In silico identification of *Corynebacterium pseudotuberculosis* antigenic targets and application in immunodiagnosis

Andrea de Fátima Silva Rezende, Alexandre Antunes Brum, Carlos Guilherme Reis, Henrique Ramos Angelo, Karen Silva Leal, Mara Thais de Oliveira Silva, Simone Simionatto, Vasco Azevedo, Anderson Santos, Ricardo Wagner Portela, Odir Dellagostin and Sibele Borsuk

1Laboratório de Pesquisa em Doenças Infecciosas, Centro de Desenvolvimento Tecnológico, Biotecnologia, UFPel, Pelotas, RS 96010-900, Brazil
2Laboratório de Biologia Molecular, Faculdade de Ciências Ambientais, UFGD, Dourados, MS 79825-070, Brazil
3Laboratório de Genética Celular e Molecular, Instituto de Ciências Biológicas, UFMG, Belo Horizonte, MG 31270-901, Brazil
4Laboratório de Imunologia e Biologia Molecular, Instituto de Ciências da Saúde, UFBA, Salvador, BA 40140-100, Brazil

Caseous lymphadenitis (CLA) is a disease caused by *Corynebacterium pseudotuberculosis*. It affects mainly small ruminants and causes significant economic losses worldwide. Because symptoms are not immediately noticeable, CLA clinical diagnosis is not effective. Numerous serological tests are being developed to detect the disease in asymptomatic animals, but currently available immunoassays have problems with sensitivity. Current ELISA formats use native bacterial antigens, and recombinant proteins could be useful for improving the immunoassay parameters. The *C. pseudotuberculosis* proteins CP0126a, CP0369 and CP1957 were identified from 2097 candidate proteins by mature epitope density (MED) analysis, expressed in *Escherichia coli* and evaluated in an indirect immunoenzymic system. The CP0126a, CP0369 and CP1957 ELISAs showed 77.5 %, 92.5 % and 92.5 % specificity and 95 %, 90 % and 85 % sensitivity, respectively. Receiver operating characteristic (ROC) curve analysis showed an area under the curve of 0.874, 0.951 and 0.881, respectively. The proteins identified in silico were recognized by antibodies in the sera from infected animals without being recognized in negative samples. The ELISA assay using the rCP0369 protein as antigen had the greatest specificity and sensitivity values, followed by rCP1957. This is an interesting strategy for seroepidemiological investigations in sheep flocks due to its significant specificity and sensitivity.

INTRODUCTION

*Corynebacterium pseudotuberculosis* is the aetiological agent of caseous lymphadenitis (CLA), a disease that primarily affects small ruminants (Dorella *et al.*, 2006). CLA has a severe economic impact on the goat and sheep industry and is found throughout the world (D’Afonseca *et al.*, 2008). Transmission is facilitated by the survival of bacteria in the environment for long periods of time (Baird & Fontaine, 2007), combined with the inability of antibiotics to reach the bacteria inside the abscesses (Costa *et al.*, 2011). The identification of positive animals is important for CLA control as a basis for appropriate management to prevent the dissemination of disease in the flock.

The clinical diagnosis of CLA is not very effective. The signs and symptoms are not immediately apparent, and the visceral form of the disease is usually only detected in slaughterhouses (Soares *et al.*, 2012). Several serological
tests have been developed to detect CLA in asymptomatic animals, including the serum agglutination test (AWAD, 1960), gel immune-diffusion (Anderson & Nairn, 1984), haemolysis inhibition (Burrel, 1980), and ELISA assays using several C. pseudotuberculosis antigenic preparations as lysate (Binns et al., 2007; Kaba et al., 2001; Rebouçãs et al., 2013), secreted proteins (Paule et al., 2003; Seyffert et al., 2010), C. pseudotuberculosis exotoxins (Baird & Malone, 2010; Dercksen et al., 2000) and recombinant phospholipase (rPLD) toxin (Menzies et al., 2004; Stapleton et al., 2009; Sting et al., 1998). Few ELISA formats use recombinant proteins for the diagnosis of CLA, which can confer higher levels of specificity associated with significant sensitivity values because they use a single purified antigen.

