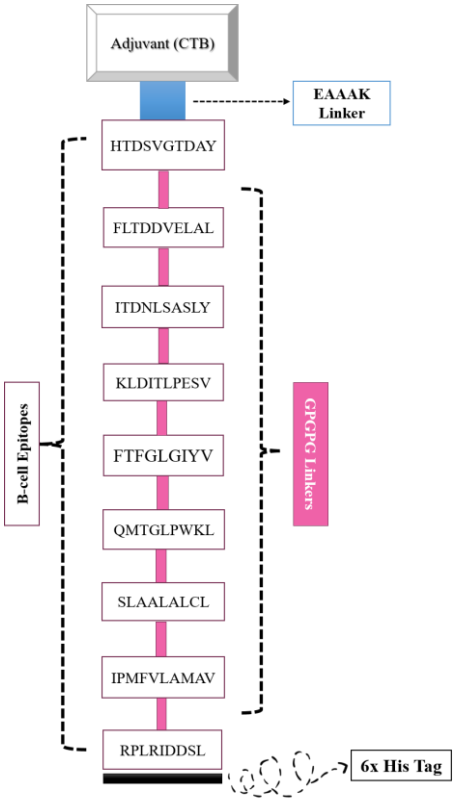

Fig. 2. Representation of the designed vaccine model. The amino terminus starts with the adjuvant attached to the first B-cell epitope with an EAAAK linker. The rest of the epitopes are conjoined together via GPGPG linkers. At the c-terminus, a 6x His Tag is attached for identification purpose.

2.9.2. Prediction of the allergenicity and solubility

To predict the allergenicity of the vaccine sequence, AlgPred v.2.0 [53] and AllerTOP v.2.0 [54] 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[55]and Protein-sol servers [56].

2.9.3. Evaluation of the physiochemical properties

For the functional characterization of the protein, ExPASy ProtParam[43] tool was used which explored the vaccine sequence on a number of physiochemical 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 [57].

2.9.4. 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 [58]. To validate

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 February 2021 doi:10.20944/preprints202102.0579.v1

the generated 3D structure, its refinement was carried out to enhance the stability of the predicted model using GalaxyRefine [59]. The refined protein was then validated through the Ramachandran plot assessment via RAMPAGE server [60].

2.9.5. 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 [61]. For the docking analysis, TLR4 3D model was retrieved form the PDB and active and passive residues were predicted for both the complexes via CPORT[94]. iMODs[95] 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.

2.9.6. Codon optimization of the vaccine construct

JCat[96] 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 BglII and Apa1 were not selected for the final optimization of the vaccine sequence.

2.9.7. 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 injections[97].

3. Results

3.1. Screening of essential and host non-homologous proteins

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

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 February 2021 doi:10.20944/preprints202102.0579.v1

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 of 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)**.

3.3. 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.

3.4. 3-D structure prediction and refinement

Prioritized proteins were checked for crystalline structure availability in Protein Data Bank. Only five potential vaccine candidate proteins (OprF, OprO, OprD, OprE, and $CupB_5$) have had available structures in both organism-specific and general PDB search. Remaining five proteins, PA1288, pra, PA4874, PA0724, and PA1956 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 except for PA4541 because its RaptorX structure was not stable. 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 has any crystalline structure and also lacked any suitable template for homology prediction thence 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. 3).**

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 February 2021 doi:10.20944/preprints202102.0579.v1

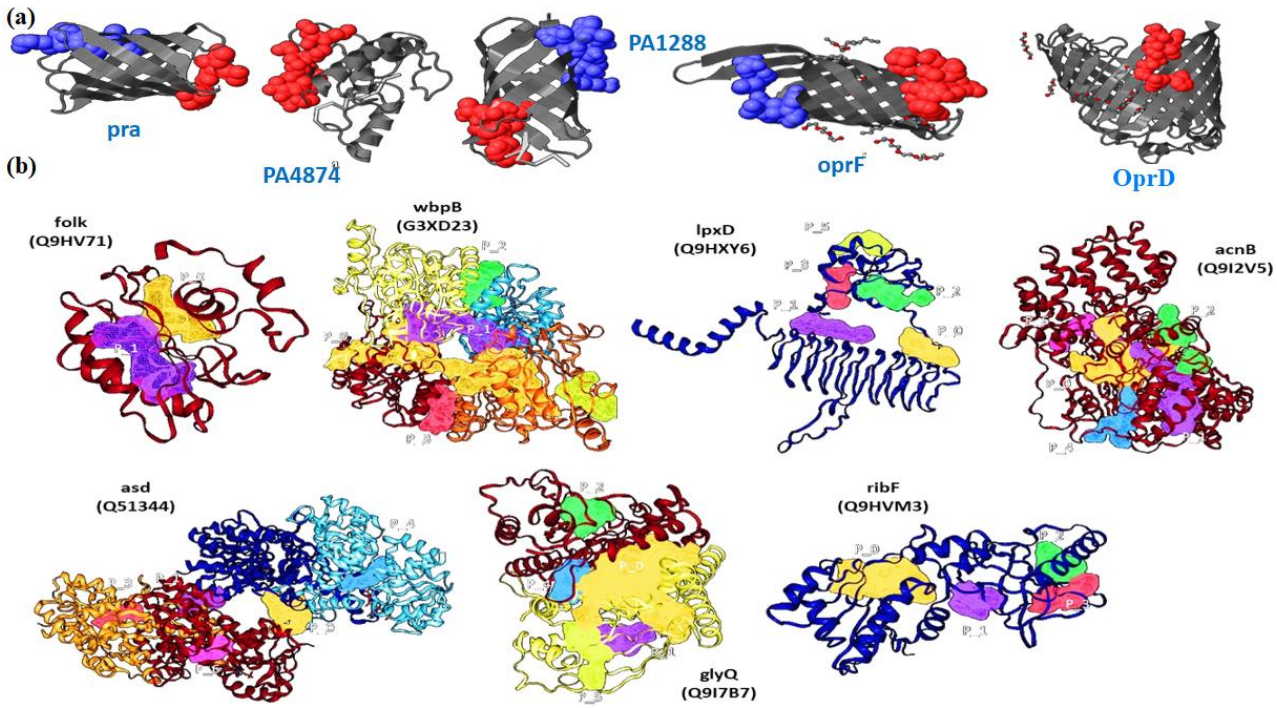

Fig. 3. Three-dimensional structure of the predicted vaccine candidate proteins along with the topological view of corresponding antigenic peptides (on top) and druggable target proteins. Figure (a) illustrates the RaptorX modeled three-dimensional structures of proteins with their corresponding surface-exposed 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.

