Applied Microbiology and Biotechnology (2021) 105:8277–8286
https://doi.org/10.1007/s00253-021-11619-x

In silico analyses and design of a chimeric protein containing epitopes of SpaC, PknG, NanH, and SodC proteins for the control of caseous lymphadenitis

Mara Thais de Oliveira Silva[1] · Rodrigo Barros de Pinho[1] · Francisco Silvestre Brilhante Bezerra[2] · Nicole Ramos Scholl[1] · Luiza Domingues Moron[1] · Mirna Samara Dié Alves[1] · Rafael dos Santos Woloski[3] · Frederico Schmitt Kremer[3] · Sibele Borsuk[1]

Received: 25 June 2021 / Revised: 23 September 2021 / Accepted: 26 September 2021 / Published online: 8 October 2021

Abstract

Caseous lymphadenitis (CLA) is an infectious disease that affects goats and sheep causing drastic impacts on milk and meat production and is caused by *Corynebacterium pseudotuberculosis*. The disease can be prevented through vaccination but currently, vaccines demonstrate limited efficacy consequently leading to a need for the development of new ones. Here, we described the *in silico* development of a new chimeric protein constructed with epitopes identified from the sequences of the genes *nanH*, *pknG*, *spaC*, and *sodC*, previously described as potential vaccinal targets against *C. pseudotuberculosis*. The chimera was expressed, purified, and its immunogenicity was evaluated using sera of immunized mice. Results indicate the chimeric protein was able to stimulate antibody production. Additionally, analysis using serum from naturally infected goats showed that the protein is recognized by sera from these animals, indicating the possibility for using this chimera in new diagnostic methods.

Key points

- *The chimera was expressed with 52 kDa and a yield of 7 mg/L after purification.*
- *The chimera was recognized by the sera of animals immunized with this formulation.*
- *Chimera reacted with the serum of goats naturally infected with C. pseudotuberculosis.*

Keywords Chimera · *C. pseudotuberculosis* · Immunoinformatics · Epitopes · Immunodominant

Introduction

Caseous lymphadenitis (CLA) is recognized as a significant cause of economic loss in small ruminants farming worldwide (Guimarães et al. 2011). In Australia, where financial losses related to this disease have been extensively studied, CLA causes an estimated annual loss of $12–$15 million to the meat industry and approximately a $17 million annual loss in wool production (Osman et al. 2018). In Brazil, data from a study aiming at determining the negative impact of CLA in a slaughterhouse revealed economic losses of 6.09% on goat production (Barnabé et al. 2020). Currently, commercially available vaccines for *Corynebacterium pseudotuberculosis* demonstrate limited efficacy and do not provide effective protection against CLA. Consequently, more effective vaccine formulations are a priority (Dorella et al. 2009).

Most of the recent attempts for developing CLA vaccines are based on the recombinant subunit approach and generally demonstrate higher levels of safety since they are based on selected purified antigens (Rezende et al. 2020; Bezerra et al. 2021). The success of this type of vaccine is closely related to the selection of vaccine targets (Droppa-Almeida et al. 2016).

✉ Sibele Borsuk
sibeleborsuk@gmail.com

[1] Laboratório de Biotecnologia Infecto-Parasitária, Centro de Desenvolvimento Tecnológico, Biotecnologia, Universidade Federal de Pelotas (UFPel), Campus Universitário s/n, Prédio 19, Pelotas, RS 96010-900, Brazil

[2] Laboratório de Biotecnologia Aplicada a Doenças Infecto-Parasitárias, Centro de Ciências Biológicas e da Saúde, UFERSA, Mossoró, RN, Brazil

[3] Laboratório de Bioinformática e Proteômica, Centro de Desenvolvimento Tecnológico, Biotecnologia, UFPel, Pelotas, RS 96010-900, Brazil

In 2010, sequencing of *C. pseudotuberculosis* genome by pyrosequencing revealed some potential virulence factors in the human isolate FCR41 (Trost et al. 2010). Based on this study, NanH, PknG, SpaC, and SodC proteins were selected for *in silico* vaccine potential characterization (Santana-Jorge et al. 2016). Considering that proteins involved in host cell adhesion and invasion, nutrient acquisition, and evasion of the host immune system could play important roles in the virulence and pathogenicity of *C. pseudotuberculosis*, these proteins are promising for the development of vaccines against CLA (Santana-Jorge et al. 2016).

In this context, structural vaccinology is a noteworthy field when it comes to vaccine development due to its impact on improving their immunogenicity and stability, and on providing information to promote the design of new and improved antigens (Dormitzer et al. 2012). By using the immunoinformatics approach, structural vaccinology has accelerated research through the identification of dominant immunogens, offering a high level of confidence in predicting target epitopes that applied in vaccines can effectively stimulate protective immune responses against several pathogens (Lata et al. 2018). Using such knowledge, chimera production comes to light as a strategy to aggregate these potential epitopes into a single improved antigen to maximize efficacy and minimize side effects in comparison to the available vaccines (Oliveira et al. 2017). In this study, we identified immunogenic epitopes on the virulence factors NanH, PknG, SpaC, and SodC, combined them *in silico* to produce a chimera which was applied in a chimeric vaccine and assessed *in vitro* for its antigenicity in goats.

Materials and methods

Protein sequences

NanH, PknG, SpaC, and SodC protein sequences (GenBank accession nos. **ADK28179.1**, **ADK29622.1**, **ADK29663.1**, and **ADK28404.1**, respectively) were recovered from GenBank, NCBI (http://ncbi.nlm.nih.gov) in FASTA format and used in further analysis. The presence of signal peptide was evaluated using the SignalP program (http://www.cbs.dtu.dk/services/SignalP/). SignalP server predicts the presence and location of signal peptide cleavage sites in amino acid sequences from different organisms based on a combination of several artificial neural networks. Sequences containing the signal peptide were treated and this region was removed.

