Brazilian Journal of Biology®

ISSN 1519-6984 (Print)
ISSN 1678-4375 (Online)

Original Article

Integrated computational approaches assisted development of a novel multi-epitope vaccine against MDR *Streptococcus pseudopneumoniae*

Abordagens computacionais integradas que auxiliaram no desenvolvimento de uma nova vacina multiepítopo contra *Streptococcus pseudopneumoniae* multirresistente

R. Attar[a*]

[a]University of Jeddah, Faculty of Science, Department of Biology, Jeddah, Saudi Arabia

Abstract

The emergence of antibiotic resistance (AR) in bacteria is becoming an alarming health concern because it allows them to adapt themselves to changing environments. It is possible to prevent the spread of AR in many ways, such as reducing antibiotic misuse in human and veterinary medicine. *Streptococcus pseudopneumoniae* is one of these AR bacterial species that can cause pneumonia in humans and is responsible for high mortality and morbidity rates. It is oval shaped gram-positive bacterium that shows resistance to several antibiotics like penicillin, tetracycline, ciprofloxacin, erythromycin, and co-trimoxazale and no approved vaccine is available to overcome diseases of the pathogen. Thus, substantial efforts are necessary to select protective antigens from a whole genome of pathogens that are easily tested experimentally. The *in silico* designed vaccine was safe and potent in immunizing individuals against the aforementioned pathogens. Herein, we utilized a subtractive genomic approach to identify potential epitope-based vaccine candidates against *S. pseudopneumoniae*. In total, 50850 proteins were retrieved from the NCBI, representing the complete genome of *S. pseudopneumoniae*. Out of the total, CD-HIT analysis identified 1022 proteins as non-redundant and 49828 proteins as redundant and further subjected for subcellular localization in which bulk of proteins was located in the cytoplasm, with seven extracellular proteins (penicillin-binding protein, alpha-amylase, solute-binding protein, hypothetical protein, CHAP domain-containing protein, polysaccharide deacetylase family protein, hypothetical protein). Six immune cells epitopes (SNLQSENDRL, RNDSLQKQAR, NPTTTSEGF, KVKKKNNKK, AYSQGSQKEH, and SVVDQVSGDF) were predicted with the help of the IEDB server. To design a multi-epitopes vaccine these immune cell epitopes were together by GPGPG and adjuvant linker to enhance immune response efficacy. The 3D structure of the designed vaccine was modeled and conducted molecular docking and dynamic simulation studies were to check the binding efficacy with immune cells receptor and dynamic behavior of the docked complex. Finally, we concluded that the designed vaccine construct can provoke a proper and protective immune response against *S. pseudopneumoniae.*

Keywords: *Streptococcus pseudopneumoniae*, immunoinformatics, epitope-based vaccine, docking, MD simulations.

Resumo

O surgimento de resistência a antibióticos (AR) em bactérias tem se tornando uma preocupação sanitária alarmante, uma vez que permite que elas se adaptem a ambientes em constante alteração. É possível prevenir a disseminação da RA de várias maneiras, como reduzir o uso indevido de antibióticos na medicina humana e veterinária. O *Streptococcus pseudopneumoniae* é uma dessas espécies bacterianas de AR que podem causar pneumonia em humanos e são responsáveis por altas taxas de mortalidade e morbidade. É uma bactéria gram-positiva de forma oval que mostra resistência a diversos antibióticos como penicilina, tetraciclina, ciprofloxacina, eritromicina e cotrimoxazale, além disso, nenhuma vacina aprovada está disponível para superar as doenças do patógeno. Assim, esforços substanciais são necessários para selecionar antígenos protetores de todo um genoma de patógenos que são facilmente testados experimentalmente. A vacina projetada *in silico* foi considerada segura e potente na imunização de indivíduos contra os patógenos mencionados. Neste trabalho, utilizamos uma abordagem genômica subtrativa para identificar potenciais candidatos a vacinas baseadas em epítopos contra *S. pseudopneumoniae*. No total, 50.850 proteínas foram recuperadas do NCBI, representando o genoma completo de *S. pseudopneumoniae*. Do total, a análise de CD-HIT identificou 1.022 proteínas como não redundantes e 49.828 proteínas como redundantes e posteriormente submetidas à localização subcelular na qual a maior parte das proteínas estava localizada no citoplasma, com sete proteínas extracelulares (proteína de ligação à penicilina, alfa- amilase, proteína de ligação a soluto, proteína hipotética, proteína contendo domínio CHAP, proteína da família polissacarídeo desacetilase,

*e-mail: rmattar@uj.edu.sa
Received: November 6, 2022 – Accepted: December 24, 2022

proteína hipotética). Seis epítopos de células imunes (SNLQSENDRL, RNDSLQKQAR, NPTTTSEGF, KVKKKNNKK, AYSQGSQKEH e SVVDQVSGDF) foram previstos com a ajuda do servidor IEDB. Para projetar uma vacina de múltiplos epítopos, esses epítopos de células imunes foram reunidos por GPGPG e um ligante adjuvante para aumentar a eficácia da resposta imune. A estrutura 3D da vacina projetada foi modelada e conduzido estudos de *docking* molecular e simulação dinâmica para verificar a eficácia da ligação com o receptor de células imunes e o comportamento dinâmico do complexo *docked*. Finalmente, concluímos que a construção da vacina projetada pode provocar uma resposta imune adequada e protetora contra *S. pseudopneumoniae*.

Palavras-chave: *Streptococcus pseudopneumoniae*, imunoinformática, vacina baseada em epítopo, atracação, simulações MD.

1. Introduction

Vaccination is the most promising and effective therapeutic/prophylactic strategy for different infectious diseases. Along with clean water and sanitation, it can reduce the burden of deadly infectious diseases that can result in millions of deaths (Clem, 2011). The development of high-throughput sequencing technology has made it easy for vaccine to treat not only communicable diseases but also non-communicable such as cancer and neurodegenerative diseases (Georgiev, 2009). In ancient times, Turks and Chinese used to treat infectious diseases like smallpox by applying crusts isolated from smallpox affected patients. These initial attempts were further proceeded experimentally with immunization by Lady Mary Wortely Montagu in 1718 and later on Edward Jenner in 1798. The experiment of Edward Jenner to induce smallpox immunity by using cowpox lesions was a better strategy than all other attempts performed at that time (Goldsby et al., 2003). Due to drug resistance and a lack of common targets against genetically diverse strains, using available drugs became a difficult endeavor (Araújo et al., 2019). Drug-resistance strains of pathogens have become a significant global public health threat (Nalamolu et al., 2019). The most effective way of reducing the number of antibiotic resistance (AR) strains is by developing protective vaccines. There is a pressing demand for suitable therapeutics against these pathogens, particularly a vaccine to protect against antibiotic resistance (AR) pathogenic bacteria. A specific and safe vaccine would be highly beneficial to stop the spread of the AR bacterial pathogens and stop resistance evolution. Additionally, the development of therapeutic monoclonal and polyclonal antibodies has helped in designing pathogen specific vaccines to protect the population from the menace of AR (Reddick and Alto, 2014).

