Short Communication

Cloning and molecular characterization of \textit{Omp31} gene from \textit{Brucella melitensis} Rev 1 strain

Yousefi, S., Sekhavati *, M.H., Tahmoorespur, M., Abbassi-Daloii, T.

Department of Animal Sciences, Faculty of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran

Received 16 January 2014; accepted 14 September 2014
Corresponding Author: hadisekhavati@gmail.com

ABSTRACT
Brucellosis, caused by the genus \textit{Brucella bacterium}, is a well-known infection among domestic animals. Considering the serious economic and medical consequences of this infection, various preventive efforts have been made through using recombinant vaccines, based on outer membrane protein (OMP) antigens of \textit{Brucella} species. The objective of the present study was to clone, analyze the sequence, and predict the epitopes of \textit{Omp31} gene as a major \textit{B. melitensis} antigen. The full-length open reading frame (ORF) for this gene was amplified by specific primers and cloned into the pTZ57R/T vector. The gene sequence of \textit{B. melitensis} Rev 1 strain was submitted to NCBI database. The results of phylogenetic analysis showed that \textit{Omp31} is almost similar in different \textit{Brucella} species. Online prediction software programs were also used to predict B- and T-cell epitopes, secondary and tertiary structures, antigenicity, and enzymatic degradation sites. The bioinformatic tools in the current study were confirmed by the results of three different experimental epitope prediction studies. Bioinformatic analysis identified one T-cell and three B-cell epitopes for Omp31 antigen. Finally, based on the antigenicity and proteosome recognition sites, common B- and T-cell epitopes were predicted for \textit{Omp31} (amino acids 191-204). Bioinformatic analysis showed that these regions had proper epitope characterization and could be useful for recombinant vaccine development.

Keywords: \textit{Brucella melitensis}, \textit{Omp31}, Bioinformatic analysis

INTRODUCTION

Brucellosis, which is regarded as a common zoonotic disease and a public health issue, has economic consequences for many developing countries (Karthik et al., 2013). This infection caused by \textit{Brucella bacterium}, a genus of Gram-negative bacteria, can primarily affect domestic animals (Cutler et al., 2005).

In animals, brucellosis is characterized by abortion and reduced fertility, while in humans, it manifests with chronic infections and symptoms such as undulant fever, arthritis, and osteomyelitis (Pappas et al., 2006). The outer membrane proteins (OMPs) of \textit{Brucella bacterium} are cell-specific surface antigens, which have extremely remarkable immunogenic characteristics. MPs are excellent candidates for the development of
brucellosis recombinant vaccines. These cell surface antigens are classified into two major groups: 1) OMP2a and OMP2b, and 2) OMP25 and OMP31 (Gupta et al., 2012). The open reading frame (ORF) of OMP31 gene was initially cloned from B. melitensis 16M strain, and its predicted amino acid sequence exhibited significant homology to OMP25 gene (34% sequence identity) (Vizcaino et al., 1996). In addition, according to previous research, Omp31 could be used as an experimental gene in protection assays to determine its potential as a vaccine candidate (Vahedi et al., 2011; Azimi et al., 2012; Ghasemi et al., 2013). The immune system in the human body produces antibodies which specifically attach to identified regions of antigens, known as epitopes (Berzofsky, 1985). In general, epitopes may be classified as B-cell (continuous and discontinuous) and T-cell (MHC-I and MHC-II) epitopes (Zhang et al., 2012). Continuous or linear epitopes are composed of consecutive amino acids, whereas discontinuous epitopes constitute spatially folded amino acids which lie far away in the primary sequence (Ponomarenko and van Regenmortel, 2009). T-cell epitopes are antigenic peptide strings, recognized by T-cell receptors (Chen et al., 2011). B- and T-cell epitopes, which could be predicted via computational tools, are widely applied in antibody production, immunodiagnostics, development of epitope-based vaccines, selective deimmunization of therapeutic proteins, and autoimmunity (Steere et al., 2011). These cost-effective, advantageous, and feasible tools could also replace experimental methods which are costly and time-consuming (Ponomarenko and van Regenmortel, 2009). Today, several epitope prediction software programs are available. The first generation of these prediction tools was supported by motif-based algorithms (Chen et al., 2011), primary amino acid sequence of antigens (Hopp and Woods, 1981), or other physiochemical protein characteristics. Recently, more sophisticated methods have been developed, using various machine learning algorithms, based on support vector machines (Donnes and Elofsson, 2002), hidden Markov model (Noguchi et al., 2002), and artificial neural networks (Buus et al., 2003). The objective of the present study was to clone, analyze the sequence, and predict the epitopes of Omp31 as a B. melitensis antigen. Finally, B- and T-cell epitopes were used for designing an epitope-based vaccine.