The performance and applicability of serological assays rely on the selection of an immunodominant antigen which is recognized in the majority of C. pseudotuberculosis-infected hosts (Huerta et al., 2013). Several strains of C. pseudotuberculosis were sequenced using new sequencing technology, providing important data about this pathogen (Pethick et al., 2012a, 2012b; Pinto et al., 2012; Silva et al., 2011). Moreover, comparative proteomics, exoproteome and pan-exoproteome analysis of C. pseudotuberculosis have been important in the identification of additional targets that can be used in CLA diagnosis or vaccine development (Pacheco et al., 2011; Silva et al., 2013). The in silico-predicted pan-exoproteome of C. pseudotuberculosis strains was also determined to find potential targets (Santos et al., 2012). Target selection by the mature epitope density (MED) score is useful because it is based on the idea of ranking proteins according to the concentrations of ninemer epitopes for major histocompatibility complex (MHC I) anchoring on the exported portions of proteins (Santos et al., 2013). This approach can be used to identify proteins for vaccines or diagnosis.

The aim of this work was to select antigenic targets in the C. pseudotuberculosis genome by MED analysis, use these in immunoenzymatic assays, and evaluate their applicability for CLA diagnosis in sheep compared with the ELISA using secreted proteins from C. pseudotuberculosis.

**METHODS**

**Computational target selection method.** The target selection method adopted in the present work was published by Santos et al. (2013) and was called the mature epitope density (MED) score. The idea behind the MED score is to rank proteins according to their concentrations of ninemer epitopes for MHC I anchoring on the exported portions of proteins, excluding the cytoplasmic and membrane portions (Fig. 1). The final result is a list of exported proteins (secreted or potentially exposed on the bacterial surface) ranked by descending epitope concentration (Fig. 2). When predicting epitopes, the MED score method makes no distinction as to what MHC allele is being used for predictions. Instead, all alleles are taken into account. The NETMHC program was used to predict epitopes (Lundegaard et al., 2008).

**Strains and culture conditions.** The C. pseudotuberculosis T1 strain (provided by Dr Robert Meyer, UFBA), Escherichia coli TOP10 and E. coli BL21 Star (Invitrogen) cells were used. C. pseudotuberculosis was cultivated in BHI (brain heart infusion; Neogen) supplemented with 0.5 % Tween 80 at 37 °C for 72 h in a shaker at 200 r.p.m., or in BHI agar 1.5 %. The E. coli strains were grown in LB medium (Luria Bertani) or LB agar 1.5 % for 16 h at 37 °C. When necessary, the LB medium was supplemented with 100 µg ampicillin ml⁻¹.

**Sera.** A total of 60 serum samples were used to establish the ELISA validation parameters. The 40 negative sera were taken from sheep in CLA non-endemic areas and the 20 positive sera from symptomatic animals with isolation of C. pseudotuberculosis from caseous lesions. Additionally, an indirect ELISA using C. pseudotuberculosis secreted antigens with 99 % specificity, but with 89 % sensitivity, was used to confirm the infection status of the animals (Rebouçãs et al., 2013). Negative animals from non-endemic areas were included in the experiment only with a negative result on this ELISA. Positive animals had to present with positive C. pseudotuberculosis isolations from caseous lesions and a positive ELISA result.

Negative sera were used as a standard for evaluating the cut-off point of the ELISA for secreted proteins (Seyffert et al., 2010). The cut-off was based on the mean absorbance of the 40 negative sheep sera plus three standard deviations.

All the procedures regarding obtaining serum samples were performed by veterinarians and approved by the Ethical Committee in Animal Experimentation of the Federal University of Bahia (ICs/UFBA 01/2009).

**Expression of the recombinant proteins CP0126, CP0369 and CP1957 in E. coli.** Genes were amplified using the primers F5'-ATAGGTACCTCGAGCAACAACGGC-3' and R5'-CCGAAAGCTTTTATTTTTG for cp0102_0126a and R5'-GCCGAAGCTTTTATTTTTG for cp0102_0369a and R5'-ACCAGGCGCCCGGACTGCTGTCGTC-3' for cp1002_1957a.

Expression of the recombinant proteins CP0126, CP0369 and CP1957 in E. coli. The cut-off was

**Western blot.** To determine the identity and antigenicity of CP0126, CP0369 and CP1957 recombinant proteins, Western blot analysis was performed. CP0126a, rCP0369 and rCP1957 were mixed with SDS gel-loading buffer [100 mM Tris/HCl, pH 6.8, 100 mM 2-mercaptoethanol, 4 % (w/v) SDS, 0.2 % (w/v) bromophenol blue and 20 % (v/v) glycerol], heated under reducing conditions at 100 °C for 10 min and then subjected to gel electrophoresis on 12 % SDS-PAGE. Afterwards, the proteins were transferred electrically to a
Predictions Predicted epitopes* [50 – Average (MHC Affinity)]
Amino acids length – Epitope length + 1

**Fig. 1.** Prediction of subcellular location of hypothetical protein present in the *C. pseudotuberculosis* genome made by TMHMM 2.0 and depicting three possible subcellular localizations for the same protein. The portions in the different localizations are split, and only the parts predicted as ‘outside’ are concatenated for analysis under the MED score.