3.5. 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. 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 precursor 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

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 February 2021 doi:10.20944/preprints202102.0579.v1

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 $3.55x10^{-08}$. 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). **qlyQ** 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).

3.6. Designing and screening of 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 IC_{50} value analysis for DRB1*0101 allele, several bound alleles by epitopes and T-cell binding affinity (prediction by tepitool server, IC50 value cut off < 100nMol). 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

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 February 2021 doi:10.20944/preprints202102.0579.v1

and recognition by both B-cell and T-cell. Finally, 8 peptides with the lowest IC_{50} values (cut off <100nMol), highest antigenicity scores, having a binding affinity with a maximum number of alleles of both classes were selected as shown in **Table 2**.

Table 2. List of prioritized antigenic and virulent epitopes from 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	Consensus (Percentage)
P13794	oprF	350	Outer Membrane	HTDSVGTDAY	65.53	4	0	0.000385	1.9872	100
P13794	oprF	350	Outer Membrane	FLTDDVELAL	2.09	18	0	0.0042	0.4731	100
P32722	oprD	443	Outer Membrane	ITDNLSASLY	4.73	15	0	0.00000231	0.7422	100
Q9I456	PA1288	424	Outer Membrane	KLDITLPESV	5.28	6	3	0.00166	1.2296	100
Q9I456	PA1288	424	Outer Membrane	FTFGLGIYV	55.34	27	1	0.00277	2.1776	100
G3XDA9	pra	162	Extracellular	QMTGLPWKL	9.23	12	0	0.0125	2.2073	QMTGLPW**T**L
G3XDA9	pra	162	Extracellular	SLAALALCL	3.85	12	0	0.0149	0.9236	100
Q9HUT9	PA4874	138	Periplasmic	IPMFVLAMAV	5.06	14	23	0.000385	1.6836	100
Q9HVN6	PA4541	1,417	Extracellular	RPLRIDDSL	86.9	9	1	0.5395	1.7311	100

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**).

3.7. 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.

3.8. Alternate drug targets assessment

3.8.1. 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.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 February 2021 doi:10.20944/preprints202102.0579.v1

The pockets with drug scores> 0.50 were selected as shown in figure 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.

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

3.8.2. 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 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-hydroxymethyldihydropteridine pyrophosphokinase (P26281) having 74 reported compounds and 51.3% identity. The most similar reported targets are described in **Table 4**.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 February 2021 doi:10.20944/preprints202102.0579.v1

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-hydroxymethyldihydropteridine pyrophosphokinase	P26281	SINGLE PROTEIN	*Escherichia coli (strain K12)*	74	51.3	164	3.00E-41
Q9I7B7	CHEMBL3086	Acetyl-Coenzyme A carboxylase	Q5SWU9	SINGLE PROTEIN	*Mus musculus*	2	42.5	28.9	1.80E+00
Q9I7B7	CHEMBL2397	Acetyl-CoA carboxylase 1	P11497	SINGLE PROTEIN	*Rattus norvegicus*	150	42.5	28.9	1.90E+00
G3XD85	CHEMBL3774301	Cysteine proteinase 1	Q01957	SINGLE PROTEIN	*Entamoeba histolytica*	3	42.11	28.9	2.20E+00
G3XD85	CHEMBL3774301	Cysteine proteinase 1	Q01957	SINGLE PROTEIN	*Entamoeba histolytica*	3	42.11	28.9	2.20E+00
Q9HV71	CHEMBL3559681	2-amino-4-hydroxy-6-hydroxymethyldihydropteridin e pyrophosphokinase	X5EH84	SINGLE PROTEIN	*Staphylococcus aureus*	60	41.55	117	5.00E-27

3.9. Designing of polypeptide vaccine construct

3.9.1. Prediction of the allergenicity and solubility

The protein sequence of the vaccine was found to be non-allergenic in nature as predicted by the AlgPred 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.

3.9.2. Secondary and tertiary structures Prediction, refinement and validation of vaccine constructs

For the characterization of the designed vaccine, ExPASy server was employed for the confirmation of various physiochemical properties. The molecular weight of the designed sequence was calculated as 26.5 kDa affirming its stability and ease of purification. Further, 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 vaccine protein. Upon the secondary structure analysis, 28% alpha-helix, 18% β sheets and, 52% random coil were predicted in the vaccine sequence as shown in **Fig. 4.** I-TASSER was used for predicting the tertiary structure of the construct. The first phase of refinement was carried out by the ModRefiner which was preceded by the GalaxyRefine treatment for the effective enhancement of the tertiary structure of the vaccine. 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 after the final refinement.

