Computational Design of a Multi-epitope Vaccine Against *Clostridium chauvoei*: An Immunoinformatics Approach

Çiğdem Yılmaz Çolak[1]

Abstract

Blackleg is an infectious disease of animals that is commonly caused by *Clostridium chauvoei* and characterized by localized muscle necrosis. In this study, proteome-mining and immunoinformatics approaches were applied to identify novel antigenic proteins and to construct a multi-epitope vaccine against *C. chauvoei*. All proteins of *C. chauvoei* strains were retrieved from the NCBI Microbial Genome Database containing both genomic and proteomic data of prokaryotes. The proteins were analyzed to exclude non-redundant sequences and to determine antigenic, virulent, and non-allergenic vaccine candidates through several online tools, resulting in seven protein candidates. Cytotoxic T and B cell epitopes of these proteins were evaluated through the tools present in the immune epitope database and the prioritized antigenic epitopes were then conjugated via appropriate linkers to construct the vaccine candidate. After the evaluation of physicochemical properties of the construct, the tertiary structure was modeled and refined through trRosetta and GalaxyRefine, respectively. The quality of the 3D structure was validated by ERRAT score, z-score, and Ramachandran plot and the construct was then docked with bovine Toll-like receptor 4 (TLR 4) using ClusPro. The docked complex was subjected to Molecular Mechanics/Generalized Born Surface Area in the HawkDock server and normal mode analysis in the iMODS simulation suite to assess the binding energy and stability of the complex, respectively. Overall, the vaccine construct was found stable and energetically feasible for bovine TLR 4 binding. Therefore, it can be used as a multi-epitope vaccine construct in clostridial vaccines to control the blackleg disease.

Keywords *Clostridium chauvoei* · Reverse vaccinology · Immunoinformatics · Multi-epitope vaccine

Introduction

In small ruminants, severe toxic infections are caused by *Clostridium* species including *C. chauvoei*, *C. septicum*, *C. novyi*, and *C. sordellii* (also known as *Paeniclostridium sordellii*) mainly due to the contamination of wounds. Blackleg is one of these infectious diseases that is widely observed in cattle and sheep with a high mortality rate. The common causative agent of blackleg is *C. chauvoei* which is a Gram-positive and strict anaerobic bacterium defined as one of the most pathogenic *Clostridium* sp. (Abreu and Uzal 2016). Due to the endospore-forming nature of *C. chauvoei*, it persistently contaminates the soil, which, in turn, causes the infection of grazing ruminants. Although *C. chauvoei*

is thought to be animal-specific, human cases with neutropenic enterocolitis and gas gangrene are also reported in the literature, indicating that animal specificity should be reconsidered (Nagano et al. 2008; Weather and Tweardy 2012). There is limited knowledge about its pathogenesis, but the studies reveal the virulence factors of the pathogen including *C. chauvoei* toxin A (hemolytic leukocidin A) (CctA), flagella, and sialidase that can play important roles in the spread of the pathogen and infection (Disasa et al. 2020). The most studied virulence factor of *C. chauvoei* is CctA which belongs to the alpha-hemolysin channel-forming toxin family and causes cell lysis through membrane disruption (Zaragoza et al. 2019). Moreover, it is a well-conserved toxin among all strains of *C. chauvoei* and has similarities with the pore-forming toxins of *C. perfringens* and *Staphylococcus aureus* (Frey et al. 2012; Abreu and Uzal 2016).

Blackleg is characterized by localized muscle necrosis, especially in the leg muscle tissues, where the disease causes a blackish-red color in the infected area. The common death

✉ Çiğdem Yılmaz Çolak
cigdem.yilmaz@tubitak.gov.tr

1 Genetic Engineering and Biotechnology Institute, TUBITAK, 41470 Gebze, Kocaeli, Turkey

cause due to blackleg is acute toxemia in animals. The disease can be prevented by vaccination and commercial vaccines consist of formalin-inactivated bacterial cultures of *C. chauvoei* formulated alone or with other clostridial ingredients (Disasa et al. 2020). Although vaccination can control the disease, blackleg outbreaks still occur frequently, possibly due to vaccine failures as a result of wrong dosage and usage of expired vaccines (Disasa et al. 2020). Besides the conventional approach, recombinant vaccine studies against *C. chauvoei* are investigated to provide new generation vaccines (Corpus et al. 2008; Frey et al. 2012). A study with recombinant CctA conferred protection in animal models against *C. chauvoei* challenge, suggesting the possible inclusion of the toxin into a toxoid vaccine formulation. Although this monovalent vaccine seems to provide high immunity, it would be better to include several antigenic proteins into toxoid or multi-epitope vaccines to confer long-lasting and efficient immunity.

Traditional antigen discovery for vaccine production is a time-consuming process and only a small portion of microorganism's proteins is handled due to the limitation of culture-based techniques. In the era of whole genomic and proteomic analysis, it is now possible to cover a huge number of proteins for the discovery of novel antigens in a fast manner via high throughput in silico methods. While proteome-mining with reverse vaccinology pipeline reveals novel vaccine candidates, the immunoinformatics approach provides the identification of T and B cell epitopes of these candidates in silico that can be further analyzed in in vitro studies. The high throughput methods also allow docking these proteins or epitopes with immune cell receptors, especially with Toll-like receptors (TLRs) to evaluate their binding affinity (Oli et al. 2020).

In the present study, by means of reverse vaccinology and immunoinformatics approaches, it is aimed to determine novel vaccine candidates from *C. chauvoei* strains and to identify their T and B cell epitopes to construct a multi-epitope vaccine against the pathogen. The vaccine construct can be used in a combined subunit vaccine against clostridial diseases of animals.

Methodology

Proteome Retrieval and Selection of Non-redundant Proteins

The Microbial Genome Database of National Center for Biotechnology Information (NCBI) (https://www.ncbi.nlm.nih.gov/genome/microbes/) contains four different strains of *C. chauvoei* with their genomic and proteomic data. Thus, the protein sequences of all strains were retrieved from the database in FASTA format. They were subjected to the cluster

database at high identity with tolerance (CD-HIT) (http://weizhong-lab.ucsd.edu/cd-hit/) analysis for the identification of non-redundant proteins and the sequence identity threshold was set to 80% along with the exclusion of the proteins smaller than 100 amino acids in length (Huang et al. 2010).

Identification of Antigenic and Virulent Vaccine Candidates

The Vaxign2 (http://www.violinet.org/vaxign2) reverse vaccinology approach was employed to determine potential vaccine candidates based on the defined criteria which were as follows: Vaxign-ML > 90, transmembrane helicase ≥ 1, no human or mouse similarity, and adhesin property ≥ 0.51 (He et al. 2010). Vaxign2 also predicts subcellular localization of the proteins through PSORTb v3.0 server. The proteins tagged with 'unknown location' in the Vaxign2 analysis were subjected to CELLO v2.5 (http://cello.life.nctu.edu.tw/) (Yu et al. 2006). Cell wall and extracellular proteins were shortlisted and they were analyzed to reveal antigenic and virulent proteins through VaxiJen v2.0 (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html) and BLASTp of virulence factor database (VFDB) (http://www.mgc.ac.cn/VFs/), respectively (Doytchinova and Flower 2007; Liu et al. 2019). While the proteins with VaxiJen score of > 0.5 were considered as antigenic, VFDB cut-off values were as follows: E-value of 1e − 04 and bit score of > 100.