Negative and positive controls consisted of pools of 10 serum samples each, taken from non-endemic CLA area sheep and animals with positive *C. pseudotuberculosis* isolation from caseous lesions, respectively.

**Indirect ELISA using CP0126a, rCP0369 or rCP1957 proteins.** Polystyrene 96-well plates (Maxisorp, Nunc) were coated with 100 µl (200 ng) of rCP0126a, rCP0369 or rCP1957 proteins in carbonate/bicarbonate buffer (pH 9.8) and incubated at 4 °C for 16 h. Washing and negative controls (1 : 100), all in duplicate, were added to the plates. After 1 h of incubation at 4 °C and washing three times with PBS-T, 100 µl per well (1 : 4000) of anti-sheep IgG conjugated with horseradish peroxidase (Sigma) was added. After a further hour of incubation at 37 °C and three washes with PBS-T, 100 µl per well of the substrate/chromogen solution [o-phenylenediamine dihydrochloride (Sigma) in 0.4 mg phosphate/citrate buffer ml⁻¹ containing 0.04 % of 30 % hydrogen peroxide, pH 5.0] was added. The plates were incubated at room temperature in the dark for 15 min. Mean OD at 450 nm was determined using a microtiter plate reader (MR-96A, Mindray).

The best recombinant protein concentration, sera and conjugate dilution were determined via the checkerboard titration method, where the highest positive control/negative control OD value correlation was used as a reference. The antigens, positive and negative controls, and the HRP-conjugated antibody were diluted and tested at different concentrations.

**Statistical analysis.** SPSS v. 12.0 software (IBM) was used to calculate the Kappa index. The level of statistical agreement between the results obtained with the ELISAs and the CLA infectious status was then classified as previously described by Landis & Koch (1977): 0 – no agreement;
0.0–0.19 – poor agreement; 0.20–0.39 – fair agreement; 0.40–0.59 – moderate agreement; and 0.60–1.00 – almost perfect agreement. The receiver operating characteristic (ROC) curve was generated using the same software. According to an arbitrary guideline (Swets, 1988), the area under the ROC curve (AUC) value was used to determine the accuracy of the assays.

RESULTS

MED score results in *C. pseudotuberculosis*

The present work shows three targets: *Cp1002_0126a*, *Cp1002_1957* and *Cp1002_0369*, the first, second and third bars shown in descending order in Fig. 2. The experimental results presented here are the MED results of three targets among 2097 candidate proteins. There is experimental evidence of exportation concerning all proteins in Fig. 2(a) but not all proteins in Fig. 2(b) have such evidence. Proteins from Fig. 2(a) are the intersection set product between the set of all proteins from the work of Pacheco *et al.* (2011) and CP0369, CP1957 and CP0126a recombinant proteins

The three proteins were expressed as inclusion bodies in the *E. coli* BL21(DE3) Star strain. The recombinant
proteins were solubilized with 8 M urea. The yield obtained after purification and re-folding was approximately 3.0 mg rCP0126a protein l⁻¹ and approximately 6.0 mg rCP0369 and rCP1957 protein l⁻¹ (Fig. 3a). The identities of the recombinant proteins were characterized by Western blot analysis with the monoclonal anti-6×His antibody (Fig. 3d). The apparent molecular masses of the reactive bands were consistent with approximately 25 kDa for CP0126a and 35 kDa for the CP0369 and CP1957 proteins; the recombinant proteins presented the expected molecular weights. rCP0126a, rCP0369 and rCP1957 protein antigenicities were evaluated by Western blot using CLA-positive sheep sera, and the three recombinant proteins were recognized by the sera of symptomatic animals with no cross-reaction observed when negative pools were tested by Western blot (Fig. 3b, c).