Epitopes prediction

Immunodominant B cells epitopes were mapped on the treated sequences using the software ABCpred (http://crdd.osdd.net/raghava/abcpred/), which predicts B cell epitopes in an antigen sequence using an artificial neural network and fixed-length patterns (a score value greater than 0.9 was used in the analysis), and BepiPred (http://www.cbs.dtu.dk/services/BepiPred/) that predicts B cell epitopes from a protein sequence using a Random Forest algorithm trained on epitopes and non-epitopes amino acids determined from crystal structures. A threshold of 0.7 was used as the cut-off point. Since protection against *C. pseudotuberculosis* is mainly dependent on an immune response involving the production of cytotoxic T lymphocytes (CTL) (Bastos et al. 2012), predictions of CTL epitopes are important for vaccine design. Additionally, the prediction of helper T lymphocytes (HTL) epitopes is also important because these cells are responsible for inducing efficient antibody response or CTL response (Lata et al. 2018). Hence, to predict these epitopes, the sequences of the four proteins were submitted to the programs: (i) TepiTool (http://tools.iedb.org/tepitool/; mouse MHC class I alleles H-2-Db, H-2-Kk, H-2-Kb, and H-2-Kd, and mouse MHC class II alleles H2-IAb, H2-IAd, and H2-IEd), which is part of the Immune Epitope Database (IEDB) and provides some of the top major histocompatibility complex (MHC) binding prediction algorithms for several species. After MHCII analyses on TepiTool, only sequences that were classified with a rank lower than 5.0 were used in this study. (ii) NetMHC (http://www.cbs.dtu.dk/services/NetMHC/; mouse MHC alleles H-2-Db, H-2-Dd, H-2-Kb, H-2-Kd, H-2-Kk, and H-2-Ld), which generates high-precision peptide binding predictions to the MHC class I through predictions based on artificial neural networks (only epitopes with an affinity value of less than 1000.0 were used in the study); and (iii) NetMHCII (http://www.cbs.dtu.dk/services/NetMHCII/; mouse MHC class II alleles H-2-IAb, H-2-IAd, H-2-IAk, H-2-IAs, H-2-IAu, H-2-IEd, and H-2-IEk), which was constructed using an extended data set of quantitative MHC-peptide binding affinity data obtained from the IEDB (only epitopes with an affinity value of less than 500.0 were used in the study). The results obtained in each analysis were tabulated in Excel spreadsheets and an algorithm was generated for the identification and labeling of these epitopes in the protein structures that were predicted using the I-TASSER server, which produces a three-dimensional structure model based on the Protein Data Bank database.

Conservation assessment

In order to evaluate homologs of NanH, PknG, SpaC, and SodC proteins in the proteomes of *Ovis aries* (UniProt Proteomes: UP000002356) and *Capra aegagrus hircus* (UniProt Proteomes: UP000291000), two of the main species affected by CLA, a BLASTP (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins) was performed. Protein sequences with *e*-value < 1 e-10 were considered as homologs. NCBI

BLASTP (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins) was also used to identify the presence of sequences homologous to the target proteins in this study (GenBank accession nos. **ADK28179.1** (NanH), **ADK29622.1** (PknG), **ADK29663.1** (SpaC), and **ADK28404.1** (SodC)) in the most representative genomes of *C. pseudotuberculosis* (1002: **CP001809**; 3–99-5: **NC_016781**; C231: **CP001829**; CP13: **CP014998**; and MIC6: **NZ_CP019769**). MUSCLE tool (https://www.ebi.ac.uk/Tools/msa/muscle/) was used to perform the multiple alignments between the analyzed sequences.

Prediction of transmembrane helices in SpaC protein

SpaC, previously described as a membrane protein (Santana-Jorge et al. 2016), was also evaluated for the presence of transmembrane helices in its structure. For this, the TMHMM (http://www.cbs.dtu.dk/services/TMHMM-2.0/) program was used. TMHMM is a method based on a Hidden Markov Model, in which models are estimated both by maximum likelihood and a discriminative method, and, additionally, a method for reassignment of the membrane helix boundaries was also applied (Sonnhammer et al. 1998).

Chimeric design and three-dimensional protein structure

After the identification of the immunodominant epitopes and selection of the most promising ones for the composition of the multi-epitope vaccine, these were ligated by glycine linkers containing five replicates of this amino acid. The chimeric construct obtained was sent for chemical synthesis by the company GenOne Biotechnologies (Rio de Janeiro, Brazil) and the gene sequence was delivered cloned into the pAE vector. The DNA coding sequence of the multi-epitope chimeric protein has been deposited at GenBank under the accession number **MZ501492**. The three-dimensional (3D) structure of the chimeric sequence was predicted using the I-TASSER server, which generates full-length atomic structural models from multiple threading alignments and iterative structural assembly simulations followed by atomic-level structure refinement (Yang and Zhang 2015). The modeled structure with the highest assigned confidence score (*C*-score) was selected and submitted to further analysis.

Physicochemical parameter, allergenicity, antigenicity, and solubility

After glycine linkers assembly, the chimera sequence was submitted to characterization analyses of its physicochemical properties. To check for allergenicity the AlgPred tool (http://crdd.osdd.net/raghava/algpred/) was used. The software analyzes the degree of allergenicity based on the potential for activating IgE immunoglobulins, which are responsible for the stimulation of the immune response. The antigenic potential of the proteins was verified by VaxiJen30 v2.0 software at http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html. Analyses on isoelectric point (pI) instability index, *in vivo* half-life, molecular weight, and hydropathicity (GRAVY) were performed on ProtParam at http://web.expasy.org/protparam. Protein solubility and other parameters, such as Protein-Sol (https://protein-sol.manchester.ac.uk/), were verified using SolPro on SCRATCH (http://scratch.proteomics.ics.uci.edu/cgi).