Prediction of T and B Lymphocyte Epitopes

The CTL epitopes were uncovered using the immune epitope database (IEDB) major histocompatibility complex (MHC) class I tool (http://tools.iedb.org/mhci/) (Vita et al. 2019). In the analysis, cow alleles including BoLA-6*01301 (HD6), BoLA-2*01201 (T2A), BoLA-3*00201 (JSP), BoLA-1*02301 (D18.4), BoLA-3*00101 (AW10), BoLA-6*04101 (T2B), BoLA-T2C and BoLA-T5 were selected to represent three bovine species (*Bos taurus taurus*, *Bos taurus indicus*, and hybrid) (Santos Junior et al. 2020). The epitopes with a score of more than 0.5 and a percentile rank lower than 3 were selected as they were considered as good binders and the epitope length was set to 9 residues (9-mer). Moreover, CTL epitopes were subjected to immunogenicity analysis in the IEDB MHC class I immunogenicity tool (http://tools.iedb.org/immunogenicity/). For the identification of linear B cell epitopes, the amino acid sequences of the shortlisted proteins were subjected to IEDB BepiPred linear epitope prediction 2.0 tool (http://tools.iedb.org/bcell/) (Vita et al. 2019). The threshold was set to 0.5 and the epitopes smaller than 30-mer were included for further analysis.

The shortlisted CTL and B cell epitopes were subjected to antigenicity and allergenicity analysis through the VaxiJen v2.0 and AllergenFP v1.0 (https://ddg-pharmfac.net/Aller

genFP/) servers, respectively (Doytchinova and Flower 2007; Dimitrov et al. 2014). The epitopes with 'allergenic' status and an antigenicity score lower than 0.5 were excluded from the study.

Construction of Multi-epitope Vaccine Candidate

Among the CTL and B cell epitopes screened according to the defined criteria, the top two epitopes with the highest antigenic scores for each protein were selected to construct a polyepitope vaccine candidate. While CTL epitopes were linked by Ala-Ala-Tyr (AAY) linkers, B cell epitopes were linked via Lys-Lys (KK) linkers (Dong et al. 2020). The antigenic nature of the vaccine construct was checked by VaxiJen and AllergenFP was used to evaluate the allergenicity of the construct. While ProtSol (https://protein-sol.manchester. ac.uk/) was employed for solubility assessment, the physicochemical properties like molecular weight, aliphatic index, and instability index were predicted using Expasy ProtParam (https://web.expasy.org/protparam/) (Gasteiger et al. 2005; Hebditch et al. 2017).

Prediction of Secondary and Tertiary Structure of the Construct

The secondary structure of the vaccine construct was evaluated using the PSIPRED server (http://bioinf.cs.ucl.ac.uk/psipred/) (Jones 1999; Buchan and Jones 2019). The 3D structure of the construct was predicted in trRosetta (https://yanglab.nankai.edu.cn/trRosetta/) and refinement of the tertiary structure was performed using the GalaxyRefine server (http://galaxy.seoklab.org/) (Ko et al. 2012; Yang et al. 2020). The refined tertiary structure of the construct was validated via the ERRAT score and z-score prediction provided by ProSA (https://prosa.services.came.sbg.ac.at/prosa.php) (Colovos and Yeates 1993; Wiederstein and Sippl 2007). Further validation of the construct quality was assessed by Ramachandran plot using PROCHEK (https://saves.mbi.ucla.edu/) (Laskowski et al. 1993).

Molecular Docking of the Construct with Bovine TLR 4

In the study, the interaction between the vaccine construct and TLR 4 was studied due to the localization of the receptor on the cell surface. The bovine TLR 4 crystal structure was not present and thus, its amino acid sequence was downloaded from the UniProt database (ID: Q9GL65) and used to model TLR 4 through homology modeling in the SWISS-MODEL server (https://swissmodel.expasy.org/) (Waterhouse et al. 2018). After refinement and validation of the structure, the docking of the vaccine construct with bovine TLR 4 was performed via ClusPro (https://cluspro.org) (Kozakov et al. 2017). The top outputted complex was subjected to the PDBsum (http://www.ebi.ac.uk/thornton-srv/databases/pdbsum/Generate.html) analysis to map the residues of the construct and TLR 4 in the interaction site and the binding free energy of the docked complex was assessed via HawkDock (http://cadd.zju.edu.cn/hawkdock/) using Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) method (Laskowski et al. 2018; Weng et al. 2019).

Normal Mode Analysis of the Docked Complex

iMODS server (http://imods.chaconlab.org/) was used to investigate structural dynamics of the vaccine-receptor complex (Lopez-Blanco et al. 2014). The server employs normal mode analysis (NMA) to compute protein internal coordinates for the stability evaluation and the results are represented as deformability plot, covariance matrix, elastic network model, B-factor value, and eigenvalue.

Codon Optimization and In Silico Cloning

Reverse translation and codon optimization were performed via the Java Codon Adaptation Tool (JCat) (http://www.jcat.de/) (Grote et al. 2005). *Escherichia coli* K12 strain was selected for the prokaryotic expression system as it was commonly used and additional options including the removal of rho-independent transcription terminators, cleavage sites of restriction enzymes, and prokaryotic ribosomal binding sites were employed. In silico cloning studies were employed through SnapGene software (https://www.snapgene.com).

Results

Subtractive Proteomic Approach

For the protein sequence retrieval, the NCBI Microbial Genome Database was used as it contains proteomic as well as genomic data and also more *C. chauvoei* strains than other databases including UniProt. The complete proteome of *C. chauvoei* SBP 07/09, DSM 7528, 12S0467, and JF4335 strains containing 10,189 protein sequences were retrieved in FASTA format. The bacterial genome usually has redundant sequences as a result of duplication during evolution, transduction, transformation, and conjugation and the removal of these sequences is appreciated as they have no influence on pathogen survival (ul Qamar et al. 2021). CD-HIT is a widely used tool to reduce sequence redundancy by clustering method (Huang et al. 2010). In the study, the proteins with 80% sequence identity were excluded as they may reflect structural and functional similarity. Moreover, the sequences smaller than 100 amino acids were eliminated

due to the low possibility of being an essential protein in pathogen survival (Rahman et al. 2020). As a result, the CD-HIT analysis revealed 2286 non-redundant proteins which were then further analyzed (Supplementary Table S1).