The Pasteur principle of vaccinology is a failure in respect of culturing some infectious pathogens *in vitro* due to the non-availability of suitable culturing media (Bambini and Rappuoli, 2009). This can be witnessed by reduced number of culture based vaccine development in the recent past. For example, this limitation of the Pasteur vaccinology was surfaced while developing vaccine candidates against *Mycobacterium leprae* and *Neisseria meningitides* (Bidmos et al., 2018). Reverse Vaccinology (RV) and subtractive proteomics (SP) study have been effectively used to design a chimeric vaccine against biological weapons that counter the pathogenicity and trigger an effective immune response against the bacteria. The approach to designing peptide-based vaccines provided a high success to stop the mortality and morbidity rate of the pathogens (Bidmos et al., 2018). Using a network of supercomputers connected to major research institutions, the system will provide access to nucleotide and protein sequences, structural information, expression profiles, immunological properties, and functional information. For the development of multicomponent meningococcal serogroup B vaccine (4CMenB), RV has been effectively used (Brooks and Brooks, 2014). Experimental vaccinology is costly and time consuming because finding vaccine targets is a daunting task (Bidmos et al., 2018). RV in this regard is highly useful in screening targets not identified by conventional vaccinology (Rappuoli et al., 2016) which not only saves time but is also straightforward and guides experimentalists in a productive way for successful vaccine design (Jansen and Anderson, 2018). The SP and RV methods can be integrated to identify antigenic epitopes for the design of a chimeric vaccine. *S. pseudopneumoniae* is often associated with chronic obstructive pulmonary and respiratory tract infections. It is oval shaped gram positive bacterium belonging to *viridans* group. The bacterium has been isolated from invasive and non-invasive sites of the host that causes many other illnesses such as middle ear infection, bacteremia, and meningitis. It shows resistance to several antibiotics like penicillin, tetracycline, ciprofloxacin, erythromycin, and cotrimoxazole. The resistance shown by above said organism has high contributed to economic losses and elevated mortality and morbidity. No approved vaccine is available to overcome diseases of the pathogen. Therefore, based on immunoinformatics tools, a study was planned to design a novel multi-epitopes vaccine against the above pathogenic bacteria to boost host immune response, which could decode both humoral and cell-mediated responses.

2. Methodology

The full methodology used for the design of a multi-epitopes vaccine against *S. psuedopnuemoniae* is given in Figure 1.

2.1. S. psuedopnuemoniae genome retrieval

For this study, the NCBI generated a total of 50 *S. psuedopnuemoniae* proteome sequences (Schoch et al., 2020). The proteome data of the strains was further analyzed once all of the data was collected in FASTA format.

2.2. Bacterial Pan-genome Analysis (BPGA)

The BPGA tool was used to process the 50 proteomes and extract the bacteria's core proteome. The core proteins are

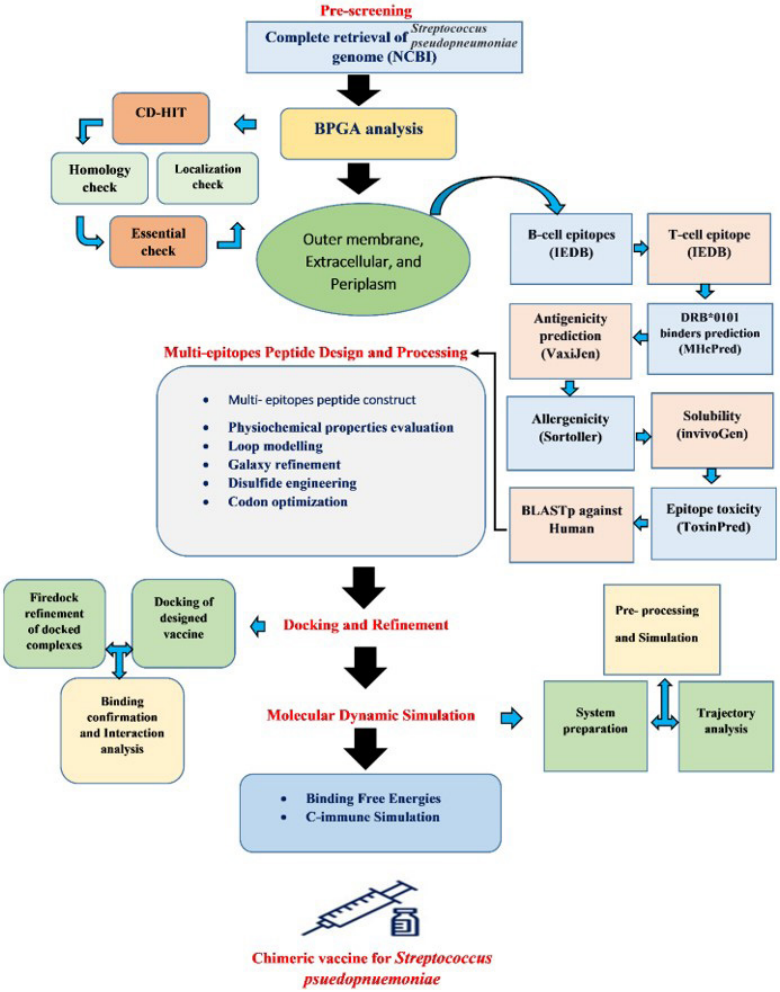

Figure 1. Flow chart diagram of vaccine construct against *S. psuedopnuemoniae*.

a group of sequences that are conserved across all strains. The redundant sequences are duplicates that occur as a result of duplications throughout the evolution process. So, therefore, these sequences are not necessary for the computational vaccine design process (Petersen et al., 2017).

2.3. CD-HIT and localization of subcellular proteins

The CD-HIT server was used to identify pathogen non-redundant core proteins. Proteins with a sequence homology of 60% or more were discarded (Butt et al., 2012). Subcellular localization analyses was performed on all non-redundant proteins. Extracellular, periplasmic, and outer membrane region proteins (OMPs) were identified in the core and non-redundant proteomes, respectively. PSORTb v 3.0 predicted outer membrane, inner membrane, extracellular, cytoplasmic, and unknown proteins (Valguarnera and Feldman, 2017).

2.4. Prediction of immune epitopes

Using the Immune Epitope Database (IEDB), the epitopes prediction step was performed on the selected proteins. The immune system of the host is stimulated/boost by these epitopes. First, we used Bepipred linear epitopes 2.0 to predict linear B-cell epitopes. T-cell epitopes were predicted by assessing B-cell epitopes for their ability to bind to molecules from both classes of major histocompatibility complexes (Vashi et al., 2020). DRB*0101 binding analysis was done using the MHcPred server, and only epitopes with IC50 values less than 100 nM for the DRB*0101 gene were processed for further analysis (Aray et al., 2019).