Expression, purification, and immunogenicity of the recombinant chimera

The plasmid containing the recombinant chimeric sequence was transformed into different strains of *Escherichia coli* in order to obtain the highest level of expression and yield of the protein. The recombinant plasmid (pAE/chimera) was introduced by electroporation into the expression strain that showed the best protein yield; One Shot™ BL21 Star™ (DE3) chemically competent *E. coli* (catalog number: C601003, Invitrogen). Induction of expression was achieved by the addition of 1 mM isopropyl-a-D-thiogalactoside, and the culture was maintained under agitation at 37 °C for 3 h. Expression of the recombinant protein was confirmed by Western blot using horseradish peroxidase (HRP)-conjugated anti-6 × His tag monoclonal antibody (Sigma-Aldrich). Purification was performed by affinity chromatography on a HisTrap™ Sepharose nickel column (GE Healthcare). The purity of the samples was determined by SDS-PAGE and the concentration defined by the BCA kit (Pierce). To determine the immunogenicity of the recombinant chimera, a Western blot analysis was performed as described by (Silva et al. 2018) and using serum samples obtained from the female Balb/C mice immunized with the following formulations: saline solution 0.9% (G1) and recombinant chimera + adjuvant saponin (G2) (data not shown). G2 mice were immunized with a total dose of 50 µg of recombinant protein combined with 7.5 µg of saponin as the adjuvant in 200 µl of a sterile saline solution. The vaccine preparations were kept under agitation at 4 °C overnight before use. The immunized mice received two vaccine doses by subcutaneous route (via s.c.) at a 21-day interval. Blood samples were collected on day 0 (preimmune), day 21 (before the second immunization), and 42 days after the first immunization. Animal experimentation was performed following the National Institute of Health guide for the care and use of laboratory animals (NIH Publications No. 8023, revised 1978) and was approved by the Ethics Committee on Animal Experimentation of the Federal University of Pelotas (protocol number 2442).

Antigenicity of recombinant chimera

Western blot analysis was performed using one sera *pool* and two individual sera samples obtained from naturally infected goats. For the test, male and female goats of no defined breed (SRD) were used. Sera samples were collected in the city of Mossoró, Rio Grande do Norte state, Brazil (lat 5° 11′ 17″ S, long 37° 20′ 39″ W), more precisely at the experimental farm of the Federal Rural University of the Semi-Arid (UFERSA), and at private rural properties. Blood sample collections were performed by jugular venipuncture, and all procedures were performed per the recommendations of the National Council for the Control of Animal Experimentation (CONCEA) and the National Institute of Health guide for the care and use of laboratory animals (NIH Publications No. 8023, revised 1978). The experiment was approved by the Ethics Committee for Animal Use (CEUA) of UFERSA through permit no. 18/2018.

Seropositive and seronegative samples were previously confirmed by indirect ELISA using *C. pseudotuberculosis* 1002 secreted antigens (Seyffert et al. 2010). A *pool* of goat serum (composed of three seronegative samples) and one other serum sample, both negative for CLA, were used as negative controls and were collected from goats from two farms with rigorous sanitary control and without CLA history in the last 3 years. Two seropositive goat samples and a positive *pool* (composed of three seropositive samples) were collected from two other farms with a CLA history.

Samples containing 10 µg of chimeric protein were mixed with a buffer (100 mM Tris–HCl, pH 6.8, 100 mM 2-mercaptoethanol, 4% SDS, 0.2% bromophenol blue, and 20% glycerol) under reducing conditions, then heated to 100 °C for 10 min, and electrophoresed on 12% SDS-PAGE gels. Proteins were electrotransferred onto a nitrocellulose membrane (GE Healthcare), which was then blocked with 5% skim milk diluted in PBS-T for 1 h at 37 °C and cut into strips. Then, the pool of negative and positive sera for CLA diluted 1:100 was added to a strip and incubated at 37 °C for 1 h. The membranes were washed with PBS-T and incubated at 37 °C for 1 h with anti-goat antibody conjugated with horseradish peroxidase (Sigma-Aldrich) diluted 1:4000 in PBS-T. Reactive bands were visualized using 3,3′-diaminobenzidine and H_2O_2.

Results

Epitopes prediction

The total number of putative epitopes identified in our study for NanH, PknG, SodC, and SpaC proteins is described in Table 1. A greater number of putative epitopes were predicted in the structure of PknG protein, whereas few epitopes were pointed for SodC protein.

The *in silico* prediction of B cell epitopes in the BepiPred program led to the identification of 19, 27, 6, and 29 putative epitopes on NanH, PknG, SodC, and SpaC, respectively. A threshold of 0.7 was used as the cut-off point in the analysis. Only sequences containing fifteen or more amino acids in their composition were chosen to proceed to the chimera construction steps because such characteristic is pivotal to their potential to behave as putative epitopes.

Aiming to increase the reliability of the results, a new analysis was performed in an additional software, the ABCPred. The results showed 67, 74, 15, and 80 putative epitopes on NanH, PknG, SodC, and SpaC, respectively. Only epitopes with the best scores (greater than 0.9), which are indicative of them being antigenic, were selected.

To identify peptide sequences capable of generating T cell responses associated with MHC class II, the NetMHCII, and TepiTool servers were used. NetMHCI and TepiTool programs were applied to investigate the ability of an epitope to be recognized by CD8 + T cells and to stimulate a cellular immune response.

Conservation assessment

The similarity analysis performed in the proteome of *O. aries* and *C. a. hircus* did not lead to any hits with an e-value below the established *e*-value threshold (1e-10) and above 40% of similarity. However, SodC presented ~ 34% of similarity when aligned both to the superoxide dismutase from *O. aries* (UniProt: P09670) and *C. a. hircus* (UniProt: A0A452EM32), while PknG presented ~ 27% when aligned to an uncharacterized kinase from *O. aries* (UniProt: W5PGI1) and *C. a. hircus* (UniProt: A0A452F862). Although not "significant" based on the proposed criteria, these alignments are in the "twilight zone" of sequence

Table 1 Total number of B cell and T cell epitopes predicted in different programs according to the criteria established for each of them. strong binding (SB) and weak binding (WB) to MHC classes I and II

Protein	NetMHC		NetMHCII		TepiTool		BepiPred	ABCPred
	WB	*SB*	*WB*	*SB*	*MHC class I*	*MHC class II*		
NanH	14	11	19	22	14	34	19	67
PknG	20	21	26	15	24	37	27	74
SodC	2	2	4	23	2	2	6	15
SpaC	18	18	21	14	24	19	29	80

alignment significance (Rost 1999). Even so, to maintain quality, these regions were not included in the final chimera.

The proteins target sequences proved to be well preserved in all strains evaluated, apart from a small change in the final residues of the SpaC sequence. No other mutated sites were observed.

Prediction of transmembrane helices in proteins

To identify the transmembrane region of the protein SpaC, we used the TMHMM program. According to the analysis carried out prior to the selection of epitopes, SpaC protein presented transmembrane helices in its C-terminal portion/residues 767–789. This region was not included in the synthesized chimeric sequence.