MATERIALS AND METHODS

Bacterial strains, growth conditions, and isolation. In the current study, B. melitensis Rev 1 strain was obtained from the B. melitensis culture collection (Razi Institute, Mashhad, Iran) and cultured, as described in the literature (Delpino et al., 2007). DNA was extracted, using a DNA extraction kit (Bioneer, Korea). The quality and purity of the extracted DNA were analyzed by agarose gel electrophoresis and a NanoDrop spectrophotometer (Sigma, USA). Escherichia coli strain DH5α was used as the host for cloning, sequencing, and maintaining different DNA fragments. T/A cloning vector pTZ57R/T (Thermo, USA) was used for cloning and sequencing the amplified gene.

PCR amplification. Genomic DNA of B. melitensis Rev 1 was used as the template for amplifying the full-length ORF of Omp31 gene (723 bp), using EX Taq DNA polymerase (Takara, Japan). Specific primers with restriction sites at the 5’ end (underlined) were designed, using Primer Premier 5 (Premier Biosoft International, USA), according to the available nucleotide sequences on NCBI GenBank database (Table 1). Polymerase chain reaction (PCR) was performed in a reaction mixture, containing 2.5 μl of 10X PCR buffer, 2 μl of MgCl2, 2 μl of dNTPs, 0.5 μl of DNA solution (50-100 ng/μl), 1.5 μl of mix primer (5 pmol/μl), and 0.125 U/μl of EX Taq DNA polymerase; also, deionized water was added to reach a final volume of 25 μl. The PCR program was performed with an initial denaturation at 94 °C for 6 min, followed by 30 cycles of denaturation at 94 °C for 30 sec, annealing at 58 °C for 30 sec, extension at 72 °C for 45 sec, and a final extension at 72°C for 10 min.
Table 1. Specific primers with restriction sites

<table>
<thead>
<tr>
<th>Gene</th>
<th>Primer sequences (5’ 3’)</th>
<th>Restriction enzyme</th>
<th>Length (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>F:</td>
<td>GAAATTCATGAAATCCGTAATTGCG</td>
<td>EcoRI</td>
<td>723</td>
</tr>
<tr>
<td>Omp31</td>
<td>GGATCCTTAGAACTTGAGAGCATGACCCG</td>
<td>BamHI</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Validation of bioinformatic software programs used in the present study

<table>
<thead>
<tr>
<th>Antigens</th>
<th>Predicted epitopes</th>
<th>Experimental epitopes</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dnak2 for Brucella species</td>
<td>40-67,78-92,210-227,357-370,523-537,609-640</td>
<td>617-637</td>
<td>Vizcaino et al., 1997</td>
</tr>
<tr>
<td>SOD3 for Brucella species</td>
<td>44-50,70-86,134-153,147-165</td>
<td>75-86,136-150,149-62</td>
<td>Tabatabai et al., 1994</td>
</tr>
</tbody>
</table>