2.5. Antigenicity, allergenicity, solubility, and toxicity analysis

Vaxijen 2.0 was used to ensure antigenicity. After that, non-allergen sequences were selected using Allertop 2.0, an online allergen prediction tool. All toxic epitopes were deleted to avoid harmful consequences via Toxinpred. Another online webserver, InvivoGen, was employed to obtain solubility prediction data. If the proteins are utilized in vaccine design, the non-homology ensures that no undesired auto-immune responses will occur. Similarly, the non-homology of certain proteins helps to prevent the suppression of helpful probiotic bacteria by mistake (Kar and Srivastava, 2018).

2.6. Multi-epitopes vaccine deigning

The B-cell derived T-cell epitopes were joined using GPGPG linkers. Finally, an EAAAK linker was used to bind the designed epitopes to the N-terminal of the B subunit of the cholera toxin adjuvant (Jia et al., 2020). The linkers used are rigid to allow efficient separation of the epitopes and ease epitopes presentation to host immune receptors. The SCRATCH protein predictor 3Dpro was used to predict the 3D structure of the multi-epitopes vaccine construct. Galaxy Loop was used to design loop structures, which were subsequently refined using Galaxy Refine (Cárdenas, 2021).

2.7. Disulfide engineering and Codon optimization

A disulfide engineering was accomplished using Design 2.0 to achieve vaccine stability, and then reverse translated to DNA sequence using JCat server. Finally, using Snap Gene tool, the vaccine construct was cloned into the pET-28a (+) vector (Peele et al., 2021).

2.8. Molecular docking

Docking methods was used in vaccine development to assess for binding interactions between vaccine molecules and immunological receptors. The immune receptors used in docking analysis were MHC-I (ID: 1AKJ), MHC-II (ID: 3L6F) and TLR-4 (ID; 4G8A). The receptors were retrieved from protein data bank and then used in UCSF Chimera where they are prepared and energy minimized using steepest descent and conjugate gradient algorithm for 1500 steps. Blind docking was performed to predict vaccine interactions with immune receptors using the PATCHDOCK server and for refinement subjected to Firedock (Raj et al., 2021). The RMSD distance used in docking with 4.0 Å. The complex with the lowest global energy was ranked on top and evaluated in UCSF Chimera 1.13 for binding conformation and interactions.

2.9. Molecular dynamic simulation

The dynamics of the constructed vaccine, comprising system creating, pre-processing, and production phases, were also displayed using Assistant model construction with Energy Refinement (AMBER) 20 for 100-ns simulations. The dimensions of a receptor with a vaccine design were constructed using an antechamber tool. The TIP3P solvation box was used to submerge complexes using the Leap software. The CPPTRAJ module of AMBER and the Visual Molecular Dynamics program were used to analyze the simulated trajectories for additional structural evaluation (Wang et al., 2021).

2.10. Docked complexes' binding free energies

The AMBER 20 package and the MMPBSA.py module were used to calculate the MMPBSA binding free energy of vaccine-TLR4. In all, 100 frames were taken from simulated trajectories and submitted to the MM/GBSA equation (Mermelstein et al., 2018).

3. Results

3.1. S. psuedopnuemoniae genome retrieval

A total 50 *S. psuedopnuemoniae* proteomes were obtained from the National Center for Biotechnology Information (NCBI) and subjected to SP analysis. Proteins that are found in the pathogen core proteome, non-host-specific, and are crucial for survival were considered. During this phase, researchers considered vaccine properties found in the literature while assessing protein candidates for vaccine development.

3.2. BPGA analysis

To process the genome, the BPGA tool was used. The major targets for vaccine development were core proteins. During this phase, redundant proteins were not chosen, resulting in biochemical function redundancy.

3.3. CD-HIT analysis

There are 50850 total proteins in the proteome data of strains. Pathogen core proteome, host non-similarity, and essential proteins are among the proteins. CD-HIT analysis identified 1022 proteins as non-redundant and 49828 proteins as redundant as mentioned in Figure 2. Immunoinformatics has been considered non-redundant proteins because of their crucial functions for subcellular localization and virulence analysis (Petersen et al., 2017).

3.4. Localization of subcellular proteins

A subcellular localization technique was used to get surface proteins capable of invasion, adhesion, and proliferation for the prediction of ex-proteome and secretome. Protein localization was determined using the Psortb tool. The bulk of proteins was located in the cytoplasm, with seven extracellular proteins (penicillin-binding protein, alpha-amylase, solute-binding protein, hypothetical protein, CHAP domain-containing protein, polysaccharide deacetylase family protein, hypothetical protein), and a few proteins expected to be unknown. The number of proteins that are surface localized is given in Figure 2.

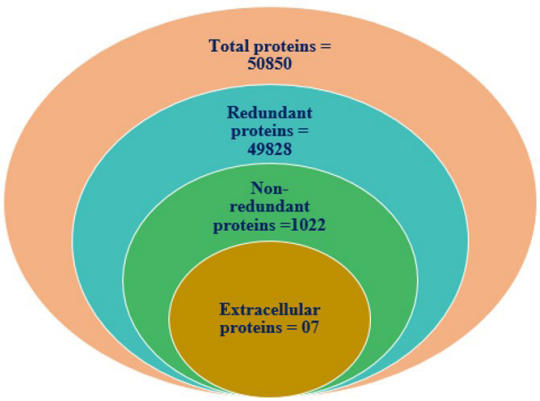

Figure 2. Number of total proteins, redundant proteins, non-redundant proteins, and extracellular proteins.

3.5. Prediction of immune epitopes

Initially, the proteins were used to predict B-cell epitopes, and those with scores greater than 0.8 were chosen. Only epitopes with consistent binding to MHC-I and MHC-II were exposed to a T-cell epitope derived from B cells (Dhanda et al., 2019). Only epitopes with a low percentile rank were chosen. DRB*0101 binding study was done using the MHcPred to find epitopes with IC50 values less than 100 nM for the DRB*0101 gene. The allergenicity, solubility, and toxicity of the epitopes were then determined. The percentile rank IC50 predicted score for all predicted B and T-cell epitopes is listed in Table 1 and Table 2, respectively.

3.6. Antigenicity, allergenicity, solubility, and toxicity analysis

Allergic and non-antigenic epitopes were excluded from these analyses, and the remaining epitopes were chosen for additional research. To prevent the harmful impact of all the toxic epitopes, the solubility was assessed using InvivoGen (Peele et al., 2021) and ToxinPred for toxicity. Six epitopes (SNLQSENDRL, RNDSLQKQAR, NPTTTSEGF, KVKKKNNKK, AYSQGSQKEH, and SVVDQVSGDF) were determined to be soluble and non-toxic and were evaluated for multi-epitopes design. To avoid an autoimmune reaction, the epitopes were aligned with reference human genome and only those epitopes that showed non-homology with

Table 1. B- cell predicted epitopes from selected proteins.