**ELISAs**

Serum samples from 60 sheep were evaluated by ELISA using the three recombinant proteins, with 20 serum samples from sheep with confirmed \textit{C. pseudotuberculosis} infection and 40 serum samples from non-endemic area sheep. Fig. 4 shows the distribution of positive and negative sera results using the ELISAs with the recombinant proteins as antigens. Sera samples were previously rated as positive or negative using a secreted protein ELISA (Rebouças et al., 2013). There were nine, three and three false positives and one, two and three false negatives when rCP0126a, rCP0369 and rCP1957 were used as antigens in the immunoenzymic assay, respectively. There was no overlap of sera with false negative results for all three ELISAs or for any of two of them combined. Two samples were false positive for both rCP0126a and rCP0369 ELISAs; one sample presented false positive for both rCP0126 and rCP1957 ELISAs. No

---

**Fig. 3.** SDS-PAGE 12 % and Western blotting (WB) evaluation of rCP0126a, rCP0369 and rCP1957 proteins. (a) SDS-PAGE 12 % of recombinant proteins after purification. (1) rCP0126a, (2) rCP0369, (3) rCP1957. (b) WB of rCP0126a, rCP0369 and rCP1957 protein antigenicity using a pool of 10 positive sheep sera for CLA. Anti-sheep IgG-peroxidase conjugate was used as the secondary antibody. (1) rCP0126a, (2) rCP0369, (3) rCP1957 (4) rCP1957. (c) WB of rCP0126a, rCP0369 and rCP1957 protein antigenicity using a pool of 10 negative sheep sera for CLA. Anti-sheep IgG-peroxidase conjugate was used as the secondary antibody. (1) rCP0126a, (2) rCP0369, (3) rCP1957. (d) WB using monoclonal anti-6×His (diluted 1 : 5000). (1) Molecular weight marker prestained (Promega), (2) rCP0126a, (3) rCP0369, (4) rCP1957 recombinant proteins purified. The reactive band shows the identity of the recombinant proteins of approximately 25 kDa for CP0126a and 35 kDa for the CP0369 and CP1957 proteins.
samples presented false positive results for both rCP0369 and rCP1957 ELISAs or for all three assays.

The rCP0126a, rCP0369 and rCP1957 ELISAs showed 77.5 %, 92.5 % and 92.5 % specificity and 95 %, 90 % and 85 % sensitivity, with a calculated cut-off of 0.08, 0.15 and 0.085, respectively. The ELISA-rCP0369 and ELISA-rCP1957 presented a Kappa index of 0.814 and 0.850 and an AUC of 0.951 and 0.881, respectively, demonstrating the high concordance with the ELISA using secreted proteins and the CLA infectious status, as well as the high accuracy of both tests. A substantial agreement and a significant accuracy were obtained for the ELISA-rCP0126a because the Kappa index and AUC were 0.659 and 0.874, respectively (Fig. 5).

On the basis of the sensitivity and specificity of the ELISAs using the rCP0126a, rCP0369 and rCP1957 proteins, the positive predictive value varied from 30 %, 57.1 % and 66.6 % (for 10 % CLA prevalence) to 97.2 %, 99 % and 99.3 % (for 90 % CLA prevalence), respectively. The negative predictive value ranged from 98 % to 46.2 %, 98.8 % to 50.6 % and 98.8 % to 51.3 %, respectively, depending on the prevalence of the disease in a particular area (Fig. 6).

**DISCUSSION**

*C. pseudotuberculosis* is the aetiological agent of CLA, a disease that has caused serious economic losses in the sheep and goat farming sector around the world. In Brazil, according to MAPA (2014), infectious diseases such as CLA interfere with productivity and cause serious losses to producers.

---

**Fig. 4.** Distribution of the individual results obtained in the ELISAs using the recombinant proteins (rCP0126a, rCP0369 and rCP1957) as antigens for the detection of specific antibodies against *C. pseudotuberculosis*. Forty negative serum samples and 20 positive samples were employed in this survey. The graphs are organized by infection status (positive = 1, or negative = 0). (a) rCP0126a-ELISA; (b) rCP0369-ELISA; (c) rCP1957-ELISA. Samples were considered positive when the OD value was greater than or equal to each calculated cut-off, represented as a horizontal line in the graphs.