Table 3. Final B- and T-cell epitopes predicted in this study

<table>
<thead>
<tr>
<th>Gene</th>
<th>No.</th>
<th>B-cell epitopes</th>
<th>Final B-cell epitopes</th>
<th>Final T-cell epitopes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Omp31</td>
<td>1</td>
<td>23 VVSEPSAPTAAPVDTF 38</td>
<td>193 YAINNNWTLELY</td>
<td>207</td>
</tr>
<tr>
<td>2</td>
<td>51 YAGGKFHKFSSHFDKEDNEQVSG 73</td>
<td>91 NWQLDNGVVL</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>168 GDDASALHTWSDKTAKGWTLS 188</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>191 AYEYAINNNWTLELY 204</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4. Protein digestion analysis of final B- and T-cell epitopes

<table>
<thead>
<tr>
<th>Gene</th>
<th>No.</th>
<th>B-cell epitopes</th>
<th>Mass (Da)</th>
<th>pI</th>
<th>Undigested enzymes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Omp31</td>
<td>1</td>
<td>YAGGKFHKFSSHFDKEDNEQVS G</td>
<td>2574.75</td>
<td>5.5</td>
<td>Clostripain, cyanogen bromide, iodojbenzoate, and trypsin R</td>
</tr>
<tr>
<td>2</td>
<td>GDDASALHTWSDKTAKGWTLS</td>
<td>2247.78</td>
<td>5.3</td>
<td>Trypsin R, proline endopeptidase, cyanogen bromide, AspN, clostripain, and staphylococcal protease</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>AYEYAINNNWTLELY T-cell epitopes</td>
<td>165.78</td>
<td>4.8</td>
<td>Trypsin R, proline endopeptidase, cyanogen bromide, AspN, and clostripain</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mass(Da)</td>
<td>pI</td>
<td>Undigested enzyme</td>
<td></td>
</tr>
<tr>
<td>Omp31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5. Antigenicity of the predicted epitopes

<table>
<thead>
<tr>
<th>Gene</th>
<th>No.</th>
<th>Final B-cell epitopes</th>
<th>VaxiJen scores</th>
<th>Final T-cell epitopes</th>
<th>VaxiJen scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Omp31</td>
<td>1</td>
<td>23 VVSEPSAPTAAPVDTF 38</td>
<td>0.2*</td>
<td>193 YAINNNWTLELY 207</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>51 YAGGKFHKFSSHFDKEDNEQVSG 73</td>
<td>0.7</td>
<td>91 NWQLDNGVVL</td>
<td>100</td>
<td>0.08*</td>
</tr>
<tr>
<td>3</td>
<td>168 GDDASALHTWSDKTAKGWTLS 188</td>
<td>0.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>191 AYEYAINNNWTLELY 204</td>
<td>1.2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Probable Non-Antigen
Cloning. The PCR product was purified from the agarose gel by Ron's Agarose Gel Miniprep Kit (BioRon, Germany), according to the manufacturer’s instructions. Based on the manufacturer’s instructions, the purified PCR product was ligated into pTZ57R/T cloning vector by T/A cloning strategy. Preparation of competent cells and transformation were performed as described in the literature (Sambrook and Russell, 2001). The recombinant vectors were transformed into competent E. coli DH5α. The recombinant clone(s),
harboring plasmid DNA with inserts, were screened, based on their ampicillin resistance. The fidelity of *E. coli* DH5α transformants was verified by PCR reaction, using M13 universal primers. The recombinant plasmids were purified, using Ron's Plasmid Mini Kit and confirmed by restriction enzyme digestion. Also, purified plasmids were subjected to sequencing at Bioneer Corporation (South Korea). The obtained sequences were studied through homology analysis and were aligned with other reported *Omp31* gene, using BLAST and CLC Main Workbench 5.5 (CLC Bio, USA), respectively.

**Prediction of secondary and tertiary structures.** The secondary structure was predicted, using the improved self-optimized prediction method with alignment (SOPMA) software (available on https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=NPSA/ npsa_sopma.html) (Geourjon and Deleage, 1995) with four conformational states (i.e., helix, sheet, turn, and coil) of candidate genes. In addition, the tertiary structure was analyzed by 3D Ligand Site, which is an online ligand-binding site prediction server (available on http://www.sbg.bio.ic.ac.uk) (Wass and Sternberg, 2009).