Proteins	B--cell epitopes
>core/307/1/Org1_Gene1774 (penicillin-binding protein)	VSKAPSLSESKLVAT
	LIADLGSERRVNA
	PNQYDPYSHPEAAQN
	YISAEQYEKAVNTSITDGLQSLKSASN
	WDIYNTDEYVAYPDD
	IDYPSLHYSNAISSN
	DGSEKELSNVGTRAM
	YTDEEIENHIKTSQFV
	YRNGQFVFQNGARPTWTETTSQSSSTENSSTSTEGSTSLAPTTSPDASNNSQNQAPATSPDASTNGQNQAPNTPGANQTPAPQTPQQQ
>core/700/1/Org1_Gene832 (alpha-amylase)	LGEFNQKGTVRTKYGFKEDYL
	ADHREAFQVIEVDPVDRTVELGEPFIINGWTSFTFDGRQDTYNDFHWHWYHFTGTDYDAKRRKSGIYLIQGDNKGWANEELVDNENGNYDYLM
	PDKEANLDYLEK
	DLAYGEQTDYFD
>core/946/1/Org1_ Gene85(solute-binding protein)	SKTADKPAESGSSEA
	GDQLGGLDKLSLDNQSG
	QLSEVKLSDGAKTEDTTKSLVT
	VFGQNGKDPKDIGLAND
	EKWPKGMQDGTAAGNL
	EIPANTEARSYAEGKNDE
	FKNAQPIPNISQMSA
>core/5127/1/Org1_ Gene385(hypothetical protein)	QLDDKGRAQVTRYHEKHSKGGAGKKERLLNLREQFLNK
>core/5047/3/Org3_ Gene306(hypothetical protein)	PISFRDADGNFVSAADVWNEKKLEELFNRLNPKRALRLARTKKEET
>core/995/4/Org4_ Gene274(CHAP domain-containing protein)	HAETTDDKITAQDKKISNLTAQQQEAQQQVDQIQEQVSAIQTEQSNLQSENDRLQAESKKLEGEITELSKNIVSRNDSLQKQARSAQTNGAATNYINTVVNSKSITEAISRVA
	GGNRTLGNKRGWFNPTTTSEGFV
>core/753/2/Org2_ Gene936(polysaccharide deacetylase family protein)	KRLPQNKVKKKNNKK
	VLNKQAFEQKIESLKKEKEDAYSQGSQKEHFRKE
	PLVGDNLISSVKDIIVKDITDKVEGKE
	YSEKGDSSLTGVENRLIKKQAYD
	TLDQLYLKEDGSIFTLDQFFTDPSTAKEKILEGVKSTLQDKKVDQSVVDQVSGDFSAAELSSWKFAYKD
	AMTNVEEIAMPISDFFDYIQTSYLTEKDAELYKKVQEEKH
	PVLTKLSLEDAKK
	SDDIRNSLDLSFIMWDVDSMDWKSKNEA
	LGNHVKPHEIYYS

Table 2. T- cell predicted epitopes with percentile rank and MHC Pred with IC_{50}.

MHC-I		MHC-II		MHC Pred (IC_{50}	
SNLQSENDRL	3	SNLQSENDRLQAES	33	NLQSENDRL	39.72
RNDSLQKQAR	2.2	NIVSRNDSLQKQAR	17	NDSLQKQAR	29.99
NPTTTSEGF	0.09	RGWFNPTTTSEGFV	14	NPTTTSEGF	58.88
KVKKKNNKK	0.01	PQNKVKKKNNKK	9.8	KVKKKNNKK	47.86
AYSQGSQKEH	2	AYSQGSQKEHFRKE	40	YSQGSQKEH	32.73
SVVDQVSGDF	0.08	KVDQSVVDQVSGDF	21	SVVDQVSGD	21.78

Table 3. Analysis of antigenicity, allergenicity, toxicity, solubility, and homology check with human.

Epitopes	Antigenicity	Allergenicity	Toxicity	Solubility	Human
NLQSENDRL	1.5005	Non-Allergen	Non-Toxin	Soluble	Non-Similar
NDSLQKQAR	0.9				
NPTTTSEGF	1.0759				
KVKKKNNKK	2.0805				
YSQGSQKEH	1.5411				
SVVDQVSGD	1.1113				

the human genome were chosen. The epitopes that satisfy all of the epitope criteria are tabulated in Table 3.

3.7. Construction of multi-epitopes vaccine

To generate strong and protective immune responses, a multi-epitopes base vaccine construct was used, rather than a single epitope. GPGPG linkers were used to join all of the top 10 screen epitopes in the multi-epitope vaccine. An adjuvant (Cholera Toxin B subunit) was added to the potential vaccine in order to enhance immune response efficacy. With the help of 3Dpro SCRATCH, the 3D structure of the multi-epitope construction was predicted, as illustrated in Figure 3. Following loops were modeled: Met1-lys5, Tyr18-Thr22, Ala31-Ile38, Leu5-Thr6, Phe63-Pro74, Thr99-Lys105, Lys137-Lys150, Glu151-Gly157, Ser158-Pro170, Ser 175-Gly 183. Then, the modelled structure was subjected to refinement in galaxy online services for refining 2. The vaccine construct antigenicity and allerginicity were revisited and found to have antigenic value of 1.21 and non-allergic, respectively.

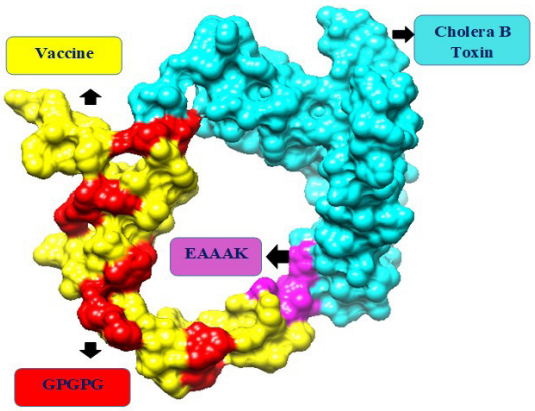

Figure 3. Schematic diagram of vaccine consisting of selected epitopes with GPGPG linkers and linked with cholera B toxin through EAAAK.

3.8. Disulfide engineering

Some of the peptide residues in multi-peptide vaccines can be degraded by enzymes. Disulfide engineering of vaccine constructs was used to replace the degradable prone areas with disulfide links to minimize this issue. Cysteine bonds have been replaced for all residues that are sensitive to enzymatic degradation as shown in Figure 4.