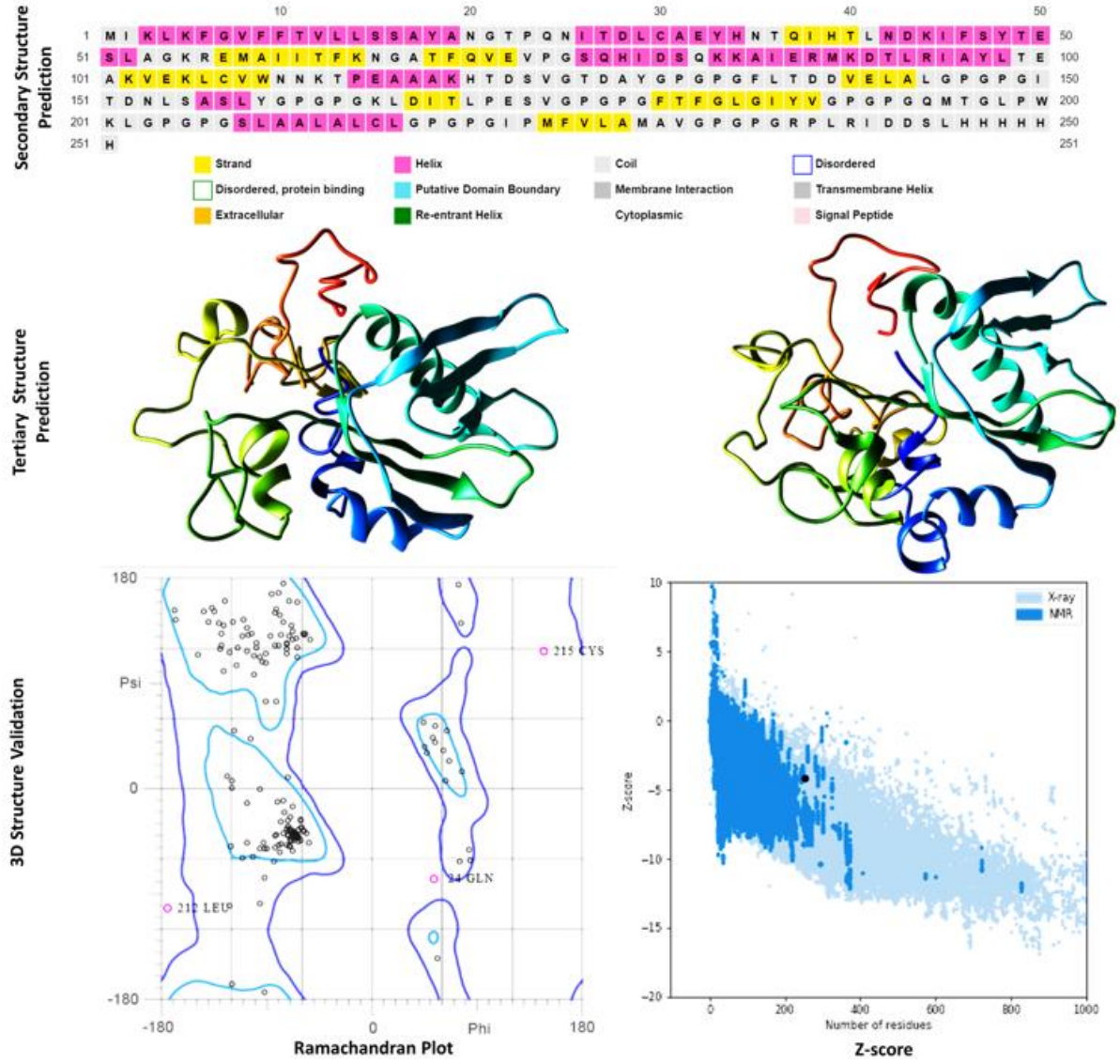

Fig. 4. Secondary and tertiary structures of the proposed vaccine. The 3D model of the vaccine protein obtained through I-TASSER based on homology modeling and the refined structure of vaccine protein obtained via Galaxy Refine after it refines on ModRefiner along with its Ramachandran plot analysis.

3.9.3. 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 $2.4784e^{-05}$. 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.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 February 2021 doi:10.20944/preprints202102.0579.v1

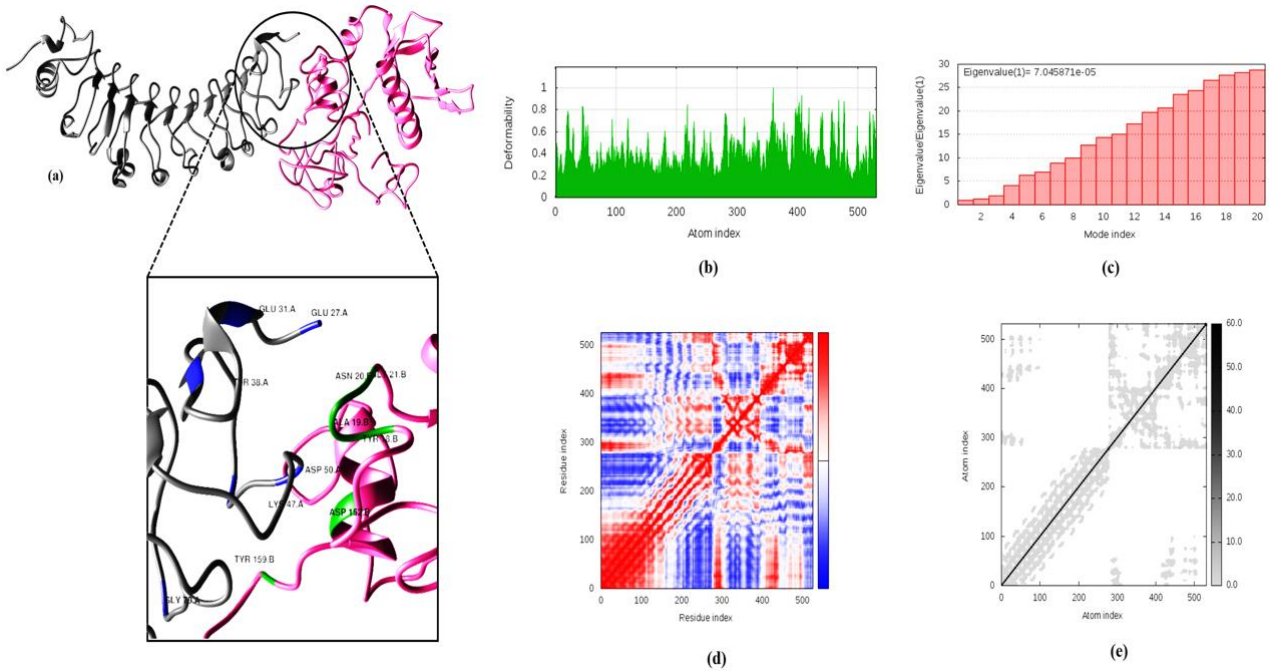

Fig. 7. Molecular docking of Peptide Vaccine with TLR4 and Normal mode analysis. The docking complex shows the binding of the protein with the immune receptor TLR4. (a) The grey molecule represents the vaccine with its active amino acid residues colored as blue whereas the pink molecule represents the TLR4 with green colored active amino acid residues. Both the complexes are directed towards each other as depicted by the arrows. (b) The B-factor value of the protein model generated (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 vaccine protein representing its flexibility.

3.9.4. 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 BglII and ApaI that indicates its safety for cloning.