Chimeric design and protein three-dimensional structure

The chimera was built in three steps: (I) immunogenic epitopes present in PknG, SpaC, SodC, and NanH proteins were identified in the software Bepipred 1.0, using a threshold higher than 0.7. (II) Epitopes containing more than 15 amino acids in their sequences were selected for the next step. (III) Of these epitopes, the regions also identified in the ABCPred software were chosen to compose the chimera. Bepipred 1.0 was chosen as the most appropriate software for this study because it evaluates linear epitopes. Since the chimeric sequence is an entirely new construction, nonexistent in nature, a conformational prediction could contribute negatively. According to the established criteria, 4, 5, 1, and 4 epitopes of PknG, SpaC, SodC, and NanH, respectively were used for constructing the chimera. The final sequence was cloned into the pAE vector between the *Bam*HI and *Eco*RI sites and forwarded for synthesis. The portion corresponding to the selected epitopes in each protein is shown in Fig. 1. As the cellular immune response is the one desired to be elicited by an effective vaccine against *C. pseudotuberculosis*, the location of the chosen epitopes is not as important since all antigens would be processed by the proteasome. Thus, the image below is only illustrative, the criteria for choosing the epitopes were only the ones mentioned above.

The protein produced at the end of the process, containing 5 glycine linkers between the epitopes (Fig. 2B), has a molecular weight of approximately 52 kDa and a molar extinction coefficient of 46,610. The chimeric sequence was submitted to I-TASSER (https://zhanglab.ccmb.med.umich.edu/I-TASSER/) and a three-dimensional structure model based on the Protein Data Bank database was obtained (Fig. 2A).

Physicochemical parameter, allergenicity, antigenicity, and solubility

The chimera sequence composed of selected epitopes from the four proteins of *C. pseudotuberculosis* (NanH, PknG, SodC, and SpaC) was submitted to the Algpred program and did not present interaction sites with E immunoglobulins, being classified as non-allergenic. This result further justifies the need to select specific epitopes for the composition of vaccine formulations since possible allergenic regions can be excluded from the final sequence.

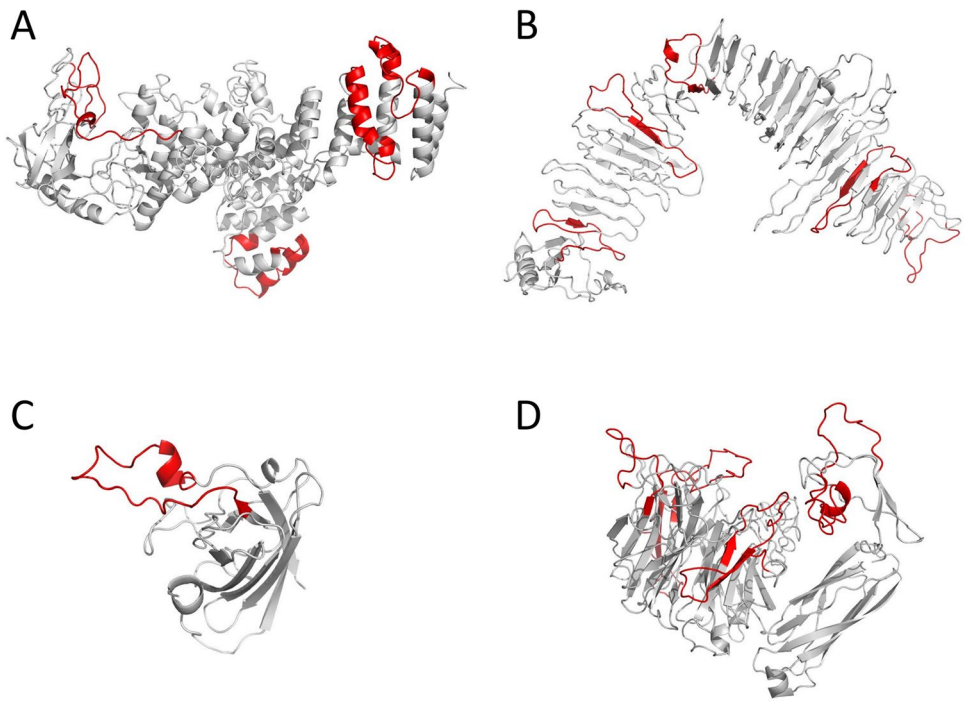

Fig. 1 Immunogenic epitopes (red) selected to compose the chimera marked on the three-dimensional structure of the proteins PknG (**A**), SpaC (**B**), SodC (**C**), and NanH (**D**)

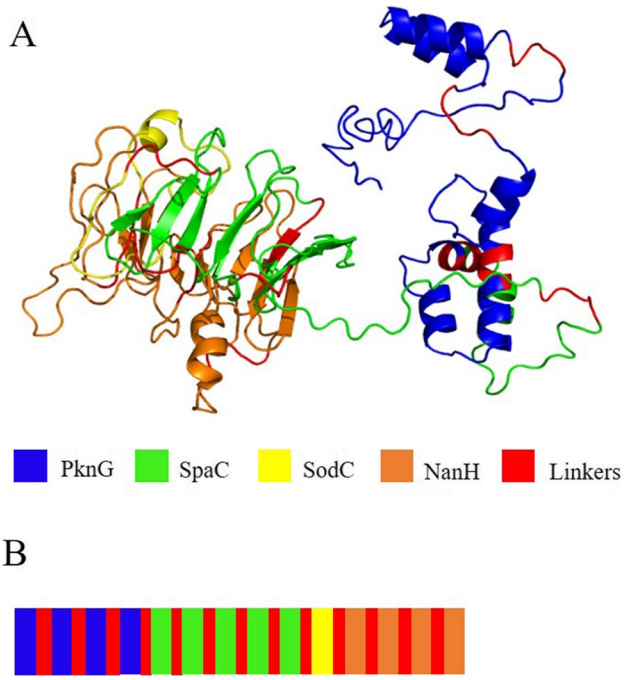

Fig. 2 Chimera model. **A** Tertiary protein structures predicted by I-TASSER were rendered using PyMOL. **B** Schematic representation of the chimera's structural organization formed by the epitopes of PknG, SpaC, SodC, and NanH proteins, distributed in this order and joined by linkers containing 5 glycines

VaxiJen was used to verify the antigenic potential of the chimera. The program uses a threshold (0.4) for this model and any value below it indicates the protein as non-antigenic. The chimera showed antigenicity greater than 1.3, which may infer a greater activation of the immune response (Table 2).