3.9. Codon optimization

Following the above steps, With the Java Codon Adaptation Tool (JCat), the sequence from the designed vaccine construct was reverse translated to a DNA sequence in order to find the highest level of vaccine expression in a vector based on codon adaptation index (CAI) and GC content. SnapGene was used to express the designed vaccine in the pET-28a (+) vector.

3.10. Molecular docking and refinement

The interaction between vaccine and receptors was predicted using blind docking. Top 20 results were generated from the Patch Dock server. After deleting several steric conflicts and intermolecular conformational errors, the FIREDOCK server re-ranked the solutions

Figure 4. Original and mutant vaccines with yellow color indicate cysteine residues.

(Andrusier et al., 2007). Solution 9 in MHC-I, 2 in MHC-II, and 4 in TLR-4 was chosen because it has lowest global energy (-8.30, -13.25, -9.15 kJ•mol1), attractive van der Waals (-21.25, -30.91, -6.35 kJ•mol1), repulsive van der Waals (7.39, 9.55, 2.68 kJ•mol1), atomic contact energy (ACE) (9.96, 10.93, -0.64 kJ•mol1), and hydrogen bonding (-3.36, -3.07, -3.07 kJ•mol1) as given in Table 4. The solutions were subjected to intermolecular structure conformation studies using UCSF Chimera 1.13.1 as shown in Figure 5. The stable vaccine interactions with MHC-I and MHC-II could lead to activation of cytotoxic T cells and helper T-cells, respectively. Similarly, the vaccine presentation by the TLR-4 will activate ND-kB signally pathway and inflammatory cytokine production, responsible for innate immunity as well as adaptive immune responses.

3.11. Chemically interactions between receptors with vaccine

MHC molecules process peptide antigens and present them to immune cells, which is necessary for adaptive immune response. To generate an appropriate immune response, foreign peptide antigens must interact with different types of immune cells before antigen processing and presentation. These intermolecular interactions between MHC-I, MHC-II, and TLR-4 are crucial to comprehend since they revealed residues that are required for vaccine identification. Multiple critical amino acid residues of immune cell receptor molecules interacted with the model vaccine design as studied in UCSF chimera and listed in Table 5.

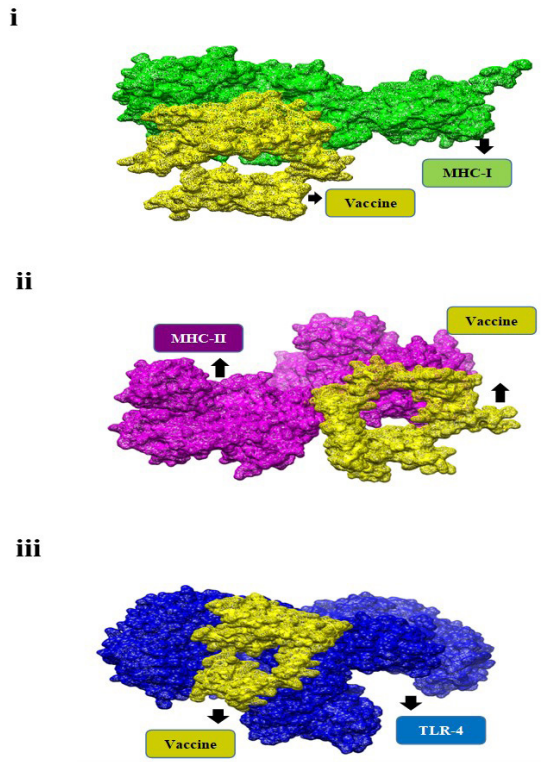

Figure 5. 3D model of docked complexes of vaccine with receptors. i) vaccine + MHC-I complex, ii) vaccine + MHC-II complex, and iii) vaccine + TLR-4 complex.

3.12. Molecular dynamic simulation

The dynamic behavior of macromolecules is evaluated via molecular dynamic simulation analysis (Hansson et al., 2002). To assess the structural stability of the systems, molecular dynamics simulations of docked complexes were run for 100 nanoseconds (ns). The root mean square deviation (RMSD) graph plot is stable, with no significant structural variations as shown in Figure 6A. The residue flexibility of receptors with the vaccine molecule was studied by root mean square fluctuation (RMSF) and it was concluded that no significant variations occur in any systems as represented in Figure 6B.

3.13. Hydrogen bonding

The hydrogen bond is formed when the electrons of the acceptor and donor share between each other. Fluorine, oxygen, and nitrogen can transfer hydrogen or accept it in the majority of cases. VMD was used to count the

Table 4. Lists the docking scores of the refine docked complexes of MHC-I, MHC-II, and TLR-4.

Rank	Solution Number	Global Energy	Attractive VdW	Repulsive VdW	ACE	Hydrogen bonding (HB)
MHC-I						
1	2	-13.25	-30.91	9.55	10.93	-3.07
2	9	5.70	-23.17	31.93	9.35	-2.32
3	3	8.18	-1.09	0.33	0.68	0.00
4	5	20.16	-1.79	0.59	2.10	0.00
5	4	26.60	-23.93	50.20	0.30	-1.61
6	10	968.95	-35.64	1250.15	3.21	-2.75
7	8	3245.59	-63.43	4121.74	14.60	-6.95
8	6	3353.30	-37.07	4266.23	1.75	-5.13
MHC-II						
1	9	-8.30	-21.25	7.39	9.96	-3.36
2	4	-0.78	-4.85	2.21	0.65	-1.91
3	5	2.16	-5.29	1.33	3.10	0.00
4	6	2.59	-1.47	0.00	0.77	0.00
5	1	8.97	-2.40	0.00	0.49	0.00
6	3	19.85	-26.16	16.06	8.73	0.00
7	7	31.12	-27.93	36.35	12.19	-5.24
8	10	205.52	-38.25	249.36	17.90	-2.40
9	2	1770.11	-26.55	2255.76	-3.17	-1.25
10	8	4303.02	-85.82	5547.87	6.94	-12.4
TLR-4						
1	4	-9.15	-6.35	2.68	-0.64	-0.89
2	7	-1.85	-20.91	5.83	14.45	-2.13
3	6	-1.19	-6.52	1.95	1.54	0.00
4	2	10.70	-26.09	13.34	14.76	-1.28
5	5	11.42	-2.80	0.12	1.43	-0.66
6	9	22.93	-22.15	11.83	7.95	-1.27
7	10	308.68	-35.13	432.86	8.06	-7.37
8	1	468.62	-71.10	667.71	23.32	-9.11
9	8	544.41	-49.44	743.36	6.97	-7.77
10	3	2956.89	-50.94	3735.59	11.13	-5.07

Table 5. Chemical wise interaction between vaccine and receptors.