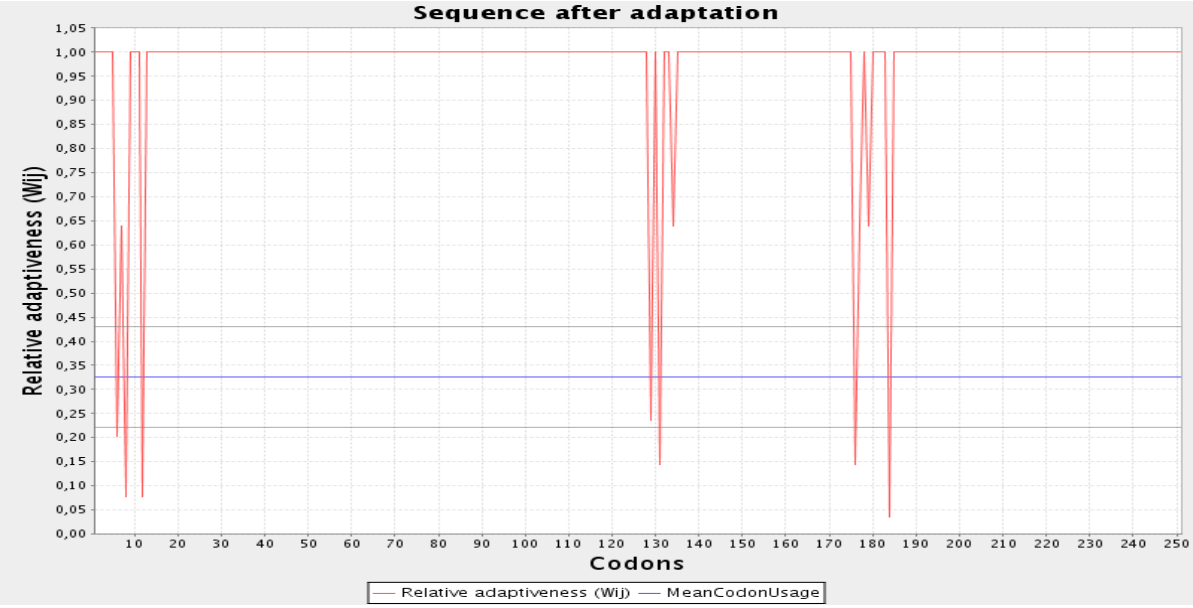

Fig. 6. The codon adaptation index (CAI). The CAI graph indicates that codon optimization of the vaccine with 0.933% and the GC content of the optimized codons (55.77%).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 February 2021 doi:10.20944/preprints202102.0579.v1

3.9.5. Analysis of the immunity of the peptide 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**).

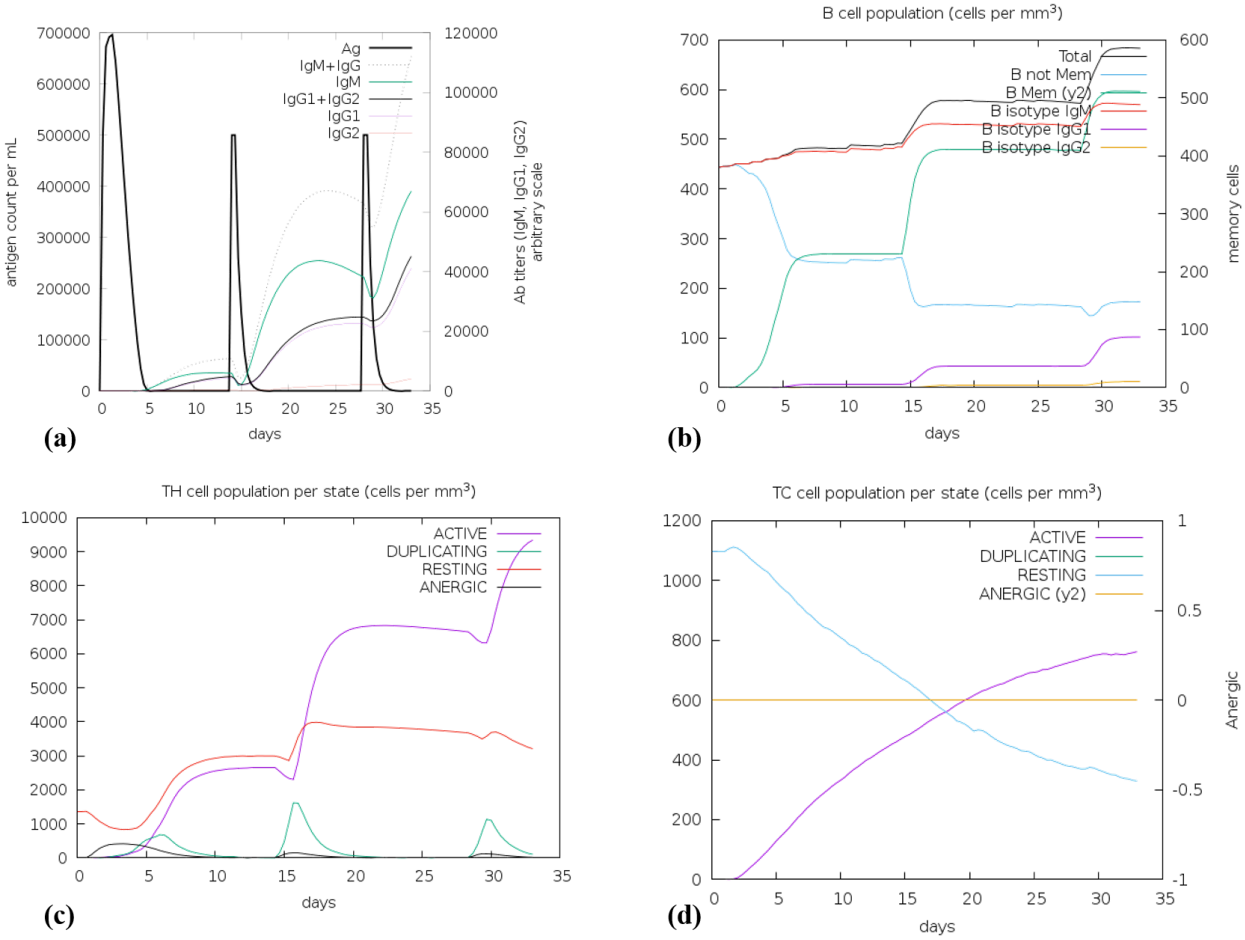

Fig. 7. The immune simulations of the peptide vaccine. Figures a, b, c and d represent the enhanced immune profile of the protein with high concentrations of antigens and immunoglobulins on secondary immune responses.

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 [62-64]. 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 [65-70]. While the lack of vaccines for superbugs was not even an area of interest for academia as well as an industry until recently [71]. The fact that no preventive vaccine is available for protecting the susceptible patients was the main reason for us to look into the pathogens genome/proteome for finding solutions to this challenge [72-74]. 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*.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 February 2021 doi:10.20944/preprints202102.0579.v1

Since the dawn of genomics and the availability of publically accessible genomic data and analysis tools, many different studies have been done to expedite the vaccine and drug development against noticeable pathogens [75]. The most prominent one was the case of the meningococcal group B vaccine which is currently available in markets [15]. Unfortunately, the situation is quite adverse for *P. aeruginosa* in terms of both prevention and treatment [74].

In our previous study, we have presented an overview of all the vaccine development studies and reports [5]. 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 [76]. 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 [77]. 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, 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 **(Supplementary Table S2)**.

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 [14]. 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.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 February 2021 doi:10.20944/preprints202102.0579.v1

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 ($CupB_5$) 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 [78]. 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 [79]. 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* [80]. 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 [81]. 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-1L, both of which are important virulence factors [82].

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 [83]. 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/II human clinical trial study in 2003 for OprF-OprI epitopes based vaccines [84].