**Fig. 5.** ROC analysis of rCP0126a, rCP0369 and rCP1957 ELISAs. (a) ROC plot for rCP0126a-ELISA; AUC = 0.874 [0.763–0.945, 95 % confidence interval (CI)]. (b) ROC plot for rCP-0369-ELISA; AUC = 0.951 (0.861–0.989, 95 % CI). (c) ROC plot for rCP1957-ELISA; AUC = 0.881 (0.771–0.950, 95 % CI).
In this work we used three proteins, rCP0126a, rCP0369 and rCP1957 that were identified by MED analysis from more than 2097 targets. They encode a hypothetical protein, phosphoesterase PA-phosphatase related protein and cmtB trehalose corynomycolyl transferase B, respectively. These proteins are secreted or potentially exposed on the bacterial surface and for this reason are interesting targets to be used in immunodagnosis. These three proteins were expressed and purified from E. coli, evaluated in an indirect ELISA, and compared with the CLA infectious status of each animal, associated with the results of an ELISA using secreted C. pseudotuberculosis proteins (Rebouças et al., 2013). All of the proteins were recognized by sera from sheep presenting classical CLA clinical symptoms, associated with the bacterial isolation from caseous lesions and a positive result in a secreted antigen-ELISA.

The CP0126a-, CP0369- and CP1957-ELISAs showed 77.5 %, 92.5 % and 92.5 % specificity and 95 %, 90 % and 85 % sensitivity, respectively, when used for the serodiagnosis of CLA in sheep. A diagnostic assay based on ELISA has been developed using antigens obtained by C. pseudotuberculosis culture sonication, showing 71 % specificity and 83 % sensitivity (Binns et al., 2007). An increase in specificity is obtained using secreted antigens, with 93.5 % sensitivity and 98.5 % specificity (Seyffert et al., 2010), and using exotoxin, with 99 % specificity and 79.5 % sensitivity (Dercksen et al., 2000). Recombinant proteins have also been used in ELISA for CLA diagnosis. Two ELISA assays were compared (using whole cell antigens and rPLD) in the diagnosis of C. pseudotuberculosis-infected goats, presenting 81 % and 97 % sensitivity and 98 % and 99 % specificity, respectively (Sting et al., 2012). A comparison of ELISA using rPLD and IFN-γ assay detected goats experimentally infected with C. pseudotuberculosis with a reliability of 81.0 % and non-infected goats with a reliability of 97.0 % (Menzies et al., 2004). Others assays using a surface plasmon resonance-based biosensor assay for the detection of anti-PLD antibodies presented 86 % and 76 % sensitivity and specificity, respectively (Stapleton et al., 2009). In our study, the ELISA-CP1957 and ELISA-rCP0369 showed a great and satisfactory level of sensitivity and specificity. Interestingly, the rCP0126a protein that had the higher MED score had a low specificity, but a high sensitivity, when used in the ELISA.

Epidemiological studies have estimated that most Brazilian herds are infected and that clinical prevalence exceeds 30 % (Guimarães et al., 2011), but there are some regions with up to 80 % prevalence (Seyffert et al., 2010). Considering the 77.5 %, 92.5 % and 95 % specificity and 95 %, 90 % and 85 % sensitivity of the ELISAs, we obtained a positive predictive value of 65 %, 83.7 % and 88 %, respectively, and a negative predictive value of 98.6 %, 95.5 % and 95.6 %, respectively, for the rCP0126a, rCP0369 and rCP1957 ELISAs (with CLA prevalence approximately 30 %). Appropriate management of flocks requires a test able to identify true negatives and positives, indicating the removal of positive animals to prevent infection of the healthy animals. In a study performed in UK, the incidence of CLA was greatly reduced in infected sheep flocks using clinical examination and regular ELISA testing, where the owners were advised to remove from the flock any sheep that demonstrated clinical signs of CLA or tested positive for anti-PLD antibodies by ELISA or Western blot. This management enabled the disease to be controlled to such an extent that the appearance of new clinical cases of CLA was effectively halted (Baird & Malone, 2010).

The C. pseudotuberculosis targets identified were assessed by indirect ELISA, and two (rCP0369 and rCP1957) may be used for seroepidemiological CLA assessment in sheep flocks because they have excellent sensitivity and specificity levels. We are currently working on chimeric antigens using the targets studied herein, with the objective to enhance the assays’ sensitivity, but it must be noted that chimeric antigens can also interfere with the assay specificity, since neo-antigenic epitopes can be formed in the development of these chimeric antigens. We are also investigating other proteins as recombinant antigens for ELISA, assessing their applicability to increase the specificity.

---

**Fig. 6.** Positive (diamonds) and negative (squares) associated predictive values with ELISA-r0126a, ELISA-r0369 and ELISA-r1957 for varying prevalence level of CLA. Values were determined by ROC analysis based on sensitivity and specificity of the ELISAs.
ACKNOWLEDGEMENTS

This work had financial support from the Foundation for Research Support of the State of Rio Grande do Sul (FAPERGS) through project no. 11/1894-0. The authors declare that they have no conflict of interest.

REFERENCES