Determination of Antigenic and Virulent Vaccine Candidates

Vaxign2 was used to determine vaccine candidates among the non-redundant proteins. It is the first web-based vaccine design tool that reverse vaccinology is employed, resulting in the prediction of many features including transmembrane helices, adhesin probability, subcellular localization, and similarity to human and mice proteins through TMHMM, SPAAN, PSORTb v3.0, and NCBI BLAST servers, respectively. Besides these predicted features, Vaxign2 assigns a score to each protein reflecting their protegenicity and induction potential for protective immunity (He et al. 2010). After the Vaxign2 analysis and determination of the location of 'unknown' proteins via CELLO v2.5, 6 cell wall and 50 extracellular proteins were shortlisted as vaccine candidates due to their possible interaction with immune cells (Supplementary Table S2). These candidates were then subjected to antigenicity analysis in VaxiJen v2.0 and their virulent potential was revealed through VFDB. VaxiJen performs an Auto Cross-Covariance (ACC) transformation via an alignment-independent algorithm and a protein with a threshold of more than 0.5 is considered as a potential immunogen (Doytchinova and Flower 2007). VFDB contains 714 experimentally verified virulence factors and provides a BLAST search against the database (Doytchinova and Flower 2007; Liu et al. 2019). According to the defined criteria, 2 cell wall and 5 extracellular proteins were identified as antigenic and virulent vaccine candidates (Table 1).

Epitope Selection for T and B Lymphocytes

An effective vaccine is expected to mimic natural immunity by inducing a long-lasting adaptive immunity where T and B cells serve as critical components of the immune response. While CTL epitopes can trigger long-lasting cellular immunity, B cell epitopes stimulate humoral immune response through the production of immunoglobulins. Although helper T lymphocytes (HTL) are critical elements for the activation of CTL and B cells, a complete sequence of the bovine MHC II locus related to HTL was not assembled in the bovine genome (Gaafar et al. 2019). Therefore, only BoLA MHC I analysis was performed in the study for the determination of the CTL epitopes. Among the selected MHC I alleles, BoLA-HD6, BoLA-T2A, BoLA-JSP, BoLA-D18.4, and BoLA-T2B represent *Bos taurus taurus* whereas BoLA–T5 and BoLA-AW10 belong to *Bos taurus indicus*. Additionally, BoLA-T2C belongs to a *Bos taurus taurus* and *Bos taurus indicus* hybrid (Santos Junior et al. 2020). For the determination of CTL and B cell epitopes, the IEDB server was employed with the defined criteria. The server provides a wide range of analysis methods for epitope prediction and it contains over one million B cell, T cell, and MHC binding elution assays and more than 260,000 epitopes (Fleri et al. 2017).

All shortlisted CTL and B cell epitopes were subjected to antigenicity analysis in the VaxiJen server to obtain the most favorable epitopes to provoke the immune system. They were also analyzed through AllergenFP for the exclusion of the allergenic epitopes to avoid allergenic reactions in the body. The detailed information about the non-allergenic and antigenic CTL and B cell epitopes was presented in Supplementary Tables S3 and S4 whereas the non-allergenic epitopes with the highest antigenic scores were listed in Table 2 as they were considered to design a multi-epitope vaccine candidate against *C. chauvoei*. In the case of flagellin and flagellar hook-basal body protein, no CTL and B cell epitopes met the pre-defined criteria for epitope selection, respectively.

In Silico Design of Multi-epitope-Based Vaccine

For the construction of the multi-epitope vaccine, the top two CTL and B cell epitopes were conjugated via AAY and

Table 1 Antigenic and virulent vaccine candidates predicted through reverse vaccinology pipeline

Protein accession	Protein name	VaxignML score	Localization (probability)	Adhesin probability	Transmembrane helices	Antigenicity	Virulent
QBJ74913.1	Flagellin	95.7	Extracellular	0.801	0	0.5876	✓
QBJ76507.1	Hypothetical protein C6H62_13090	97.7	Extracellular	0.827	0	0.5266	✓
ATD57902.1	Alpha hemolysin (CctA)	96.8	Extracellular	0.836	0	0.5776	✓
QBJ74926.1	Flagellar hook-basal body protein	99.7	Extracellular	0.774	0	0.6971	✓
CDG00977.1	Putative cell wall surface anchor family protein	97.5	Extracellular	0.663	0	0.6630	✓
QBJ74801.1	Hypothetical protein C6H62_03650	99.9	Cell wall	0.744	1	0.5720	✓
QBJ75069.1	Exo-alpha-sialidase	99.9	Cell wall	0.762	1	0.6266	✓

Table 2 The prioritized CTL and B cell epitopes of the selected vaccine candidates

Protein accession	Protein name	CTL epitope	B cell epitope
QBJ74913.1	Flagellin	–	RLEHTINNLNTSSENLTAAESRVRDVD ATAGKSMEKLSSGLRINRAGD
QBJ76507.1	Hypothetical protein C6H62_13090	SGIQINYPK TVGDHELTF	TKVDSTNPNYYI YAEDGPGSIK
ATD57902.1	Alpha hemolysin (CctA)	TTTIFSLTL KGTEESTVK	SIEYGKEGPKAGGGINGSYTAQ WYGENNRNSRIDS
QBJ74926.1	Flagellar hook-basal body protein	DIAGVTITL ELNGYSFEI	–
CDG00977.1	Putative cell wall surface anchor family protein	NLAGATFEI GKLRFQLPY	TKTDKQTGKPI ANVKNKKIAEQSNTG
QBJ74801.1	Hypothetical protein C6H62_03650	ENAEPGYTM RVGLGFRIK	LNPTEGETA IEGGNRVVLTTENRGDEEGAYK
QBJ75069.1	Exo-alpha-sialidase	STSGGVGQL STAGDTNNK	TSSQPGEGADKAIDNNTST DAGNRSNGT

KK linkers, respectively (Table 2; Fig. 1a). The vaccine construct was then evaluated in terms of physicochemical properties using Expasy Protparam. The final construct consisted of 356 amino acids and the molecular weight was predicted as 38.28 kDa with a theoretical isoelectric point (pI) of 9.51 and the instability index (II) score of 19.08. Any value of pI higher than 7 indicates basic nature of proteins and an instability index greater than 40 represents unstable structures (Ali et al. 2017). Therefore, it is suggested that the vaccine construct is alkaline and stable in nature. Moreover, the aliphatic index defined as the relative volume maintained by the aliphatic side chains was 60.76. A higher value of aliphatic index indicates greater thermostability and thus, the construct has thermostable nature. Furthermore, a negative score of grand average of hydrophobicity index (−0.808) suggested hydrophilic nature of the construct that is a desirable feature of a vaccine candidate (Ali et al. 2017).

The antigenic feature of the vaccine construct was assessed by VaxiJen and the analysis confirmed the antigenicity with a score of 1.2137. Moreover, allergenicity was evaluated to avoid allergic reactions in the body and AllergenFP confirmed that the vaccine construct had a non-allergenic nature.