The data obtained in ProtParam, including molecular weight, are available in Table 2. The chimera sequence showed a pI of approximately 6.7, considered to be acidic. In addition, GRAVY, a tool that provides the hydropathicity of proteins, was also used and the protein obtained a negative GRAVY value, which represents hydrophilic characteristics compatible with aqueous media. Furthermore, the chimeric protein proved to be unstable. This result can be explained by the fact that the construction is a synthetic form that does not exist in nature. Nevertheless, other promising results obtained here reaffirm the potential of its applicability in vaccine construction. ProtParam also identified the half-life for this protein *in vivo* in *E. coli*. The protein had a relevant period of more than 10 h of estimated half-life.

In Table 2, the solubility value of the scale (QuerySol) represents the predicted solubility for the protein of interest. In this analysis, the population mean for the experimental data set (PopAvrSol) is 0.45 and, therefore, any solubility value greater than 0.45 is predicted to have a higher solubility than the average soluble *E. coli* protein. The QuerySol value obtained by the chimera was approximately 0.6, which indicates expression in a soluble form and not in the form of inclusion bodies.

Expression, purification, and immunogenicity of the recombinant chimera

Expression of the recombinant chimera was confirmed by Western blot, performed using a monoclonal anti-6 × His antibody, through the observation of a reactive band at the expected size of 52 kDa (Fig. 3A). The protein was expressed in *E. coli* BL21 Star (DE3) and obtained in its soluble form, with a yield of 7 mg/L after purification. The results obtained by Western blot using the anti-chimeric polyclonal serum for immunogenicity evaluation demonstrated that the recombinant chimera was recognized by the sera of animals immunized with the proposed vaccine formulation (Fig. 3B).

Antigenicity of recombinant chimera

In the analysis of antigenicity by Western blot, the recombinant chimera was identified by sera of goat naturally infected with *C. pseudotuberculosis* (Fig. 4), suggesting that the expressed protein maintained epitopes similar to those of native NanH, PknG, SodC, and SpaC proteins. Negative CLA sera did not react with the protein.

Discussion

Corynebacterium pseudotuberculosis is a gram-positive facultative intracellular pathogen and the etiological agent of caseous lymphadenitis in both sheep and goats (Guimarães et al. 2011). In Brazil, CLA is one of the most prevalent diseases of goats and, consequently, has an economic impact because it leads to reduced milk production and condemnation of carcasses of infected animals (Bastos et al. 2012). Since the vaccines currently available for CLA have low efficacy rates, the search for a

Table 2 Physicochemical characterization of the chimera

Antigens	Algpred	VaxiJen	PI	Instability index	Molecular weight	GRAVY	Estimated half-life in *Escherichia coli*, *in vivo*	QuerySol
Chimera	Non-allergenic	1.3781	6.76	Unstable	52,314.47	−1.205	> 10 h	≅0.6

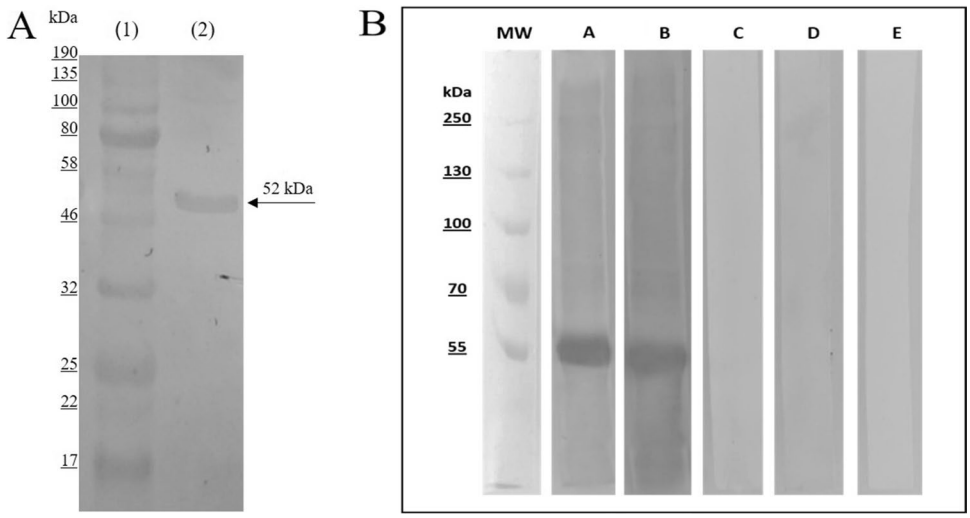

Fig. 3 Western blot for chimera characterization. **A** Characterization of the identity of the recombinant chimera using a monoclonal anti-6×His antibody. Lanes: [1] pre-stained protein ladder; [2] recombinant chimera. **B** Characterization of chimeric protein-induced immunogenicity through Western blot using polyclonal serum from mice immunized with recombinant chimera. **A**, **B**: sera *pools* of the animals immunized with recombinant chimera; **C**, **D** sera *pools* of the animals immunized with saline solution 0.9%; **E** negative control (serum absent in the reaction)

new vaccine formulation is constant and several potential antigens have been reported (Dorella et al., 2006). In this context, NanH, PknG, SodC, and SpaC proteins have already been identified as promising molecules for the development of vaccines against CLA (Santana-Jorge et al. 2016), and based on our conservation analysis, these proteins presented highly conserved sequences when compared to other *Corynebacterium* strains by multiple sequence alignment and no significant similarity when compared to the host (*O. aries*) proteome by BLAST.

NanH, also called sialidase, is an extracellular neuraminidase (Trost et al. 2010) that contributes to the recognition of sialic acids exposed on the surface of host cells (Kim et al. 2010). In pathogenic bacteria, these enzymes are considered important virulence factors as they assist with the propagation or invasion of bacteria within the host (Sheu et al. 2002). PknG protein has already been identified in species of the genus *Corynebacterium*, *Mycobacterium*, *Nocardia*, and *Rhodococcus*, and it is said to be responsible for inhibiting the formation of phagolysosomes, consequently allowing the intracellular survival of mycobacteria (Niebisch et al. 2006).