Receptors	Residues
MHC-I	Arg 194, Asp 166, Gln 24, Glu 101,His 78, Ile 60, Leu 29, Lys 137, Met 89, Phe 69, Pro 114, Thr 62, Tyr 144, Val 103,
MHC-II	Arg 131, Asn 174, Gln 37, Glu 177, His 115, Lys 137, Met 122, Ser 132, Pro 142, Trp 109, Val 108
TLR-4	Ala193, Asn 200, Gln 24, Glu 206, His 20, Phe 208, Pro 210, Ser 188, Val 33

hydrogen bonds generated during simulation. Between the vaccine receptors complex, a strong clustering pattern of hydrogen bonds was observed. This shows a significant complex formation and a high affinity of the vaccine for receptors as shown in Figure 7.

3.14. Docked complexes' binding free energies

The MM-GBSA and MMPBSA methods are used to compute the binding free energies of docked complexes, as well as to evaluate the docked complexes' binding efficiency. The total binding free energy of TLR-4, MHC-I, and MHC-II complex were -112.14 kcal/mol, -92.26 kcal/mol, and -89.1 kcal/mol, respectively. The van der Waals energy, as well as electrostatic energy, contributed to the net binding energy. Furthermore, the minimal energy contribution from polar salvation was observed as illustrated in Table 6.

4. Discussion

The bacterial evolution process allows bacteria to make themselves resistant to antibiotics. It is possible to generate an adequate immunological response against specific pathogens by vaccination to prevent infection with AR pathogens. *S. pseudopneumoniae* is one of the etiological respiration tract infectious pathogens that

Table 6. Binding free energies between designed vaccine with receptors.

Energy Parameter	TLR-4-Vaccine Complex	MHC-I-Vaccine Complex	MHC-II-Vaccine Complex
MM-GBSA			
VDWAALS	-217	-182	-114
Electrostatic	-181	-131.74	-100
Delta G solv	42.74	44	31.45
Delta Total	-355.26	-269.74	-182.55
MM-PBSA			
VDWAALS	-217	-182	-114
EEL	-181	-131.74	-100
Delta G solv	52.00	30	27.78
Delta Total	-346	-283.74	-186.22

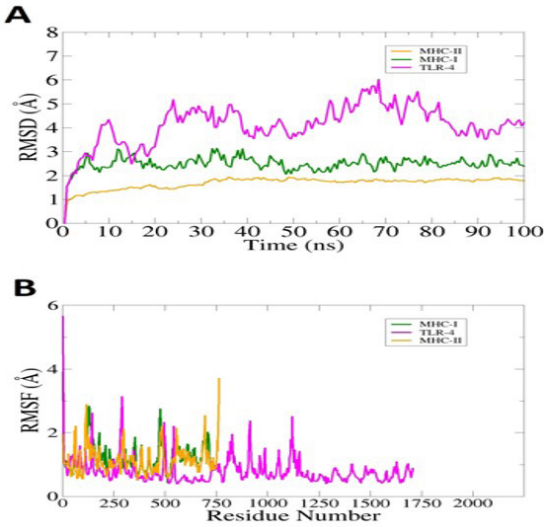

Figure 6. Molecular dynamic simulation analysis three output values are depicted here, (A) RMSD and (B) RMSF.

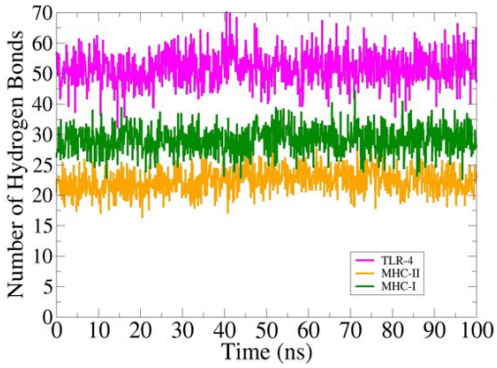

Figure 7. Hydrogen bonding of vaccine with receptor (MHC-I, MHC-II, TLR4) complex.

causes pneumonia belonging to viridans group. It is an oval shaped gram-positive bacterium that shows resistance to several antibiotics like penicillin, tetracycline, ciprofloxacin, erythromycin, and co-cotrimoxazole [3]. The bacterium has been isolated from invasive and non-invasive sites of the host that causes many other illnesses such a middle ear infection, bacteremia, and meningitis with high mortality and morbidity [2]. Therefore, we employed RV approaches to designed vaccine construct against *S. psuedopnuemoniae* to improve health issues. In RV, putative surface-associated proteins are identified without culturing the microorganisms, which is safe, specific, and potent (Abdullah et al., 2021). Pan-genomic reverse vaccination (PGRV) in particular is more effective than conventional RV because it screens highly conserved targets rather than specific ones. Conventional vaccines are time-consuming, expensive, and require a lot of human resources (Caniça et al., 2019).

We suggested RV over Pasteur's rules of vaccinology because of the inability to culture some pathogens *in vitro*, extensive antigenic variability, and molecular mimicry limit the broad applicability of traditional culture-based techniques in the development of vaccines targeting economically important *S. pseudopnueminiae* pathogens. In computational vaccination, immune cell epitopes are predicted, potential and safe antigens are identified and processed, population coverage and conservation are analyzed, toxicity of antigens is predicted, allergenicity is assessed, docking and simulation approaches are utilized, and binding energies are estimated (Ismail et al., 2020). The whole genome was analyzed for core genome containing those sequences/proteins present in all strains and other is dispensable genome having sequences present in one or more individuals but not in all strains. The proteins are non-redundant and their sequence shows non-similarity to human normal flora and further processed them for physicochemical properties (Naz et al., 2015). Different analysis was performed to achieve physiochemical properties like molecular weight, Gravy, solubility index, and transmembrane helicases number through Htmmtop sever. 07 extracellular (penicillin-binding protein, alpha-amylase, solute-binding protein, hypothetical

protein, CHAP domain-containing protein, polysaccharide deacetylase family protein, hypothetical protein) proteins were selected by applying subcellular localization strategy. The B and T-cell epitopes were predicted that stimulate the production of cytokines, interleukin, and immunoglobulin (Vita et al., 2015). MHcPred analysis of the predicated epitopes was performed and observed that these T-cell epitopes are specially bound to HLA alleles which can stimulate cell-mediated and humoral responses. 06 epitopes (SNLQSENDRL, RNDSLQKQAR, NPTTTSEGF, KVKKKNNKK, AYSQGSQKEH, and SVVDQVSGDF) were selected and joined together by GPGPG linker and linked with cholera B toxin via EAAAK to promote enhanced immune responses against the pathogen. The 3D structure was vaccine modeled by the 3DScratch tool and then processed for loop modeling and refinement. Docking was performed to check the binding affinity of the vaccine construct and then proceed for molecular simulation to check the immune responses. The best docked complex with the lowest binding energy and strong stability between protein and receptor (MHC-I, MHC-II, TLR4) was achieved using different servers. The servers include those of molecular dynamics simulation and binding free energies to re-affirm complexes intermolecular stability versus time and highlight atomic level binding energies.