Structural Analysis and Validation of the Construct and TLR 4

Among the TLRs, TLR 4 and TLR 11 have been documented to recognize the proteins and only TLR 4 was selected to test the binding ability of the vaccine construct due to the absence of TLR 11 in bovine (Kar and Srivastava 2018). Based on the human TLR 4 as a template, a homology modeling of bovine TLR 4 was employed through SWISS-MODEL. The server provides global model quality estimation (GMQE) score reflecting the certainty of a model and greater GMQE indicates better quality of models (Shahid et al. 2020). The homology modeling result of bovine TLR 4 has a GMQE score of 0.46, suggesting higher quality of the TLR 4 model. The secondary structure of the vaccine construct was employed by PSIPRED (Fig. 1b). The server provides position-specific iterated-blast for the secondary structure prediction based on amino acid sequence. The tertiary structure of the construct was modeled through trRosetta using residual deep neural networks for the prediction of location distribution and inner-residue space (Fig. 1c) (Zhao et al. 2020). The structural quality of the 3D-modeled vaccine construct through ERRAT score, z-score, and Ramachandran plot was presented. The overall quality factor predicted by ERRAT was 80.24, indicating a high quality of the modeled structure as the value over 50 has been in the generally accepted range (Messaoudi et al. 2013). PROCHECK server was used to generate the Ramachandran plot which resulted in 89.0% residues in the favorable region, 10.1% residues in the allowed region, and 0.9% residues in the outlier region (Fig. 1d). Ideally, residues more than 85% in the favored region confirm the reliability of models. The 3D model of the construct was within the expected range (Yang et al. 2021). The solubility was checked with the ProtSol server and the analysis resulted in a value of 0.65 which indicated a high solubility of the modeled vaccine construct (Fig. 1e). Moreover, ProSA resulted in a z-score of −5.3, implying the reliability of the construct (Fig. 1f). The validation of the bovine TRL 4 model was also employed as in the case of the vaccine construct (data not shown). The overall results have suggested that the modeled construct and bovine TLR 4 were valid and further analysis could be carried out with these models.

Fig. 1 **a** Plain format of the vaccine construct (bold letters indicate linkers), **b** Schematic representation of the secondary structure provided by PSIPRED, **c** 3D structure of the refined construct model, **d** Ramachandran plot of the tertiary structure, **e** solubility analysis of the construct using ProtSol, and **f** z-score (−5.3) of the vaccine construct presented by ProSA
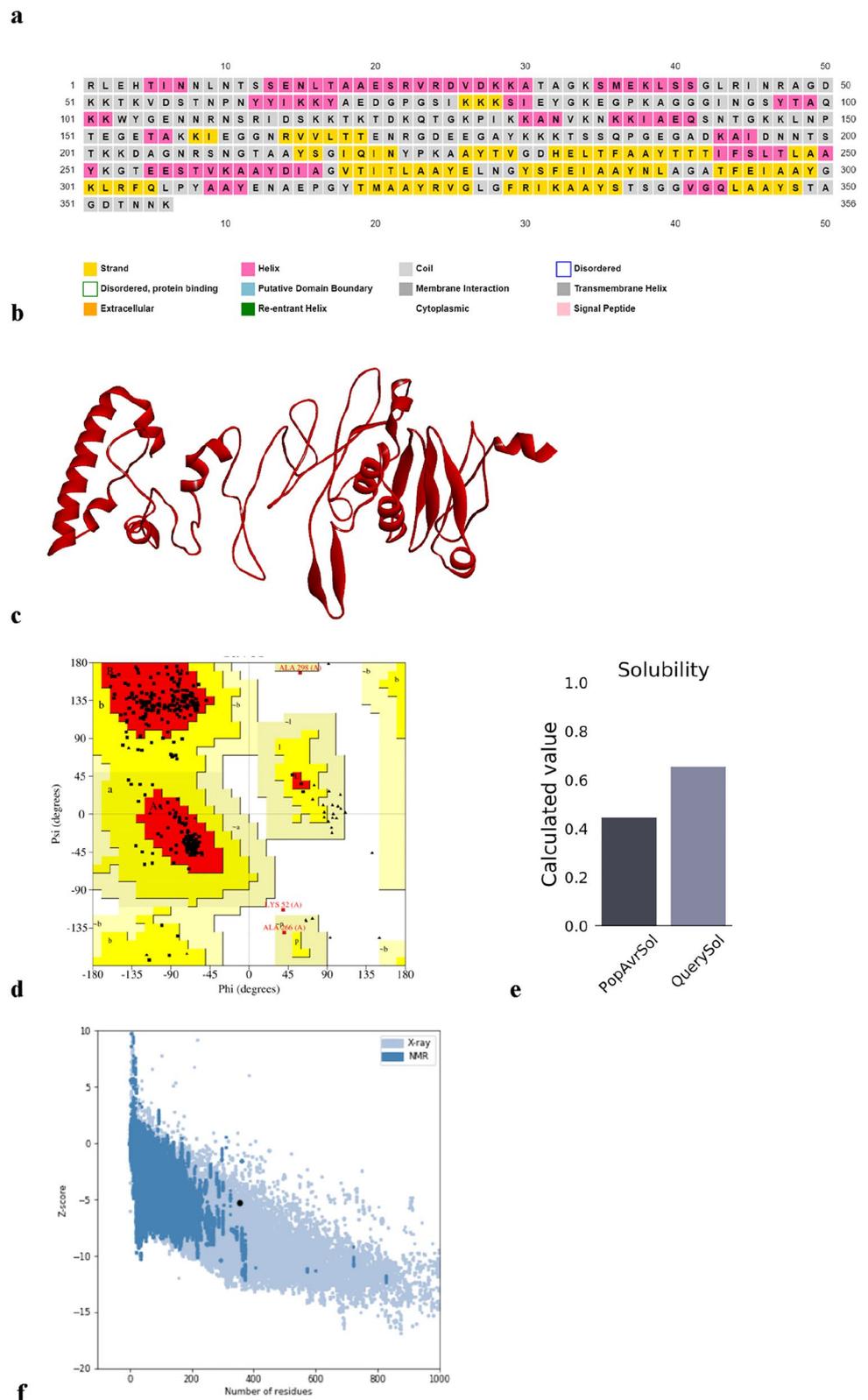
Molecular Docking and Molecular Dynamics Simulations

ClusPro is a fast rigid-body protein-protein docking tool and its fully automated method provides docking according to three algorithms: (1) Fast Fourier Transform correlation approach, (2) clustering the top outputted energy conformations, and (3) refinement of models through medium-range optimization method SDU and short Monte Carlo simulations (Kozakov et al. 2017). After refinement of the 3D structures of the vaccine construct and bovine TLR 4, they were uploaded to ClusPro for protein-protein docking. The top outputted docked complex was selected and the visualization of the complex was carried out with Maestro 12.4 of Schrodinger (Schrödinger Release 2020-3: Maestro, Schrödinger, LLC, New York, NY, 2020) (Fig. 2a). The binding free energy of the docked complex was evaluated through MM/GBSA approach over the HawDock server, providing an efficient scoring of protein-protein interactions. The analysis of the complex revealed − 124.51 kcal/mol, indicating that the docking was energetically achievable. Moreover, the interaction was deeply analyzed through PDBsum and a total of 33 residues of the construct interact with 43 residues of bovine TLR 4. The interaction consists of 21 hydrogen bonds along with 8 salt bridges and 200 nonbonded contacts. All the distance of the H-bonds between bovine TLR 4 and the construct lie within the range of 2–3.5 Å, suggesting strong interactions (Minch 1999). The detailed interacting residues are illustrated in Fig. 2b and

the interaction may indicate that the multi-epitope vaccine candidate can favorably bind to bovine TLR 4.