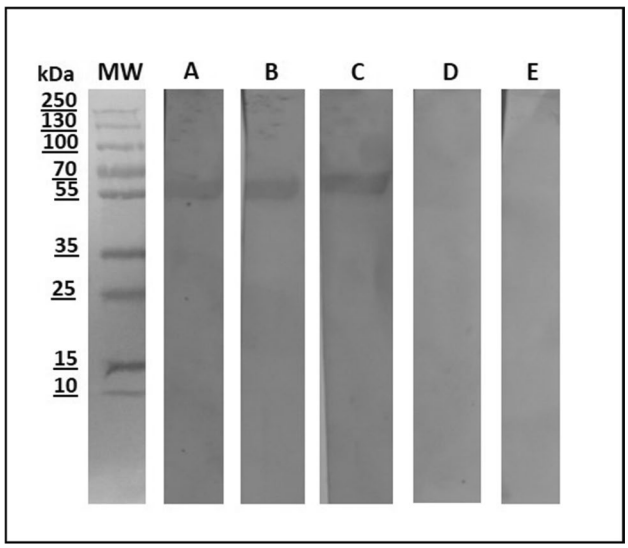

Fig. 4 Western blot to determine the chimeric protein-induced antigenicity. **A** CLA-positive goat serum *pool*; **B** positive sample; **C** positive sample; **D** CLA-negative goat serum *pool*; **E** negative sample. MW: pre-stained protein ladder

Furthermore, SodC is a zinc-dependent superoxide dismutase predicted as a protein located on the bacteria cytoplasmic membrane (Santana-Jorge et al. 2016). This location suggests that it can protect the surface of *C. pseudotuberculosis* against superoxides generated externally by mammalian host cells (Trost et al. 2010). SODs are involved in the pathogenicity of countless organisms. In *Mycobacterium tuberculosis*, they contribute to the pathogen's resistance against oxidative products generated by activated macrophages, which produce reactive oxygen species with potent activity (Piddington et al. 2001). In turn, SpaC is an adhesion protein located at the end of the additive pili of *C. pseudotuberculosis* (Santana-Jorge et al. 2016), such location is consistent with the identification of a transmembrane region at its C-terminus region. Adhesins, as part of the pili, assist in the initial contacts between bacteria and host cells by promoting ligand-receptor interactions, favoring the delivery of virulence factors, essentially toxins, and facilitating intracellular invasion (Mandlik et al. 2007).

Recently, NanH and PknG were evaluated in recombinant subunit vaccines, and promoted a 60% and 20% survival rate, respectively, after mice were challenged with virulent *C. pseudotuberculosis* (Silva et al. 2020). So, we hypothesized that these rates could be increased using improved chimeric antigens built from the rational design of immunodominant epitopes, as proposed in our study.

The approach used by immunological bioinformatics allows us to reduce the time and cost necessary for the development of vaccines, mainly by eliminating the need

to cultivate the microorganism of interest (Kazi et al. 2018). Currently, immunoinformatics is the way forward in the identification of vaccine candidates and epitope-based vaccines have become an attractive methodology to trigger protective immune responses in hosts (Oli et al. 2020).

Before the production of recombinant protein vaccines for clinical application, extensive testing is needed. Recombinant protein vaccines must be stable and safe, have high immunogenicity, stability in various conditions, and high specificity, and should preferably be soluble (Validi et al. 2018). Recent advances in the field of immunoinformatics seem to ensure the rapid and safe development of subunit vaccines and, therefore, this approach has been preferred for the development of a multi-epitope vaccine for CLA. In our study, we report for the first time the construction and production of a recombinant chimera containing multi-epitopes of the proteins NanH, PknG, SodC, and SpaC of *C. pseudotuberculosis* by using immunoinformatics tools.

Most immunoinformatics tools have been developed for the human species and, therefore, their use for other animal species is still a challenge. Nevertheless, when choosing to use some tools, especially those for the detection of epitopes, we took into consideration that the immune system of these animals tends to have certain similarities and also the fact that there are no adequate tools for such study in goats. However, when it comes to the tools to predict B cell epitopes, some attention is still needed since an ideal tool should consider compositional biases in datasets. A realistic prediction of antigenic epitopes, which are preferably critically recognized, depends on the composition of the protein. Most of the epitopes prediction methods available are based on amino acid properties, including hydrophilicity, solvent accessibility, secondary structure, flexibility, and antigenicity (Soria-guerra et al. 2014; Osman et al. 2018).

After analyses of the selected epitopes, a polypeptide sequence of 492 aa was constructed, in which a six-histidine selection label was later added, as a consequence of cloning and expression on pAE vector, to facilitate purification. For designing our multivalent vaccine using immuno-informatics approaches, 14 regions from these antigens were selected as B cell and T cell immunodominant epitopes. Of these selected epitopes, four are from PknG protein, five from SpaC, one from SodC, and four from NanH protein. These epitopes were linked by linkers containing 5 glycines. Gly-rich linkers are naturally occurring docs, connecting domains within proteins, and allowing discrete functions of domains (Priyanka et al. 2013). In addition, codon optimization was performed to improve transcriptional and translational efficiency and to facilitate high-level expression of the recombinant protein in *E. coli*.

Physicochemical parameters analyses indicated that the protein would be expressed in a soluble way and that its half-life in *E. coli* would be 10 h. This was confirmed after the expression and purification of the protein and Western blot performed with the monoclonal anti-6 × His antibody which indicated the successful expression of the chimeric protein (Fig. 3A).

The Western blot performed with polyclonal serum obtained from the immunized animals also produced consistent results (Fig. 3B). These data show that an immune response was elicited against the chimeric protein in the experimental group, while in the negative control no reactive band was generated, demonstrating the immunogenic activity of the chimeric construction.