Though multi-epitopes vaccine construction is an efficient approach of vaccine development against infectious pathogens however there are still some challenges that need to be highlighted. The epitopes ordering used in the study required extensive work as the optimal ordering is essential for boosting good immune reactions against the vaccine antigens. Secondly, the selection of adjuvant that can be best fit to the used epitopes is another area on which can be focused. The selection of B and T-cell epitopes either alone or derived form can also be evaluated. Nevertheless, the analyses done in the study concluded good docked stability of the vaccine with immune receptors and efficient representation of the vaccine epitopes to the host immune system. Thus it can be summarized the vaccine is suitable candidate for additional investigations.

5. Conclusions and Limitations

In this study, we developed an antigen-based vaccine against *S. pseudopneumoniae*, one of the most etiological human pathogens and highly resistant to several antibiotics, by combining RV, SP, and immunoinformatics approach. Furthermore, there is no licensed vaccine available for this disease. To create non-allergic, antigenic, non-toxic, and soluble epitopes, nonhomology, non-redundant proteins were used. GPGPG linkers were created between epitopes to link them together and to link with adjuvants via an EAAAK linker to enhance the potency of the vaccine. In order to prevent future outbreaks of *S. pseudopneumoniae* disease, the strongest candidates were prioritized for future vaccine development attempts. As a result of this work, a ready-to-use peptide vaccine will be available for experimentalists to examine its effectiveness *in vivo*. As a result of the study, future vaccine development against *S. pseudopneumoniae* will be accelerated by the study's findings. Additionally, the vaccine is likely to provide cross-protection against all sequenced strains of the bacteria since it is designed based on proteins found in the pathogen's core genome. The study has some limitations that need to be overcome in future research. The first step is to conduct an experimental evaluation in order to determine the best epitope combination in the vaccine construct to maximize immune response levels. Second, methods for predicting MHC molecule epitopes need to refined. Furthermore, testing both *in vivo* and *in vitro* was necessary for the vaccine to provide real immunity protection.

References

ABDULLAH, M., KADIVELLA, M., SHARMA, R., FAISAL, S.M. and AZAM, S., 2021. Designing of multiepitope-based vaccine against Leptospirosis using Immuno-Informatics approaches. *bioRxiv*. In press.

ANDRUSIER, N., NUSSINOV, R. and WOLFSON, H.J., 2007. FireDock: fast interaction refinement in molecular docking. *Proteins*, vol. 69, no. 1, pp. 139-159. http://dx.doi.org/10.1002/prot.21495. PMid:17598144.

ARAÚJO, C.L., ALVES, J., NOGUEIRA, W., PEREIRA, L.C., GOMIDE, A.C., RAMOS, R., AZEVEDO, V., SILVA, A. and FOLADOR, A., 2019. Prediction of new vaccine targets in the core genome of *Corynebacterium pseudotuberculosis* through omics approaches and reverse vaccinology. *Gene*, vol. 702, pp. 36-45. http://dx.doi.org/10.1016/j.gene.2019.03.049. PMid:30928361.

ARAY, Y., AGUILERA-GARCÍA, R. and IZQUIERDO, D.R., 2019. Exploring the nature of the H-bonds between the human class II MHC protein, HLA-DR1 (DRB* 0101) and the influenza virus hemagglutinin peptide, HA306-318, using the quantum theory of atoms in molecules. *Journal of Biomolecular Structure & Dynamics*, vol. 37, no. 1, pp. 48-64. http://dx.doi.org/10.1080/07391102.2017.1418432. PMid:29246090.

BAMBINI, S. and RAPPUOLI, R., 2009. The use of genomics in microbial vaccine development. *Drug Discovery Today*, vol. 14, no. 5-6, pp. 252-260. http://dx.doi.org/10.1016/j.drudis.2008.12.007. PMid:19150507.

BIDMOS, F.A., SIRIS, S., GLADSTONE, C.A. and LANGFORD, P.R., 2018. Bacterial vaccine antigen discovery in the reverse vaccinology 2.0 era: progress and challenges. *Frontiers in Immunology*, vol. 9, p. 2315. http://dx.doi.org/10.3389/fimmu.2018.02315. PMid:30349542.

BROOKS, B.D. and BROOKS, A.E., 2014. Therapeutic strategies to combat antibiotic resistance. *Advanced Drug Delivery Reviews*, vol. 78, pp. 14-27. http://dx.doi.org/10.1016/j.addr.2014.10.027. PMid:25450262.

BUTT, A.M., TAHIR, S., NASRULLAH, I., IDREES, M., LU, J. and TONG, Y., 2012. Mycoplasma genitalium: a comparative genomics study of metabolic pathways for the identification of drug and vaccine targets. *Infection, Genetics and Evolution*, vol. 12, no. 1, pp. 53-62. http://dx.doi.org/10.1016/j.meegid.2011.10.017. PMid:22057004.

CANIÇA, M., MANAGEIRO, V., ABRIOUEL, H., MORAN-GILAD, J. and FRANZ, C.M.A.P., 2019. Antibiotic resistance in foodborne bacteria. *Trends in Food Science & Technology*, vol. 84, pp. 41-44. http://dx.doi.org/10.1016/j.tifs.2018.08.001.

CÁRDENAS, P., 2021. Starting from scratch: a workflow for building truly novel proteins. *Synthetic Biology*, vol. 6, no. 1, p. ysab005. http://dx.doi.org/10.1093/synbio/ysab005. PMid:33758784.

CLEM, A.S., 2011. Fundamentals of vaccine immunology. *Journal of Global Infectious Diseases*, vol. 3, no. 1, pp. 73-78. http://dx.doi.org/10.4103/0974-777X.77299. PMid:21572612.

DHANDA, S.K., MAHAJAN, S., PAUL, S., YAN, Z., KIM, H., JESPERSEN, M.C., JURTZ, V., ANDREATTA, M., GREENBAUM, J.A., MARCATILI, P., SETTE, A., NIELSEN, M. and PETERS, B., 2019. IEDB-AR: immune epitope database— analysis resource in 2019. *Nucleic Acids Research*, vol. 47, no. W1, pp. W502-W506. http://dx.doi.org/10.1093/nar/gkz452. PMid:31114900.

GEORGIEV, V.S., 2009. *National Institute of Allergy and Infectious Diseases, NIH*. Totowa: Humana Press, vol. 2. http://dx.doi.org/10.1007/978-1-60327-297-1.

GOLDSBY, R.A., KINDT, T.J., OSBORNE, B.A. and KUBY, J., 2003. *Immunology*. New York: Freeman.

HANSSON, T., OOSTENBRINK, C. and VAN GUNSTEREN, W., 2002. Molecular dynamics simulations. *Current Opinion in Structural Biology*, vol. 12, no. 2, pp. 190-196. http://dx.doi.org/10.1016/S0959-440X(02)00308-1. PMid:11959496.