**Servers and software programs for B- and T-cell epitope prediction.** B- and T-cell epitopes of candidate genes were predicted, using different servers and software programs. ABCPred, BepiPred, BCpred, SVMTrip, and LEPS were used for B-cell prediction, while IEDB, SYFPEITHI, ProPred I, and ProPred were employed for T-cell prediction.

**Verification of the bioinformatic analysis approach.** In order to validate the predicted outputs, the results of three experimental epitope prediction studies were evaluated by bioinformatic tools, applied in the present study.

**Characterization of epitopes.** Final B- and T-cell epitope predictions were evaluated, using VaxiJen 2.0 server (available on http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html) for the prediction of protective antigens. In addition, enzymatic degradation sites, mass (Da), and Isoelectric point (pI) were determined, using the protein digest server (available on http://db.Systemsbiology.net: 8080/ proteomics Toolkit / protein Digest .html).

**RESULTS AND DISCUSSION**

**PCR amplification, cloning, and nucleotide sequencing analysis.** In this study, *OMP31* gene from *B. melitensis* Rev 1 strain was amplified. The accuracy of the fragment was visualized on 0.8% agarose gel electrophoresis (Figure 1). The PCR products were successfully ligated into pTZ57R/T cloning vector and transformed into competent *E. coli* DH5α cells. After the selection of positive screen colonies by colony PCR, the integrity of recombinant plasmids was confirmed through restriction digestion. The sequencing of recombinant plasmids was performed with specific primers and universal M13 primer. The sequence was submitted to NCBI database under the accession number, KJ193851. The obtained sequence was analyzed, using BLAST and CLC Main Workbench 5.5. Based on the findings, *Omp31* gene sequence had 100% homology to *B. melitensis* species (*B. melitensis* NI, *B. melitensis* M28, *B. melitensis* 183, *B. melitensis* 16M, and *B. melitensis* ATCC), while exhibiting less similarity to *B. ovis* (Figure 2). In addition, the phylogenic tree was drawn for confirming the results of genetic distance matrix. It was revealed that *B. melitensis* Rev 1 strain had a close homology to *B. melitensis* NI, *B. melitensis* M28, *B. melitensis* 183, *B. melitensis* 16M, and *B. melitensis* ATCC, as similarly observed in the pairwise comparison matrix (Figure 3).

**Prediction of the secondary structure.** We predicted the secondary structure of *Omp31*, using SOPMA server software. The results revealed that random coils, β turns, α helices, and extended strands (β folds) accounted for 59.17%, 2.5%, 15.42%, and 22.92% of the structure, respectively. The greater proportion of the extended strands and random coils in the secondary structure of the protein was associated
with the increased likelihood of antigenic epitope formation.

Verification of the bioinformatic analysis approach. In order to validate all software programs used in the present study, three antigens, whose epitopes were experimentally determined (available on http://www.iedb.org), were selected and their epitopes were predicted, using bioinformatic tools. The predicted epitopes were compared with the results of experimental studies. The findings revealed that our in silico predicted epitopes were similar to the findings of experimental studies for all the selected antigens (Table 2).

Prediction of B- and T-cell epitopes. B-cell epitopes, as well as MHC-I (A-0101, A0201, and B-2705) and MHC-II (DRB1-0101 and DRB1-0401) classes of T-cell epitopes, were predicted, using different online software programs. For each program, epitopes with the highest score were selected as proper epitopes. Finally, four B-cell and two T-cell epitopes were selected with respect to the most conserved sequences in all the proposed epitopes (Table 3).