The flexibility and stability of the vaccine-TLR 4 complex were assessed using NMA in the iMODS simulation suite. First, the deformability of the complex was compared with the monomeric bovine TLR 4 protein and a significant reduction in the distortions was observed in the vaccine-TLR 4 complex, suggesting the stabilization of the complex (Fig. 3a, d). In Fig. 3b, e and B-factor analysis, exposing the minor atomic distortions, was represented for the monomeric bovine TLR 4 and the docked complex, respectively. In addition, the eigenvalue of the monomeric bovine TLR 4 was 1.196886e − 05 whereas it was 1.024499e − 05 for the complex which was slightly higher than the monomeric protein, indicating that much more energy is needed to deform the vaccine-TLR 4 complex (Fig. 3c, f). Moreover, the variance analysis, inversely proportional to eigenvalue, was also employed and it also signified the stiffness of the complex (Fig. 3g). The coupling between the residues of the vaccine-TLR 4 complex was assessed by covariance analysis where correlated, uncorrelated, and anti-correlated motions were represented in red, white, and blue, respectively (Fig. 3h). Furthermore, the elastic network analysis was performed where each spring represented the corresponding pair of atoms. The analysis revealed the dark gray springs demonstrating the compactness and rigidity of the complex (Fig. 3i). All these results suggest stable binding interactions with compact conformation and minor fluctuations in the vaccine–TLR4 complex.

Fig. 2 Molecular docking of the vaccine construct with bovine TLR 4. **a** The schematic representation of the docked complex visualized by Maestro 12.4 of Schrodinger. The vaccine construct and bovine TLR 4 are shown in green and red, respectively. **b** The detailed interaction scheme between the vaccine construct (chain A) and the receptor (chain B) obtained from the PDBsum analysis (Color figure online)
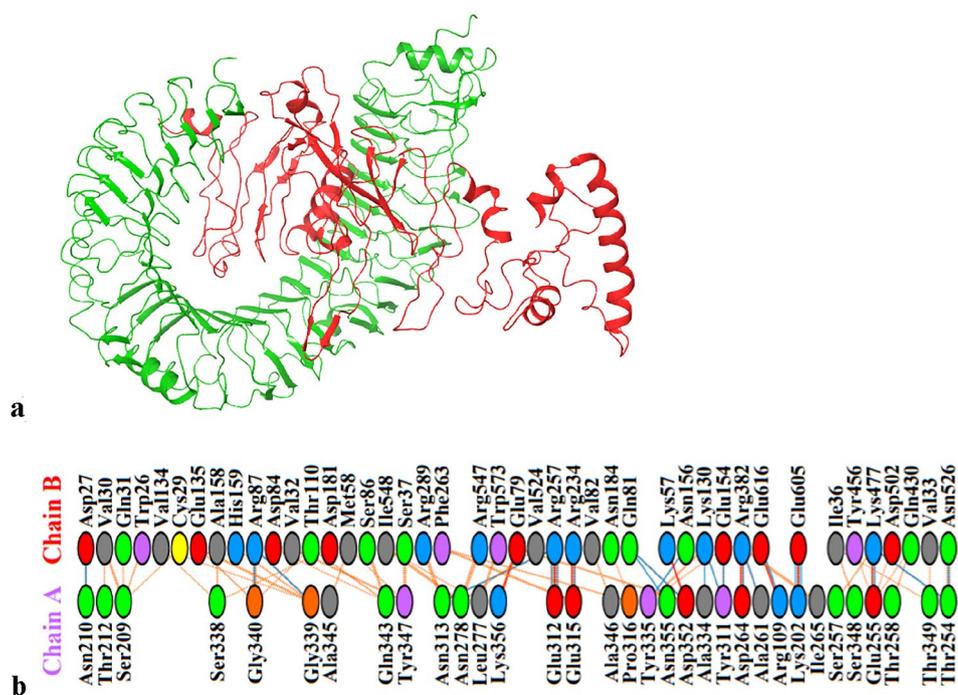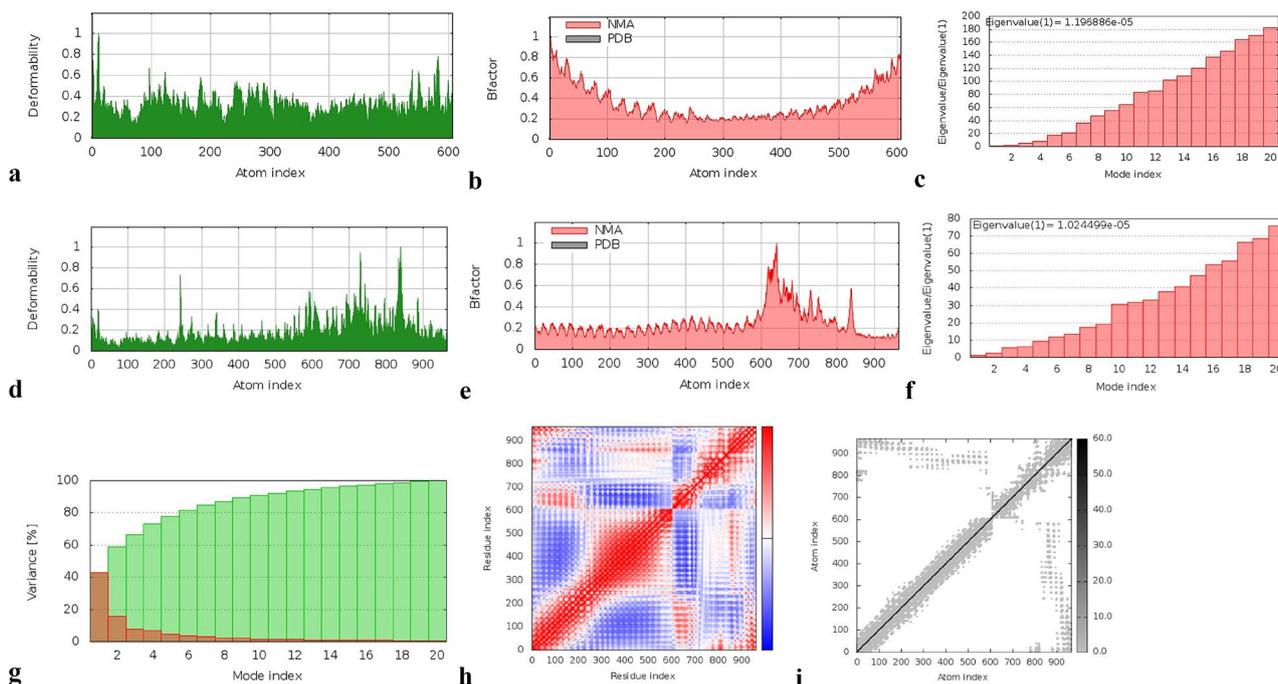
Fig. 3 Molecular dynamics simulation analysis. **a–c** The deformability, B-factor, and eigenvalue of TLR 4, **d–f** the deformability, B-factor, and eigenvalue of the vaccine-TLR 4 complex, **g** variance analysis, **h** covariance map, and **i** elastic network of the vaccine construct-TLR 4 complex