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 [85]. 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

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 February 2021 doi:10.20944/preprints202102.0579.v1

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 [86]. 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 [87]. 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 [92]. 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) [93], 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 [88]. 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 [89]. 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

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 February 2021 doi:10.20944/preprints202102.0579.v1

group focused on Constraint-based flux balance analysis for predicting novel targets for *P. aeruginosa* [90]. 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

In the study, we report novel preventive as well as therapeutic targets against *P. aeruginosa* an emerging top priority pathogen. The targeted proteins play a fundamental role in bacterial survival, pathogenesis, and progression of the infection. We screened 8 epitopes, being components of five essential proteins with the highest immunogenic potential among the selected protein pool. The potent vaccine candidates have been computationally prioritized and modeled to devise a multi-epitopic vaccine construct that could be further tested in-vivo. 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 semialdehyde dehydrogenase (Q51344), 2-amino-4-hydroxy-6-hydroxymethyldihydropteridine 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-hydroxylauroyl] 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.

Conflict of interest

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

References

1. *Global Priority List of Antibiotic Research Bacteria to Guide Research, Discovery, and Development of New Antibiotics.* ; World Health Organization: 25th Feb. 2017, 2017.
2. Wang, K.; Chen, Y.-q.; Salido, M. M.; Kohli, G. S.; Kong, J.-l.; Liang, H.-j.; Yao, Z.-t.; Xie, Y.-t.; Wu, H.-y.; Cai, S.-q., The rapid in vivo evolution of Pseudomonas aeruginosa in ventilator-associated pneumonia patients leads to attenuated virulence. *Open biology* **2017,** *7* (9), 170029.
3. Lister, P. D.; Wolter, D. J.; Hanson, N. D., Antibacterial-resistant Pseudomonas aeruginosa: clinical impact and complex regulation of chromosomally encoded resistance mechanisms. *Clinical microbiology reviews* **2009,** *22* (4), 582-610.
4. Rizvi, M.; Ahmad, J.; Khan, F.; Shukla, I.; Malik, A.; Sami, H., Synergy of drug combinations in treating multidrug-resistant Pseudomonas aeruginosa. *The Australasian medical journal* **2015,** *8* (1), 1.
5. Rashid, M. I.; Naz, A.; Ali, A.; Andleeb, S., Prediction of vaccine candidates against Pseudomonas aeruginosa: An integrated genomics and proteomics approach. *Genomics* **2017,** *109* (3), 274-283.
6. Gellatly, S. L.; Hancock, R. E., Pseudomonas aeruginosa: new insights into pathogenesis and host defenses. *Pathogens and disease* **2013,** *67* (3), 159-173.
7. Yang, F.; Gu, J.; Yang, L.; Gao, C.; Jing, H.; Wang, Y.; Zeng, H.; Zou, Q.; Lv, F.; Zhang, J., Protective efficacy of the trivalent Pseudomonas aeruginosa vaccine candidate PcrV-OprI-Hcp1 in murine pneumonia and burn models. *Scientific reports* **2017,** *7* (1), 3957.
8. Waters, V.; Smyth, A., Cystic fibrosis microbiology: advances in antimicrobial therapy. *Journal of cystic fibrosis* **2015,** *14* (5), 551-560.
9. Priebe, G. P.; Goldberg, J. B., Vaccines for Pseudomonas aeruginosa: a long and winding road. *Expert review of vaccines* **2014,** *13* (4), 507-19.
10. François, B.; Luyt, C.-E.; Dugard, A.; Wolff, M.; Diehl, J.-L.; Jaber, S.; Forel, J.-M.; Garot, D.; Kipnis, E.; Mebazaa, A., 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. *Critical care medicine* **2012,** *40* (8), 2320-2326.
11. Westritschnig, K.; Hochreiter, R.; Wallner, G.; Firbas, C.; Schwameis, M.; Jilma, B., 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. *Human vaccines & immunotherapeutics* **2014,** *10* (1), 170-183.
12. Vincent, J.-L., Vaccine development and passive immunization for Pseudomonas aeruginosa in critically ill patients: a clinical update. *Future microbiology* **2014,** *9* (4), 457-463.
13. Kanampalliwar, A.; Rajkumar, S.; Girdhar, A.; Archana, T. Reverse Vaccinology: Basics and Applications. J Vaccines Vaccin 4: 194. doi: 10.4172/2157-7560.1000 194 Page 2 of 5 Volume 4• Issue 6• 1000194 J Vaccines Vaccin ISSN: 2157-7560 JVV an open access journal how it is important for the preparation of vaccines against pathogens which are difficult to culture in laboratory. Until now vaccines against some diseases like AIDS, hepatitis C are not present. Reverse vaccinology can make a breakthrough for this. Is reverse vaccinology a hope for preparation of the vaccine against such pathogens or a myth?? Evidently, vaccines against some disease have been provoked due to infection and death of the individuals after the vaccination. these vaccines were prepared by using the conventional approach that …, 2013.
14. Delany, I.; Rappuoli, R.; Seib, K. L., Vaccines, reverse vaccinology, and bacterial pathogenesis. *Cold Spring Harbor perspectives in medicine* **2013,** *3* (5), a012476.
15. Giuliani, M. M.; Adu-Bobie, J.; Comanducci, M.; Aricò, B.; Savino, S.; Santini, L.; Brunelli, B.; Bambini, S.; Biolchi, A.; Capecchi, B., A universal vaccine for serogroup B meningococcus. *Proceedings of the National Academy of Sciences* **2006,** *103* (29), 10834-10839.
16. Brunham, R. C.; Plummer, F. A.; Stephens, R. S., Bacterial antigenic variation, host immune response, and pathogen-host coevolution. *Infection and immunity* **1993,** *61* (6), 2273.