Characterization of epitopes. The results of analysis by the protein digest server for mass (Da), pI, and enzymatic degradation sites are presented in Table 4. The findings indicated that the majority of predicted epitopes lacked proteasome recognition sites. The candidate proteins with a score of 0.64 were identified as antigens by VaxiJen 2.0 server (threshold: 0.5). The antigenicity of the final predicted B- and T-cell epitopes is shown in Table 5. Furthermore, the results of VaxiJen 2.0 analysis indicated that five predicted epitopes showed antigenicity. Finally, the 3D structure of candidate epitopes was illustrated, using 3D Ligand Site server (Figure 4). The 3D structure analysis showed that all the predicted B- and T-cell epitopes were located outside the candidate antigen.

Several studies have predicted epitopes of desired antigens via computational approaches and have applied the findings in experimental studies in order to design epitope-based vaccines (Simon et al., 2010). In the present study, the dominant B. melitensis Rev 1 antigen was a candidate for cloning, molecular analysis, and epitope prediction. The results showed that the candidate gene could be successfully cloned. Moreover, molecular analysis revealed that Omp31 sequence of B. melitensis Rev 1 is nearly similar to other Brucella species. The strong nucleotide identity of Omp31 might be attributed to the high degree of genetic relatedness among Brucella species (Rajagunalan et al., 2013). Comprehensive bioinformatic analyses were performed on the candidate antigen by online B- and T-cell epitopic prediction servers. To confirm the results of our bioinformatic approaches, we analyzed three different antigens and compared the obtained computational outputs with the experimental results. The bioinformatic analysis for GroEL, Dnak, and SOD antigens was successfully confirmed, using experimentally achieved epitopes. In the present study, the final epitopic prediction results proposed five Omp31 epitopes, which could be used as immune-dominant epitopes for the development of recombinant subunit vaccine against brucellosis. In this regard, Wang et al. (2014) in a recent epitope mapping experiment on Omp31 antigen (as a B. melitensis OMP) found that amino acid residue 48-74 was the dominant epitope; this region was also predicted in the present study via bioinformatic analysis. Moreover, similar results have been reported for Omp31 epitopes by Vizcaino et al. (1997)and Cassataro et al. (2005). Several studies have attempted to predict epitopes of the desired antigen by computational approaches and have used these findings to design epitopic-based vaccines (Li et al., 2013; Sekhavati et al., 2015; Yousefi et al., 2015). The results of secondary and tertiary structure analyses showed that the common predicted B- and T-cell epitopes were located in the random coil regions on the surface structure of candidate antigens. Random coil regions are located on the surface of the protein, where it is necessary for the surface structure to properly bind to ligands, with a high possibility of epitope formation (Li et al., 2013). To prevent degradation of peptides during antigen processing, epitopes should lack proteasome
recognition sites (Toes et al., 2001). In the present study, the predicted B- and T-cell epitopes were analyzed in terms of enzymatic degradation. The proposed epitopes lacked proteasome digestion sites for several dominant enzymes.

Although brucellosis is a common zoonotic disease among domestic animals, but up to now no recombinant vaccine has been developed. Therefore, the aim of the present study was to clone, analyze the sequence, and predict the epitopes of two candidate \textit{B. melitentis} Omp31 antigens. Additionally, we aimed to confirm the predicted epitopes, using the results of experimental epitope prediction studies in order to design a suitable recombinant vaccine. Phylogenetic analysis showed that this gene was nearly similar in different \textit{Brucella} species, and common B- and T-cell epitopes were predicted for \textit{Omp31} (amino acids 191-204). \textit{In vitro} synthesis of the determined peptides and experimental validation are required for using the predicted epitopes as effective vaccines against \textit{Brucella melitensis} pathogen. It is worth mentioning that our laboratory (Animal biotechnology lab of Animal science group of Ferdowsi university of Mashhad) has initiated research on this subject.

\section*{Ethics}

I hereby declare all ethical standards have been respected in preparation of the submitted article.

\section*{Conflict of Interest}

The authors declare that they have no conflict of interest.

\section*{Grant Support}

The present study was funded by INFS (infs.gov.ir) of I.R.I with grant No. 92041813.

\section*{Acknowledgments}

The authors would like to express their gratitude to INFS for their full support.

\section*{References}