In Silico Cloning of the Vaccine Construct

For the expression of the multi-epitope vaccine construct in *E. coli* K12, codon optimization was employed through the JCat tool. The improved sequence had a value of codon adaptive index (CAI) of 0.97 and a GC content of 47.0%. In general, CAI greater than 0.8 and a GC content between 30 and 70% are considered ideal for efficient protein expression (Kar et al. 2020). Following, NdeI and XhoI restriction sites were added to the sequence and it was cloned into the pET-28a(+) vector (Fig. 4).

Discussion

C. chauvoei is a highly pathogenic spore-forming bacterium that causes a severe disease called blackleg in animals. Rapid death of the animals without symptoms is commonly observed and thus, vaccination is the most important prophylactic measure to prevent the disease. Commercially available vaccines contain chemically inactivated bacteria that are supposed to include the main toxins such as CctA, sialidase, and flagellin. Besides, subunit vaccine development against blackleg is also studied to avoid potential adverse reactions of inactivated vaccines. A recent study investigating 20 strains of *C. chauvoei* isolated from different continents over 64 years revealed that the genome of the strains was greatly homogenous and many metabolic and structural genes were

conserved (Rychener et al. 2017). Thus, the products of the conserved genes can be used to develop subunit and multi-epitope vaccines.

In the present study, the complete proteome of *C. chauvoei* strains found in NCBI was retrieved and subjected to reverse vaccinology approach, resulting in seven prioritized vaccine candidates, namely alpha hemolysin (CctA), flagellin, exo-alpha-sialidase, flagellar hook-basal body protein, putative cell wall surface anchor family protein, and two hypothetical proteins. Among these candidates, alpha hemolysin (CctA), flagellin, and sialidase have been already considered as potential targets for subunit vaccines against the pathogen (Disasa et al. 2020). Moreover, the immunogenicity of flagellar hook-basal body proteins including FlgE and FlgG were also revealed in both in silico and in vitro studies (Wilhelm et al. 2006; Wan et al. 2021). Besides these previously known candidates, three novel antigenic proteins, putative cell wall surface anchor family protein and two hypothetical proteins, were discovered in the current study.

Multi-epitope vaccines have a potential to induce both cellular and humoral immune responses through T and B cells, respectively, and they are suggested as a more valuable alternative to monovalent vaccines. Therefore, a multi-epitope vaccine containing prioritized CTL and B cell epitopes was constructed against *C. chauvoei* after epitope-mapping of the selected vaccine candidates. During the construction, AAY sequence was used to link CTL epitopes as the studies showed that flanking Ala and Tyr

Fig. 4 In silico cloning of the vaccine construct into the pET-28a(+) expression vector via SnapGene free-trial software. While the black part represents the vector backbone, the vaccine construct is represented in red (Color figure online)
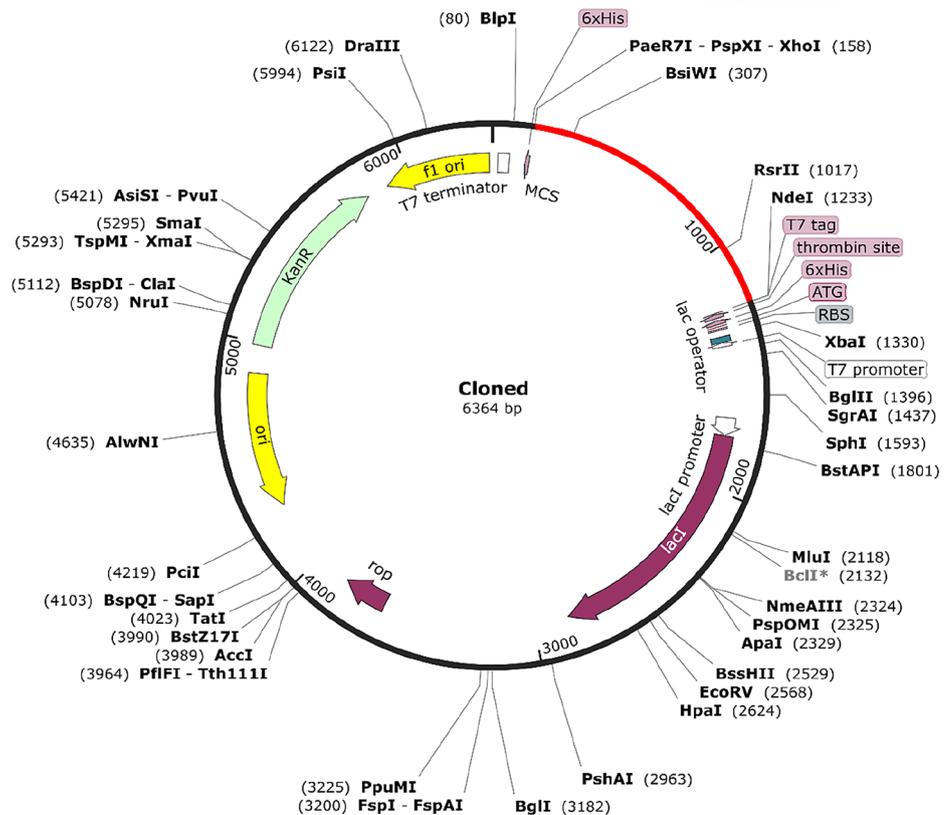
residues did not inhibit proteasomal digestion, resulting in enhanced CTL epitope presentation (Bergman et al. 1996; Wang et al. 2004). Moreover, AAY linker was used to provide better separation of the short CTL epitopes (Pandey et al. 2018). B cell epitopes were joined via KK linker that is the target for lyzosomal protease, cathepsin B, which has an important role in antigen processing. KK linker also prevents antibody induction for junctional epitopes formed when individual epitopes are linked linearly (Ayyagari et al. 2020). After the construction, in silico analysis revealed that the construct was stable, hydrophilic, and the quality of the model was in the accepted range. Moreover, docking studies demonstrated the interaction between the vaccine construct and bovine TLR 4, suggesting that the multi-epitope vaccine construct may be used to stimulate an effective immune response.