17. Montor, W. R.; Huang, J.; Hu, Y.; Hainsworth, E.; Lynch, S.; Kronish, J.-W.; Ordonez, C. L.; Logvinenko, T.; Lory, S.; LaBaer, J., Genome-wide study of Pseudomonas aeruginosa outer membrane protein immunogenicity using self-assembling protein microarrays. *Infection and immunity* **2009,** *77* (11), 4877-4886.
18. Mosquera-Rendón, J.; Rada-Bravo, A. M.; Cárdenas-Brito, S.; Corredor, M.; Restrepo-Pineda, E.; Benítez-Páez, A., Pangenome-wide and molecular evolution analyses of the Pseudomonas aeruginosa species. *BMC genomics* **2016,** *17* (1), 45.
19. Pennington, J. E.; Reynolds, H. Y.; Wood, R. E.; Robinson, R. A.; Levine, A. S., Use of a Pseudomonas aeruginosa vaccine in patients with acute leukemia and cystic fibrosis. *The American journal of medicine* **1975,** *58* (5), 629-636.
20. Hopkins, A. L.; Groom, C. R., The druggable genome. *Nature reviews Drug discovery* **2002,** *1* (9), 727.
21. Adams, C. P.; Brantner, V. V., Estimating the cost of new drug development: is it really $802 million? *Health affairs* **2006,** *25* (2), 420-428.
22. Craik, D. J.; Fairlie, D. P.; Liras, S.; Price, D., The future of peptide-based drugs. *Chemical biology & drug design* **2013,** *81* (1), 136-147.
23. Hedgecoe, A.; Martin, P., The drugs don't work: expectations and the shaping of pharmacogenetics. *Social studies of science* **2003,** *33* (3), 327-364.
24. Lenz, G. R.; Nash, H. M.; Jindal, S., Chemical ligands, genomics and drug discovery. *Drug Discovery Today* **2000,** *5* (4), 145-156.
25. Roses, A. D., Pharmacogenetics and future drug development and delivery. *The Lancet* **2000,** *355* (9212), 1358-1361.
26. Keller, T. H.; Pichota, A.; Yin, Z., A practical view of 'druggability'. *Current opinion in chemical biology* **2006,** *10* (4), 357-361.
27. Bilofsky, H. S.; Christian, B., The GenBank® genetic sequence data bank. *Nucleic acids research* **1988,** *16* (5), 1861-1863.
28. Bairoch, A.; Apweiler, R., The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic acids research* **2000,** *28* (1), 45-48.
29. Li, W.; Godzik, A., Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **2006,** *22* (13), 1658-1659.
30. Xiang, Z.; He, Y., Vaxign: a web-based vaccine target design program for reverse vaccinology. *Procedia in Vaccinology* **2009,** *1* (1), 23-29.
31. Yu, C. S.; Lin, C. J.; Hwang, J. K., Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n-peptide compositions. *Protein science* **2004,** *13* (5), 1402-1406.
32. Tusnady, G. E.; Simon, I., The HMMTOP transmembrane topology prediction server. *Bioinformatics* **2001,** *17* (9), 849-50.
33. Sachdeva, G.; Kumar, K.; Jain, P.; Ramachandran, S., SPAAN: a software program for prediction of adhesins and adhesin-like proteins using neural networks. *Bioinformatics* **2004,** *21* (4), 483-491.
34. Li, L.; Stoeckert, C. J.; Roos, D. S., OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome research* **2003,** *13* (9), 2178-2189.
35. Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J., Basic local alignment search tool. *Journal of molecular biology* **1990,** *215* (3), 403-410.
36. Calabriso, N.; Scoditti, E.; Massaro, M.; Pellegrino, M.; Storelli, C.; Ingrosso, I.; Giovinazzo, G.; Carluccio, M. A., 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. *European journal of nutrition* **2016,** *55* (2), 477-489.
37. Winsor, G. L.; Lam, D. K.; Fleming, L.; Lo, R.; Whiteside, M. D.; Yu, N. Y.; Hancock, R. E.; Brinkman, F. S., Pseudomonas Genome Database: improved comparative analysis and population genomics capability for Pseudomonas genomes. *Nucleic Acids Res* **2011,** *39* (Database issue), D596-600.

38. Luo, H.; Lin, Y.; Gao, F.; Zhang, C.-T.; Zhang, R., DEG 10, an update of the database of essential genes that includes both protein-coding genes and noncoding genomic elements. *Nucleic Acids Res* **2014,** *42* (D1), D574-D580.
39. Weigt, S.; Wang, X.; Palchevskiy, V.; Patel, N.; Ross, D.; Shino, Y.; Reynolds, J.; Shah, P.; Singer, L.; Budev, M., Gene Expression Profiling of BAL Cells and PBMC During Lung Allograft Acute Cellular Rejection. In *C32. LUNG TRANSPLANTATION: CLINICAL STUDIES*, American Thoracic Society: 2018; pp A4740-A4740.
40. Chen, L.; Xiong, Z.; Sun, L.; Yang, J.; Jin, Q., VFDB 2012 update: toward the genetic diversity and molecular evolution of bacterial virulence factors. *Nucleic Acids Res* **2011**, gkr989.
41. Garg, A.; Gupta, D., VirulentPred: a SVM based prediction method for virulent proteins in bacterial pathogens. *BMC bioinformatics* **2008,** *9* (1), 62.
42. Doytchinova, I. A.; Flower, D. R., VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC bioinformatics* **2007,** *8* (1), 4.
43. Gasteiger, E.; Hoogland, C.; Gattiker, A.; Wilkins, M. R.; Appel, R. D.; Bairoch, A., *Protein identification and analysis tools on the ExPASy server*. Springer: 2005.
44. Berman, H.; Henrick, K.; Nakamura, H.; Markley, J. L., The worldwide Protein Data Bank (wwPDB): ensuring a single, uniform archive of PDB data. *Nucleic acids research* **2006,** *35* (suppl_1), D301-D303.
45. Arnold, K.; Bordoli, L.; Kopp, J.; Schwede, T., The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics (Oxford, England)* **2006,** *22* (2), 195-201.
46. Källberg, M.; Wang, H.; Wang, S.; Peng, J.; Wang, Z.; Lu, H.; Xu, J., Template-based protein structure modeling using the RaptorX web server. *Nature protocols* **2012,** *7* (8), 1511.
47. Volkamer, A.; Kuhn, D.; Rippmann, F.; Rarey, M., DoGSiteScorer: a web server for automatic binding site prediction, analysis and druggability assessment. *Bioinformatics* **2012,** *28* (15), 2074-2075.
48. Szklarczyk, D.; Franceschini, A.; Kuhn, M.; Simonovic, M.; Roth, A.; Minguez, P.; Doerks, T.; Stark, M.; Muller, J.; Bork, P., The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Res* **2011,** *39* (suppl 1), D561-D568.
49. He, Y.; Xiang, Z.; Mobley, H. L., Vaxign: the first web-based vaccine design program for reverse vaccinology and applications for vaccine development. *J Biomed Biotechnol* **2010,** *2010*, 297505.
50. Paul, S.; Sidney, J.; Sette, A.; Peters, B., TepiTool: A Pipeline for Computational Prediction of T Cell Epitope Candidates. *Curr Protoc Immunol* **2016,** *114*, 18 19 1-18 19 24.
51. Guan, P.; Doytchinova, I. A.; Zygouri, C.; Flower, D. R., MHCPred: a server for quantitative prediction of peptide–MHC binding. *Nucleic Acids Res* **2003,** *31* (13), 3621-3624.
52. Mayrose, I.; Penn, O.; Erez, E.; Rubinstein, N. D.; Shlomi, T.; Freund, N. T.; Bublil, E. M.; Ruppin, E.; Sharan, R.; Gershoni, J. M.; Martz, E.; Pupko, T., Pepitope: epitope mapping from affinity-selected peptides. *Bioinformatics* **2007,** *23* (23), 3244-6.
53. Saha, S.; Raghava, G. P., AlgPred: prediction of allergenic proteins and mapping of IgE epitopes. *Nucleic Acids Res* **2006,** *34* (Web Server issue), W202-9.
54. Dimitrov, I.; Bangov, I.; Flower, D. R.; Doytchinova, I., AllerTOP v.2--a server for in silico prediction of allergens. *J Mol Model* **2014,** *20* (6), 2278.
55. Smialowski, P.; Doose, G.; Torkler, P.; Kaufmann, S.; Frishman, D., PROSO II--a new method for protein solubility prediction. *FEBS J* **2012,** *279* (12), 2192-200.
56. Hebditch, M.; Carballo-Amador, M. A.; Charonis, S.; Curtis, R.; Warwicker, J., Protein-Sol: a web tool for predicting protein solubility from sequence. *Bioinformatics* **2017,** *33* (19), 3098-3100.
57. Buchan, D. W. A.; Jones, D. T., The PSIPRED Protein Analysis Workbench: 20 years on. *Nucleic Acids Res* **2019,** *47* (W1), W402-W407.
58. Yang, J.; Zhang, Y., I-TASSER server: new development for protein structure and function predictions. *Nucleic Acids Res* **2015,** *43* (W1), W174-81.