In addition, the recombinant chimera was recognized by CLA-positive goat sera (Fig. 4), evidencing that this construct was able to maintain epitopes that are similar to the ones found on native NanH, PknG, SodC, and SpaC proteins. NanH and PknG proteins, when evaluated in recombinant subunit vaccines for CLA, had already shown good antigenic potential, as they were capable of eliciting high levels of IgG antibodies in immunized animals (Silva et al. 2020). So far, SpaC and SodC proteins have not been evaluated in vaccine trials. Altogether, results showed that the recombinant chimera can be potentially used as an antigen for the detection of specific antibodies of *C. pseudotuberculosis* in ELISA tests and/or in vaccine research against caseous lymphadenitis. The high conservation observed when comparing different strains of *C. pseudotuberculosis* indicates that the chimera might be protective also in heterologous challenges, as no variations were observed in the sites predicted as immunogenic during its design.

In recent years, immunoinformatics has gained prominence due to its numerous advantages. In this context, the design of new and improved antigens by assembling immunodominant epitopes on chimeric proteins has allowed the improvement of the immunogenicity, stability, and safety of vaccines against different pathogens. The chimera built in this work was characterized as non-allergenic and presented antigenic potential. In addition, it was able to induce the production of antibodies in mice immunized with it, a fact confirmed by Western blot using sera from these animals. Still, the chimera was also recognized by sera from goats naturally infected by CLA, which demonstrates its promising role as a vaccine target and use in diagnostics. Therefore, the next step is to evaluate the chimera built here in vaccine formulations and immunodiagnostic strategies.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00253-021-11619-x.

Acknowledgements We acknowledge Dr. Luciano da Silva Pinto, coordinator of the Bioinformatics and Proteomics Laboratory of the Federal University of Pelotas, for providing technical and intellectual support and for offering additional information and comments that helped improve and clarify this manuscript. We gratefully acknowledge the Coordination for the Improvement of Higher Education Personnel (CAPES, Brazil) and the National Council for Scientific and Technological Development (CNPq, Brazil) for the scholarships provided to some of the authors.

Author contribution SB, FK, and MS conceived the experiments. MS, RP, FB, NS, and LM conducted the experiments. MS, RW, and FK performed the analyses *in silico*. MS, RW, and FK performed the data analysis and wrote the manuscript. SB, MA, and FK critically revised and helped writing the work. FB contributed with inputs for conducting *in vitro* experiments. All authors read and approved the manuscript.

Data availability The authors declare they have no relevant data and material to disclose.

Code availability Not applicable.

Declarations

Ethics approval The animal experiment was performed following the National Institutes of Health guide for the care and use of laboratory animals (NIH Publications No. 8023, revised 1978) and was approved by the Ethics Committee on Animal Experimentation of the Federal University of Pelotas (protocol number 2442) and the Ethics Committee for Animal Use (CEUA) of UFERSA through permit no. 18/2018.

Consent to participate Not applicable.

Consent for publication Not applicable.

Conflict of interest The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: patent application for the recombinant chimera in the National Institute of Industrial Property (INPI, Brazil) under the number BR 10 2020 006421 5.

References

Bastos BL, Portela RWD, Dorella FA, Ribeiro D, Seyffert N, de Castro TL, P, Miyoshi A, Oliveira SC, Meyer R, Azevedo V (2012) *Corynebacterium pseudotuberculosis*: immunological responses in animal models and zoonotic potential. Immunol, J Clin Cell S4:1–15. https://doi.org/10.4172/2155-9899.S4-005

Bezerra FSB, de Oliveira Silva MT, de Fátima Silva Rezende A, Sena-Lopes A, de Pinho RB, Seixas FK, CollaresPortela TVRWD, de Carvalho Azevedo VA, Borsuk S (2021) Saponin-adjuvanted recombinant vaccines containing rCP00660, rCP09720 or rCP01850 proteins against *Corynebacterium pseudotuberculosis* infection in mice. Vaccine 28:2568–2574. https://doi.org/10.1016/j.vaccine.2021.03.062

da Costa Barnabé NN, Alves JRA, de Farias AEM, Alves FSF, Faccioli-Martins PY, Pinheiro RR, de Azevedo SS, Alves CJ (2020) Assessment of caseous lymphadenitis in goats in a slaughterhouse in the Brazilian semi-arid region and estimates of economic losses due to carcass condemnation. Semin Agrar 41:2655–2668. https://doi.org/10.5433/1679-0359.2020v41n6p2655

de Oliveira NR, Jorge S, Gomes CK, Rizzi C, Pacce VD, Collares TF, Monte LG, Dellagostin OA (2017) A novel chimeric protein composed of recombinant *Mycoplasma hyopneumoniae* antigens as a vaccine candidate evaluated in mice. Vet Microbiol 201:146–153. https://doi.org/10.1016/j.vetmic.2017.01.023

de Oliveira Silva MT, de Pinho RB, da Rocha Fonseca B, Bezerra FSB, Sousa FSS, Seixas FK, Collares T, Nascimento RJM, Portela RW, Azevedo VAC, Borsuk S (2020) NanH and PknG putative virulence factors as a recombinant subunit immunogen against *Corynebacterium pseudotuberculosis* infection in mice. Vaccine 38:8099–8106. https://doi.org/10.1016/j.vaccine.2020.11.010

de Sá Guimarães A, do Carmo FB, Pauletti RB, Seyffert N, Ribeiro D, Lage AP, Heinemann MB, Miyoshi A, Azevedo V, Guimarães Gouveia AM (2011) Caseous lymphadenitis: epidemiology, diagnosis, and control. IIOAB J 2:33–43

Dorella FA, Pacheco LG, Seyffert N, Portela RW, Meyer R, Miyoshi A, Azevedo V (2009) Antigens of *Corynebacterium pseudotuberculosis* and prospects for vaccine development. Expert Rev Vaccines 8:205–213. https://doi.org/10.1586/14760584.8.2.205

Dorella FA, Pacheco LGC, Oliveira SC, Miyoshi A, Azevedo V (2006) *Corynebacterium pseudotuberculosis*: microbiology, biochemical properties, pathogenesis and molecular studies of virulence. Vet Res 37:201–218. https://doi.org/10.1051/vetres:2005056

Dormitzer PR, Grandi G, Rappuoli R (2012) Structural vaccinology starts to deliver. Nat Publ Gr 10:807–813. https://doi.org/10.1038/nrmicro2893