In conclusion, reverse vaccinology and immunoinformatics approach paved the way for designing a chimeric subunit vaccine that could be an ideal vaccine candidate against *C. chauvoei*. The therapeutic and prophylactic effects of the chimeric vaccine need to be further evaluated in pre-clinical studies.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10989-021-10279-9.

Declarations

Conflict of interest The author declares that there is no conflict of interest.

References

Abreu CC, Uzal FA (2016) Blackleg. In: Uzal FA, Songer JG, Prescott JF, Popoff MR (eds) Clostridial diseases of animals. Wiley, New York, pp 229–242

Ali M, Pandey RK, Khatoon N, Narula A, Mishra A, Prajapati VK (2017) Exploring dengue genome to construct a multi-epitope based subunit vaccine by utilizing immunoinformatics approach to battle against dengue infection. Sci Rep 7:9232. https://doi.org/10.1038/s41598-017-09199-w

Ayyagari VS, Venkateswarulu TC, Abraham Peele K, Srirama K (2020) Design of a multi-epitope-based vaccine targeting M-protein of SARS-CoV2: an immunoinformatics approach. J Biomol Struct Dyn. https://doi.org/10.1080/07391102.2020.1850357

Bergmann CC, Yao Q, Ho CK, Buckwold SL (1996) Flanking residues alter antigenicity and immunogenicity of multi-unit CTL epitopes. J Immunol 157:3242–3249

Buchan DWA, Jones DT (2019) The PSIPRED protein analysis workbench: 20 years on. Nucleic Acids Res 47(W1):W402–W407. https://doi.org/10.1093/nar/gkz297

Colovos C, Yeates TO (1993) Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci 2:1511–1519. https://doi.org/10.1002/pro.5560020916

Corpus MLO, Andrade LH, Mendez JL, Gutierrez VT (2008) Prevention of blackleg by an immunogen of *Clostridium chauvoei*. Ann NY Acad Sci 1149(1):303–305. https://doi.org/10.1196/annals.1428.013

Dimitrov I, Naneva L, Doytchinova I, Bangov I (2014) AllergenFP: allergenicity prediction by descriptor fingerprints. Bioinformatics 30(6):846–851. https://doi.org/10.1093/bioinformatics/btt619

Disasa DD, Balcha MT, Negewo SM, Mamo ME, W/Sanbat TB, Disasa WK (2020) Review on the blackleg disease in domestic animal. GSJ 8(8):1133–1148

Dong R, Chu Z, Yu F, Zha Y (2020) Contriving multi-epitope subunit of vaccine for COVID-19: immunoinformatics approach. Front Immunol 11:1784. https://doi.org/10.3389/fimmu.2020.01784

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

Fleri W, Paul S, Dhanda SK, Mahajan S, Xu X, Peters B, Sette A (2017) The immune epitope database and analysis resource in epitope discovery and synthetic vaccine design. Front Immunol 8(278):278. https://doi.org/10.3389/fimmu.2017.00278

Frey J, Johansson A, Bürki S, Vilei EM, Redhead K (2012) Cytotoxin CctA, a major virulence factor of *Clostridium chauvoei* conferring protective immunity against myonecrosis. Vaccine 30:5500–5505. https://doi.org/10.1016/j.vaccine.2012.06.050

Gaafar BBM, Ali SA, Abd-elrahman KA, Almofti YA (2019) Immunoinformatics approach for multiepitope vaccine prediction from H, M, F, and N proteins of Peste des Petits ruminants virus. J Immunol Res. https://doi.org/10.1155/2019/6124030

Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A (2005) Protein identification and analysis tools on the ExPASy server. In: Walker JM (ed) The proteomics protocols handbook. Humana Press, Totowa, pp 571–607

Grote A, Hiller K, Scheer M, Münch R, Nörtemann B, Hempel DC, Jahn D (2005) JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. Nucleic Acids Res 33:W526–W531. https://doi.org/10.1093/nar/gki376

He Y, Xiang Z, Mobley HLT (2010) Vaxign: the first web-based vaccine design program for reverse vaccinology and an application for vaccine development. J Biomed Biotechnol. https://doi.org/10.1155/2010/297505

Hebditch M, Carballo-Amador MA, Charonis S, Curtis R, Warwicker J (2017) Protein-sol: a web tool for predicting protein solubility from sequence. Bioinformatics 33(19):3098–3100. https://doi.org/10.1093/bioinformatics/btx345

Huang Y, Niu B, Gao Y, Fu L, Li W (2010) CD-HIT Suite: a web server for clustering and comparing biological sequences. Bioinformatics 26:680–682. https://doi.org/10.1093/bioinformatics/btq003

Jones DT (1999) Protein secondary structure prediction based on position-specific scoring matrices. J Mol Biol 292(2):195–202. https://doi.org/10.1006/jmbi.1999.3091

Kar PP, Srivastava A (2018) Immuno-informatics analysis to identify novel vaccine candidates and design of a multi-epitope based vaccine candidate against *Theileria* parasites. Front Immunol 9:2213. https://doi.org/10.3389/fimmu.2018.02213

Kar T, Narsaria U, Basak S, Deb D, Castiglione F, Mueller DM, Srivastava AP (2020) A candidate multi-epitope vaccine against SARS-CoV-2. Sci Rep 10:10895. https://doi.org/10.1038/s41598-020-67749-1

Ko J, Park H, Heo L, Seok C (2012) GalaxyWEB server for protein structure prediction and refinement. Nucleic Acids Res 40(W1):W294–W297. https://doi.org/10.1093/nar/gks493

Kozakov D, Hall DR, Xia B, Porter KA, Padhorny D, Yueh C, Beglov D, Vajda S (2017) The ClusPro web server for protein-protein docking. Nat Protoc 12(2):255–278. https://doi.org/10.1038/nprot.2016.169

Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK—a program to check the stereochemical quality of protein structures. J Appl Crystallogr 26:283–291. https://doi.org/10.1107/S0021889892009944

Laskowski RA, Jabłońska J, Pravda L, Vařeková RS, Thornton JM (2018) PDBsum: structural summaries of PDB entries. Protein Sci 27:129–134. https://doi.org/10.1002/pro.3289

Liu B, Zheng DD, Jin Q, Chen LH, Yang J (2019) VFDB 2019: a comparative pathogenomic platform with an interactive web interface. Nucleic Acids Res 47(D1):D687–D692. https://doi.org/10.1093/nar/gky1080

López-Blanco JR, Aliaga J, Quintana-Ortí ES, Chacón P (2014) iMODS: internal coordinates normal mode analysis server. Nucleic Acids Res 42:W271–W276. https://doi.org/10.1093/nar/gku339