59. Heo, L.; Park, H.; Seok, C., GalaxyRefine: Protein structure refinement driven by side-chain repacking. *Nucleic Acids Res* **2013,** *41* (Web Server issue), W384-8.
60. Lovell, S. C.; Davis, I. W.; Arendall, W. B., 3rd; de Bakker, P. I.; Word, J. M.; Prisant, M. G.; Richardson, J. S.; Richardson, D. C., Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins* **2003,** *50* (3), 437-50.
61. de Vries, S. J.; van Dijk, M.; Bonvin, A. M., The HADDOCK web server for data-driven biomolecular docking. *Nat Protoc* **2010,** *5* (5), 883-97.
62. Pires, D. P. P.; Costa, A. R.; Duarte, A.; Monteiro, R.; Azeredo, J., Bacteriophage interaction with P. aeruginosa and A. baumannii biofilms and colonizing human airway epithelium. **2018**.
63. Moradali, M. F.; Ghods, S.; Rehm, B. H., Pseudomonas aeruginosa lifestyle: a paradigm for adaptation, survival, and persistence. *Frontiers in cellular and infection microbiology* **2017,** *7*, 39.
64. Juan, C.; Peña, C.; Oliver, A., Host and pathogen biomarkers for severe Pseudomonas aeruginosa infections. *The Journal of infectious diseases* **2017,** *215* (suppl_1), S44-S51.
65. Xipell, M.; Bodro, M.; Marco, F.; Martínez, J. A.; Soriano, A., Successful treatment of three severe MDR or XDR Pseudomonas aeruginosa infections with ceftolozane/tazobactam. *Future microbiology* **2017,** *12* (14), 1323-1326.
66. McLean, K.; Lee, D.; Holmes, E. A.; Penewit, K.; Waalkes, A.; Ren, M.; Lee, S. A.; Gasper, J.; Manoil, C.; Salipante, S. J., Genomic analysis identifies novel Pseudomonas aeruginosa resistance genes under selection during inhaled aztreonam therapy in vivo. *Antimicrobial agents and chemotherapy* **2019,** *63* (9), e00866-19.
67. Pang, Z.; Raudonis, R.; Glick, B. R.; Lin, T.-J.; Cheng, Z., Antibiotic resistance in Pseudomonas aeruginosa: mechanisms and alternative therapeutic strategies. *Biotechnology advances* **2019,** *37* (1), 177-192.
68. Frieri, M.; Kumar, K.; Boutin, A., Antibiotic resistance. *Journal of infection and public health* **2017,** *10* (4), 369-378.
69. Qiao, M.; Ying, G.-G.; Singer, A. C.; Zhu, Y.-G., Review of antibiotic resistance in China and its environment. *Environment international* **2018,** *110*, 160-172.
70. Karkman, A.; Do, T. T.; Walsh, F.; Virta, M. P., Antibiotic-resistance genes in waste water. *Trends in microbiology* **2018,** *26* (3), 220-228.
71. Rappuoli, R.; Bloom, D. E.; Black, S., Deploy vaccines to fight superbugs. Nature Publishing Group: 2017.
72. Rashid, M. I. Study of Pyocyanin Induced Virulence & Pathogenicity of Nosocomial Infectious Agent Pseudomonas Aeruginosa and its Prevention. National University of Science & Technology, Islamabad (NUST), 2019.
73. Baker, S. M.; McLachlan, J. B.; Morici, L. A., Immunological considerations in the development of Pseudomonas aeruginosa vaccines. *Human vaccines & immunotherapeutics* **2019**, 1-7.
74. Priebe, G. P.; Meluleni, G. J.; Coleman, F. T.; Goldberg, J. B.; Pier, G. B., Protection against fatal Pseudomonas aeruginosa pneumonia in mice after nasal immunization with a live, attenuated aroA deletion mutant. *Infection and immunity* **2003,** *71* (3), 1453-1461.
75. Sette, A.; Rappuoli, R., Reverse vaccinology: developing vaccines in the era of genomics. *Immunity* **2010,** *33* (4), 530-541.
76. Dalsass, M.; Brozzi, A.; Medini, D.; Rappuoli, R., Comparison of open-source reverse vaccinology programs for bacterial vaccine antigen discovery. *Frontiers in immunology* **2019,** *10*, 113.
77. Ranganathan, S.; Garg, G., Secretome: clues into pathogen infection and clinical applications. *Genome medicine* **2009,** *1* (11), 113.
78. Juhas, M., Pseudomonas aeruginosa essentials: an update on investigation of essential genes. *Microbiology* **2015,** *161* (11), 2053-2060.
79. Foged, C.; Hansen, J.; Agger, E. M., License to kill: Formulation requirements for optimal priming of CD8+ CTL responses with particulate vaccine delivery systems. *European Journal of Pharmaceutical Sciences* **2012,** *45* (4), 482-491.