Droppa-Almeida D, Vivas WLP, Silva KKO, Rezende AFS, Simionatto S, Meyer R, Lima-Verde IB, Delagostin O, Borsuk S, Padilha FF (2016) Recombinant CP40 from *Corynebacterium pseudotuberculosis* confers protection in mice after challenge with a virulent strain. Vaccine 34:1091–1096. https://doi.org/10.1016/j.vaccine.2015.12.064

Kazi A, Chuah C, Majeed ABA, Leow CH, Lim BH, Leow CY (2018) Current progress of immunoinformatics approach harnessed for cellular- and antibody-dependent vaccine design. Pathog Glob Health 112:123–131. https://doi.org/10.1080/20477724.2018.1446773

Kim S, Oh DB, Kwon O, Kang HA (2010) Identification and functional characterization of the NanH extracellular sialidase from *Corynebacterium diphtheriae*. J Biochem 147:523–533. https://doi.org/10.1093/jb/mvp198

Lata KS, Kumar S, Va V, Sharma P (2018) Exploring Leptospiral proteomes to identify potential candidates for vaccine design against Leptospirosis using an immunoinformatics approach. Sci Rep 8:1–15. https://doi.org/10.1038/s41598-018-25281-3

Mandlik A, Swierczynski A, Das A, Ton-that H (2007) Pili in Gram-positive bacteria: assembly, involvement in colonization and biofilm development. Trends Microbiol 16:33–40. https://doi.org/10.1016/j.tim.2007.10.010

Niebisch A, Kabus A, Schultz C, Weil B, Bott M (2006) Corynebacterial protein kinase G controls 2-oxoglutarate dehydrogenase activity via the phosphorylation status of the OdhI protein. J Biol Chem 281:12300–12307. https://doi.org/10.1074/jbc.M512515200

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

Osman AY, Nordin ML, Kadir AA, Saharee AA (2018) The epidemiology and pathophysiology of caseous lymphadenitis: a review. J Vet Med Res 5:1129

Piddington DL, Fang FC, Laessig T, Cooper M, Orme IM, Buchmeier NA, Cooper AM (2001) Cu, Zn superoxide dismutase of *Mycobacterium tuberculosis* contributes to survival in activated macrophages that are generating an oxidative burst. Infect Immun 69:4980–4987. https://doi.org/10.1128/IAI.69.8.4980

Priyanka V, Chichili R, Kumar V, Sivaraman J (2013) Linkers in the structural biology of protein-protein interactions. Protein Sci 22:153–167. https://doi.org/10.1002/pro.2206

Rezende AFS, Brum AA, Bezerra FSB, Braite DC, de Sá GL, Thurow HS, Seixas FK, Azevedo VAC, Portela RW, Borsuk S (2020) Assessment of the acid phosphatase CP01850 from *Corynebacterium pseudotuberculosis* in DNA and subunit vaccines formulations against caseous lymphadenitis. Brazilian J Vet Anim Sci 72:199–207

Rost B (1999) Twilight zone of protein sequence alignments. Protein Eng 12:85–94. https://doi.org/10.1093/protein/12.2.85

Santana-Jorge KTO, Santos TM, Tartaglia NR, Aguiar Edgar L, Souza RFS, Mariutti RB, Eberle RJ, Arni RK, Portela RW, Meyer R, Azevedo V (2016) Putative virulence factors of *Corynebacterium pseudotuberculosis* FRC41: vaccine potential and protein expression. Microb Cell Fact 15:1–13. https://doi.org/10.1186/s12934-016-0479-6

Seyffert N, Guimarães AS, Pacheco LGC, Portela RW, Bastos BL, Dorella FA, Heinemann MB, Lage AP, Gouveia AMG, Meyer R, Miyoshi A, Azevedo V (2010) High seroprevalence of caseous lymphadenitis in Brazilian goat herds revealed by *Corynebacterium pseudotuberculosis* secreted proteins-based ELISA. Res Vet Sci 88(50):55. https://doi.org/10.1016/j.rvsc.2009.07.002

Sheu SY, Tseng HJ, Huang SP, Chien CH (2002) Cloning, expression, and deletion analysis of large nanH of *Clostridium perfringens* ATCC 10543. Enzyme Microb Technol 31:794–803. https://doi.org/10.1016/S0141-0229(02)00177-1

Silva MTDO, Bezerra FSB, de Pinho RB, Begnini KR, Seixas FK, Collares T, Portela RD, Azevedo V, Dellagostin O, Borsuk S (2018) Association of *Corynebacterium pseudotuberculosis* recombinant proteins rCP09720 or rCP01850 with rPLD as immunogens in caseous lymphadenitis immunoprophylaxis. Vaccine 36:74–83. https://doi.org/10.1016/j.vaccine.2017.11.029

Sonnhammer EL, von Heijne G, Krogh A (1998) A hidden Markov model for predicting transmembrane helices in protein sequences. Intell Syst Mol Biol 6:175–182

Soria-Guerra RE, Nieto-Gomez R, Govea-Alonso DO, Rosales-Mendoza S (2014) An overview of bioinformatics tools for epitope prediction: implications on vaccine development. J Biomed Inform 53:405–414. https://doi.org/10.1016/j.jbi.2014.11.003

Trost E, Ott L, Schneider J, Schröder J, Jaenicke S, Goesmann A, Husemann P, Stoye J, Dorella FA, Rocha FS, de Castro Soares S, D'Afonseca V, Miyoshi A, Ruiz J, Silva A, Azevedo V, Burkovski A, Guiso N, Join-Lambert OF, Kayal S, Tauch A (2010) The complete genome sequence of *Corynebacterium pseudotuberculosis* FRC41 isolated from a 12-year-old girl with necrotizing lymphadenitis reveals insights into gene-regulatory networks contributing to virulence. BMC Genomics 11:728. https://doi.org/10.1186/1471-2164-11-728

Validi M, Karkhah A, Prajapati VK, Nouri HR (2018) Immuno-informatics based approaches to design a novel multi epitope-based vaccine for immune response reinforcement against Leptospirosis. Mol Immunol 104:128–138. https://doi.org/10.1016/j.molimm.2018.11.005

Yang J, Zhang Y (2015) Protein structure and function prediction uI-TASSER. Curr Protoc Bioinformatics 52:1–15. https://doi.org/10.1002/0471250953.bi0508s52