Messaoudi A, Belguith H, Ben Hamida J (2013) Homology modeling and virtual screening approaches to identify potent inhibitors of VEB-1 beta-lactamase. Theor Biol Med Model 10:22. https://doi.org/10.1186/1742-4682-10-22

Minch JM (Jeffrey GA (1999) An introduction to hydrogen bonding. J Chem Educ 76(6):759. https://doi.org/10.1021/ed076p759.1

Nagano N, Isomine S, Kato H, Sasaki Y, Takahashi M, Sakaida K, Nagano Y, Arakawa Y (2008) Human fulminant gas gangrene caused by *Clostridium chauvoei*. J Clin Microbiol 46:1545–1547. https://doi.org/10.1128/JCM.01895-07

Oli AN, Obialor WO, Ifeanyichukwu MO, Odimegwu DC, Okoyeh JN, Emechebe GO, Adejumo SA, Ibeanu GC (2020) Immunoinformatics and vaccine development: an overview. Immunotargets Ther 9:13–30. https://doi.org/10.2147/ITT.S241064

Pandey RK, Ojha R, Mishra A, Prajapati VK (2018) Designing B- and T-cell multi-epitope based subunit vaccine using immunoinformatics approach to control Zika virus infection. J Cell Biochem 119(9):7631–7642. https://doi.org/10.1002/jcb.27110

Rahman N, Ajmal A, Ali F, Rastrelli L (2020) Core proteome mediated therapeutic target mining and multi-epitope vaccine design for *Helicobacter pylori*. Genomics 112:3473–3483. https://doi.org/10.1016/j.ygeno.2020.06.026

Rychener L, In-Albon S, Djordjevic SP, Chowdhury PR, Nicholson P, Ziech RE, de Vargas AC, Frey J, Falquet L (2017) *Clostridium chauvoei*, an evolutionary dead-end pathogen. Front Microbiol 8:1054. https://doi.org/10.3389/fmicb.2017.01054

Santos Junior MNS, Santos RS, Neves WS, Fernandes JM, de Brito Guimaraes BC, Barbosa MS, Silva LSC, Gomes CM, Rezende IS, Oliveria CNT, Neres NSM, Campos GB, Bastos BL, Timenetsky J, Marques LM (2020) Immunoinformatics and analysis of antigen distribution of *Ureaplasma diversum* strains isolated from different Brazilian states. BMC Vet Res 16:379. https://doi.org/10.1186/s12917-020-02602-1

Shahid F, Ashfaq UA, Javaid A, Khalid H (2020) Immunoinformatics guided rational design of a next generation multi epitope based peptide (MEBP) vaccine by exploring Zika virus proteome. Infect Genet Evol 80:104199. https://doi.org/10.1016/j.meegid.2020.104199

ul Qamar MT, Ahmad S, Fatima I, Ahmad F, Shahid F, Naz A, Abbasi SW, Khan A, Mirza MU, Ashfaq UA, Chen LL (2021) Designing multi-epitope vaccine against *Staphylococcus aureus* by employing subtractive proteomics, reverse vaccinology and immunoinformatics approaches. Comput Biol Med 132:104389. https://doi.org/10.1016/j.compbiomed.2021.104389

Vita R, Mahajan S, Overton JA, Dhanda SK, Martini S, Cantrell JR, Wheeler DK, Sette A, Peters B (2019) The Immune Epitope Database (IEDB): 2018 update. Nucleic Acids Res 47(D1):D339–D343. https://doi.org/10.1093/nar/gky1006

Wan C, Gao C, Xie Q, Wang Y, Cheng X, Fang Y, Liu Z, Zhang W, Zou Q, Lu G, Gu J (2021) Flagella hook protein FlgE is a novel vaccine candidate of *Pseudomonas aeruginosa* identified by a genomic approach. Vaccine 39(17):2386–2395. https://doi.org/10.1016/j.vaccine.2021.03.051

Wang QM, Sun SH, Hu ZL, Zhou FJ, Yin M, Xiao CJ, Zhang JC (2004) Epitope DNA vaccines against tuberculosis: spacers and ubiquitin modulates cellular immune responses elicited by epitope DNA vaccine. Scand J Immunol 60(3):219–225. https://doi.org/10.1111/j.0300-9475.2004.01442.x

Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L, Lepore R, Schwede T (2018) SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res 46(W1):W296–W303. https://doi.org/10.1093/nar/gky427

Weatherhead JE, Tweardy DJ (2012) Lethal human neutropenic enterocolitis caused by *Clostridium chauvoei* in the United States: tip of the iceberg? J Infect 64:225–227. https://doi.org/10.1016/j.jinf.2011.09.004

Weng GQ, Wang EC, Wang Z, Liu H, Li D, Zhu F, Hou TJ (2019) HawkDock: a web server to predict and analyze the structures of protein-protein complexes based on computational docking and MM/GBSA. Nucleic Acids Res 47(W1):W322–W330. https://doi.org/10.1093/nar/gkz397

Wiederstein M, Sippl MJ (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res 35:W407–W410. https://doi.org/10.1093/nar/gkm290

Wilhelm V, Miquel A, Burzio LO, Rosemblatt M, Engel E, Valenzuela S, Parada G, Valenzuela PDT (2006) A vaccine against the salmonid pathogen *Piscirickettsia salmonis* based on recombinant proteins. Vaccine 24(23):5083–5091. https://doi.org/10.1016/j.vaccine.2006.03.027

Yang J, Anishchenko I, Park H, Peng Z, Ovchinnikov S, Baker D (2020) Improved protein structure prediction using predicted interresidue orientations. PNAS 117:1496–1503. https://doi.org/10.1073/pnas.1914677117

Yang Z, Bogdan P, Nazarian S (2021) An *in silico* deep learning approach to multi-epitope vaccine design: a SARS-CoV-2 case study. Sci Rep 11:3238. https://doi.org/10.1038/s41598-021-81749-9

Yu CS, Chen YC, Lu CH, Hwang JK (2006) Prediction of protein subcellular localization. Proteins 64:643–651. https://doi.org/10.1002/prot.21018

Zaragoza NE, Orellana CA, Moonen GA, Moutafis G, Marcellin E (2019) Vaccine production to protect animals against pathogenic Clostridia. Toxins 11(9):525. https://doi.org/10.3390/toxins11090525

Zhao J, Yuan Q, Wang H, Liu W, Liao X, Su Y, Wang X, Yuan J, Li T, Li J, Qian S, Hong C, Wang F, Liu Y, Wang Z, He Q, Li Z, He B, Zhang T, Fu Y, Ge S, Liu L, Zhang J, Xia N, Zhang Z (2020) Antibody responses to SARS-CoV-2 in patients with novel Coronavirus disease 2019. Clin Infect Dis 71(16):2027–2034. https://doi.org/10.1093/cid/ciaa344