80. Sorichter, S.; Baumann, U.; Baumgart, A.; Walterspacher, S.; von Specht, B.-U., Immune responses in the airways by nasal vaccination with systemic boosting against Pseudomonas aeruginosa in chronic lung disease. *Vaccine* **2009,** *27* (21), 2755-2759.
81. Fito-Boncompte, L.; Chapalain, A.; Bouffartigues, E.; Chaker, H.; Lesouhaitier, O.; Gicquel, G.; Bazire, A.; Madi, A.; Connil, N.; Véron, W., Full virulence of Pseudomonas aeruginosa requires OprF. *Infection and immunity* **2011,** *79* (3), 1176-1186.
82. Lücking, H.; Blum, B.; Schmitt, A.; Hungerer, K.; Domdey, H., The Pseudomonas aeruginosa outer membrane protein I vaccine: immunogenicity and safe administration in man. *Behring Institute Mitteilungen* **1997,** (98), 326-337.
83. Knapp, B.; Hundt, E.; Lenz, U.; Hungerer, K.-D.; Gabelsberger, J.; Domdey, H.; Mansouri, E.; Li, Y.; von Specht, B.-U., A recombinant hybrid outer membrane protein for vaccination against Pseudomonas aeruginosa. *Vaccine* **1999,** *17* (13-14), 1663-1666.
84. Mansouri, E.; Blome-Eberwein, S.; Gabelsberger, J.; Germann, G.; von Specht, B.-U., Clinical study to assess the immunogenicity and safety of a recombinant Pseudomonas aeruginosa OprF-OprI vaccine in burn patients. *FEMS Immunology & Medical Microbiology* **2003,** *37* (2-3), 161-166.
85. Ni, Z.; Chen, Y.; Ong, E.; He, Y., Antibiotic resistance determinant-focused Acinetobacter baumannii vaccine designed using reverse vaccinology. *International journal of molecular sciences* **2017,** *18* (2), 458.
86. Harmer, C.; Alnassafi, K.; Hu, H.; Elkins, M.; Bye, P.; Rose, B.; Cordwell, S.; Triccas, J. A.; Harbour, C.; Manos, J., Modulation of gene expression by Pseudomonas aeruginosa during chronic infection in the adult cystic fibrosis lung. *Microbiology* **2013,** *159* (Pt_11), 2354-2363.
87. Bavandpour, A. K.; Bakhshi, B.; Najar-Peerayeh, S., 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. *International journal of pharmaceutics* **2020,** *573*, 118868.
88. Sakharkar, K. R.; Sakharkar, M. K.; Chow, V. T., A novel genomics approach for the identification of drug targets in pathogens, with special reference to Pseudomonas aeruginosa. *In silico biology* **2004,** *4* (3), 355-360.
89. Perumal, D.; Lim, C. S.; Sakharkar, K. R.; Sakharkar, M. K., Differential genome analyses of metabolic enzymes in Pseudomonas aeruginosa for drug target identification. *In silico biology* **2007,** *7* (4, 5), 453-465.
90. Perumal, D.; Samal, A.; Sakharkar, K. R.; Sakharkar, M. K., Targeting multiple targets in Pseudomonas aeruginosa PAO1 using flux balance analysis of a reconstructed genome-scale metabolic network. *Journal of drug targeting* **2011,** *19* (1), 1-13.

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 (uniport.org) using their accession numbers.

Author(s) Contributions

Sidra Irum and Muhammad Ibrahim Rashid designed, performed analysis and wrote the draft. Mahnoor Majid and Zeeshan Mustafa helped in analysis. Amjad Ali edited and revised the draft. Saadia Andleeb reviewed and approved the final draft.

Competing Interests

The author(s) declare no competing interests.

Supplementary Figures/ Supplementary Table Legends

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 February 2021 doi:10.20944/preprints202102.0579.v1

Supplementary Fig. S1 (a,b): (a)The breakup of *Pseudomonas aeruginosa* proteome: Sub-cellular localization obtained from Pseudomonas database. The Pseudomonas database uses a multilevel confidence classification system for subcellular localization: class 1 is the experimentally demonstrated localization in the same species; class 2 is based on similarity with experimentally determined either orthologous or paralogous genes. BLAST search criteria used are e value of 10e-10 for query within 80-120% of subject length while class 3 contains computationally predicted subcellular localization by PSORTb V3.0 and CELLO servers. **(b) *P. aeruginosa* proteome localization split**: Subcellular localization predicted by Vaxign pipeline and CELLO V2.5. The sorted proteins comprise 4628 proteins with <2 transmembrane helices, 1863 essential proteins, 1032 virulence factors and 403 proteins with extracellular localization. The prioritized 68 proteins fulfill all preliminary criteria for protein selection.

Supplementary Fig. S2. Relative location of prioritized proteins *P. aeruginosa* PAO1 genome: The position of selected potential proteins is indicated by black dots while purple dots indicate curated virulence factors. The genes involved in antibiotic resistance are represented as pink dots. The integers outside circle corresponds to genome position. The orange, green and blue bars represent the algorithms used by server for prediction of genomic islands i.e. SIGI-HMM, IslandPick and IslandPath-DIMOB respectively.

Supplementary Fig. S3 (a,b): (a) Protein-protein interaction (PPI) analysis of putative vaccine candidates: PPI revealed porin proteins **OprD**, **OprF**, OprP, OprB, OprI and OprH functionally interacting with each other. Blue lines indicate gene co-occurrence while green lines represent gene neighborhood. **Pra** protein has a number of virulence factors of *P. aeruginosa* as indirect functional partners i.e. phzH, lasB, lap and napD. **PA1288** and **PA4874** being proteins of unknown functions have been found interacting on molecular functioning grounds with a long chain fatty-acid-CoA ligase fadD2 protein, eraR protein, acyl-CoA dehydrogenase and sporulation protein SpoVR, flagellar hook protein flgK and a unique metal binding pterin Molybdopterin-binding protein respectively. (b**) Protein-protein interaction (PPI) analysis of the screened druggable target proteins.**

Supplementary Fig. S4. Conservation analysis of broad spectrum epitopes amongst *P. aeruginosa* Taxonomic group (TAXID: 136841): Protein sequences of identified targets were aligned using CLC main workbench. Multiple sequence alignment and conserved epitopic regions amongst Taxonomic group (TAXID: 136841) members is indicated in black boxes.

Supplementary Table S1. Prediction of localization and trans-membrane helices of the *P. aeruginosa* proteins.
Supplementary Table S2. Initial screening of the proteins based upon the Vaxign dynamic analysis.
Supplementary Table S3. Filtering of the proteins which fulfill the criteria for prediction of putative vaccine candidates.
Supplementary Table S4. BLASTp analysis of the *P. aeruginosa* proteins for presence in *Homo. sapiens.*
Supplementary Table S5. Analysis of the selected proteins in Pseudomonas Database.
Supplementary Table S6. Analysis of the selected proteins from PATRIC Database.
Supplementary Table S7. List of screened essential drug targets.
Supplementary Table S8. List of selected proteins and their molecular weights.
Supplementary Table S9. List of screened epitopes for the finalized proteins and their antigenicity scores.
Supplementary Table S10. MHC I & MHC II alleles tested in the study.