ISMAIL, S., AHMAD, S. and AZAM, S.S., 2020. Vaccinomics to design a novel single chimeric subunit vaccine for broad-spectrum immunological applications targeting nosocomial *Enterobacteriaceae* pathogens. *European Journal of Pharmaceutical Sciences*, vol. 146, p. 105258. http://dx.doi.org/10.1016/j.ejps.2020.105258. PMid:32035109.

JANSEN, K.U. and ANDERSON, A.S., 2018. The role of vaccines in fighting antimicrobial resistance (AMR). *Human Vaccines & Immunotherapeutics*, vol. 14, no. 9, pp. 2142-2149. http://dx.doi.org/10.1080/21645515.2018.1476814. PMid:29787323.

JIA, S., HUANG, X., LI, H., ZHENG, D., WANG, L., QIAO, X., JIANG, Y., CUI, W., TANG, L., LI, Y. and XU, Y., 2020. Immunogenicity evaluation of recombinant Lactobacillus casei W56 expressing bovine viral diarrhea virus E2 protein in conjunction with cholera toxin B subunit as an adjuvant. *Microbial Cell Factories*, vol. 19, no. 1, p. 186. http://dx.doi.org/10.1186/s12934-020-01449-3. PMid:33004035.

KAR, P.P. and SRIVASTAVA, A., 2018. Immuno-informatics analysis to identify novel vaccine candidates and design of a multi-epitope based vaccine candidate against Theileria parasites. *Frontiers in Immunology*, vol. 9, p. 2213. http://dx.doi.org/10.3389/fimmu.2018.02213. PMid:30374343.

MERMELSTEIN, D.J., LIN, C., NELSON, G., KRETSCH, R., MCCAMMON, J.A. and WALKER, R.C., 2018. Fast and flexible gpu accelerated binding free energy calculations within the amber molecular dynamics package. *Journal of Computational Chemistry*, vol. 39, no. 19, pp. 1354-1358. http://dx.doi.org/10.1002/jcc.25187. PMid:29532496.

NALAMOLU, R.M., PASALA, C., KATARI, S.K. and AMINENI, U., 2019. Discovery of common putative drug targets and vaccine candidates for *Mycobacterium tuberculosis* sp. *Journal of Drug Delivery and Therapeutics*, vol. 9, no. 2-S, pp. 67-71. http://dx.doi.org/10.22270/jddt.v9i2-s.2603.

NAZ, A., AWAN, F.M., OBAID, A., MUHAMMAD, S.A., PARACHA, R.Z., AHMAD, J. and ALI, A., 2015. Identification of putative vaccine candidates against *Helicobacter pylori* exploiting exoproteome and secretome: a reverse vaccinology based approach. *Infection, Genetics and Evolution*, vol. 32, pp. 280-291. http://dx.doi.org/10.1016/j.meegid.2015.03.027. PMid:25818402.

PEELE, K.A., SRIHANSA, T., KRUPANIDHI, S., AYYAGARI, V.S. and VENKATESWARULU, T.C., 2021. Design of multi-epitope vaccine candidate against SARS-CoV-2: an in-silico study. *Journal of Biomolecular Structure & Dynamics*, vol. 39, no. 10, pp. 3793-3801. http://dx.doi.org/10.1080/07391102.2020.1770127. PMid:32419646.

PETERSEN, M., MEUSEMANN, K., DONATH, A., DOWLING, D., LIU, S., PETERS, R.S., PODSIADLOWSKI, L., VASILIKOPOULOS, A., ZHOU, X., MISOF, B. and NIEHUIS, O., 2017. Orthograph: a versatile tool for mapping coding nucleotide sequences to clusters of orthologous genes. *BMC Bioinformatics*, vol. 18, no. 1, p. 111. http://dx.doi.org/10.1186/s12859-017-1529-8. PMid:28209129.

RAJ, V.S.S., MONISHA, M. and PARAMASIVAM, G., 2021. In-silico screening of synthetic inhibitors for human poly (Adp-ribose) polymerase 2 enzyme using Patch Dock software for ovarian cancer therapy. Revista Gestão. *Inovação e Tecnologias*, vol. 11, no. 1, pp. 5910-5923.

RAPPUOLI, R., BOTTOMLEY, M.J., D'ORO, U., FINCO, O. and GREGORIO, E., 2016. Reverse vaccinology 2.0: human immunology instructs vaccine antigen design. *The Journal of Experimental Medicine*, vol. 213, no. 4, pp. 469-481. http://dx.doi.org/10.1084/jem.20151960. PMid:27022144.

REDDICK, L.E. and ALTO, N.M., 2014. Bacteria fighting back: how pathogens target and subvert the host innate immune system. *Molecular Cell*, vol. 54, no. 2, pp. 321-328. http://dx.doi.org/10.1016/j.molcel.2014.03.010. PMid:24766896.

SCHOCH, C.L., CIUFO, S., DOMRACHEV, M., HOTTON, C.L., KANNAN, S., KHOVANSKAYA, R., LEIPE, D., MCVEIGH, R., O'NEILL, K., ROBBERTSE, B., SHARMA, S., SOUSSOV, V., SULLIVAN, J.P., SUN, L., TURNER, S. and KARSCH-MIZRACHI, I., 2020. NCBI taxonomy: a comprehensive update on curation, resources and tools. *Database*, vol. 2020, p. baaa062. http://dx.doi.org/10.1093/database/baaa062.

VALGUARNERA, E. and FELDMAN, M.F., 2017. Glycoengineered outer membrane vesicles as a platform for vaccine development. *Methods in Enzymology*, vol. 597, pp. 285-310. http://dx.doi.org/10.1016/bs.mie.2017.06.032. PMid:28935107.

VASHI, Y., JAGRIT, V. and KUMAR, S., 2020. Understanding the B and T cell epitopes of spike protein of severe acute respiratory syndrome coronavirus-2: a computational way to predict the immunogens. *Infection, Genetics and Evolution*, vol. 84, p. 104382. http://dx.doi.org/10.1016/j.meegid.2020.104382. PMid:32473352.

VITA, R., OVERTON, J.A., GREENBAUM, J.A., PONOMARENKO, J., CLARK, J.D., CANTRELL, J.R., WHEELER, D.K., GABBARD, J.L., HIX, D., SETTE, A. and PETERS, B., 2015. The immune epitope database (IEDB) 3.0. *Nucleic Acids Research*, vol. 43, no. D1, pp. D405-D412. http://dx.doi.org/10.1093/nar/gku938. PMid:25300482.

WANG, Y., KIZILTAS, A., BLANCHARD, P. and WALSH, T.R., 2021. Calculation of 1D and 2D densities in VMD: a flexible and easy-to-use code. *Computer Physics Communications*, vol. 266, p. 108032. http://dx.doi.org/10.1016/j.cpc.2021.108032.