Development and immunoreactivity evaluation of a chimeric recombinant protein encoding \textit{Brucella} antigen: \textit{In silico} to \textit{in vitro}

Abbas Abdollahi$^{1,2}$, Shahla Mansouri$^{1*}$, Jafar Amani$^2$, Mahdi Fasihi-Ramandi$^3$ and Mohammad Moradi$^1$

$^1$Department of Microbiology, Kerman University of Medical Sciences, Kerman 93630, Iran
$^2$NCD Research Center, Department of Microbiology, Fasa University of Medical Sciences, Fasa, Fars 74616-8-6688, Iran
$^3$Applied Microbiology Research Center and $^4$Molecular Biology Research Center, Baqiyatallah Medical Science University, Tehran 19166, Iran

Received 13 July 2015; revised 22 November 2015; accepted 5 December 2015

Brucellosis is an important health problem in developing countries and no vaccine is available for the prevention of infection in humans. Because of clinically infectious disease and its economic consequences in human and animals, designing a proper vaccine against \textit{Brucella} is desirable. In the present study, we evaluated the immune responses induced by a designed recombinant chimera protein and investigated the immunogenic potential of some immune reactive antigens of \textit{Brucella}. Three immune dominant antigens of \textit{Brucella} including trigger factor (TF), Omp31 and Bp26 (have been characterized as potential immunogenic and protective antigens) were fused together by EAAAK linkers to produce a chimera. Recombinant chimeric protein was synthesized, cloned and expressed in \textit{Escherichia coli} BL21 (structure were designed \textit{in silico}). The purification of recombinant protein was performed by using Ni-NTA agarose, and anti-His antibody was used for confirmation (Western blot). The recombinant chimeric protein could be a new potential antigen candidate for the development of a subunit vaccine against \textit{Brucella}. These results demonstrate the role of the Bioinformatics in vaccine design, assisted by experimental procedures.

\textbf{Keywords:} \textit{Brucella}, bioinformatics, immunity, vaccine

\textbf{Introduction}

Brucellosis is caused by \textit{B. reculla}, a Gram negative, facultative, intracellular, partially acid fast coccobacillus, lacking capsule or flagella. Brucellosis is among the most common zoonosis infecting approx half a million people annually around the world. The disease is endemic in many developing countries, especially in Latin America, Middle East, Africa, Asia and the Mediterranean Basin$^{1,2}$. Brucellosis leads to clinically infectious diseases and economic consequences in humans and many domestic animals are huge. Infection causes abortion and infertility in the animals, and undulant fever in humans (brucellosis). At present, there is no commercially available vaccine against human brucellosis and the disease is prevented by immunization of uninfected animals and elimination of the infected ones$^{3,4}$. In animals, immunization against \textit{Brucella} infections is usually performed by administration of the live attenuated smooth \textit{Brucella} strains: \textit{B. abortus} strain S19, \textit{B. melitensis} strain Rev.1 and non-smooth strain \textit{B. abortus} RB51. \textit{B. abortus} S19 is proven to be effective against \textit{B. abortus} infection in cattle and \textit{B. melitensis} Rev.1 is effective against \textit{B. melitensis} and \textit{B. ovis} infection in sheep and goats. Despite the efficacy of vaccination, these vaccines have some disadvantages, such as, the ability to cause disease in humans and abortion in pregnant animals$^{5-7}$. However, to develop a human \textit{Brucella} vaccine, those \textit{Brucella} proteins that exist in \textit{Brucella} strains and are pathogenic to humans would be ideal for vaccine development. Because of the problems derived from the utilization of attenuated and killed vaccines in humans and animals, the development of a beneficial subunit vaccine against brucellosis is desirable$^{6-8}$.

In \textit{Brucella} spp., only few antigenic components of intracellular and cell surface components have recently been considered as protective antigens, and have suitable immunogenic activity, for example, \textit{Brucella} lumazine synthase, BLS (cytoplasm); ribosomal protein L7/L12 (cytoplasm); sugar-binding 39-kDa protein, p39 (periplasm); Bp26 periplasmic immunogenic protein, Bp26 (periplasm); molecular chaperone, DnaK (cytoplasm); outer membrane
protein, Omp16,19,25,31 (outer membrane); Cu/Zn superoxide dismutase, SodC (periplasm); SurA Peptidyl-prolylcis-trans isomerase, SurA (periplasm); and Trigger factor, Tig/TF (cytoplasm). Despite the immunogenicity of these antigens, the desirable protection against bacteria could be improved using a multiple subunit vaccine. Omp31, TF and Bp26 have been characterized as potential immunogenic and protective antigens and have been previously studied in whole and portion form to determine their protective immunogenicity.\(^9,10\)

In the present study, we have developed a new structural model containing three putative antigenic determinants of Brucella, Omp31, TF and Bp26 (in silico designed structure) and evaluated the in vitro experiments to cloning and expression of this structure prediction.

Materials and Methods

We selected three antigenic determinants\(^11-19\), TF, Bp26 and Omp31 with 485 aa, 25 (87-111aa residues) and 27 (48-74aa residues), respectively and they were fused together by hydrophobic rigid linkers (EAAAK). Restriction enzyme (RE) sites were added at 5' and 3' ends.\(^20\) Codon of this chimera was optimized to the best efficiency of expression in Escherichia coli.

In Silico Prediction

Databank Collection, Antigen Designing and Physicochemical Parameters

The identification and analysis of gene sequences and information gathering was carried out by searching the literature from the NCBI PubMed (http://www.ncbi.nlm.nih.gov/pubmed) and UniProt (http://www.uniprot.org) databases. The alignment and sequences identity of component to identify a conserved region in all the required sequences were performed using BLAST (http://www.uniprot.org/blast/) and ClustalW (http://www.ebi.ac.uk/Tools/clustalw2) softwares. With aim of best optimization of the cloning and expression designed gene in E. coli, amino acids codons were predicted by SwissProt reverse translation (http://www.bioinformatics.org/sms2/rev_trans.html) and Codon Optimization online service (https://eu.idtdna.com/CodonOpt). DNA/RNA GC Content Calculator (http://www.endmemo.com/bio/gc.php) were used to calculate G/C% before and after optimization. Antigenicity, linear epitope, \(\beta\)-turn, surface accessibility and flexibility of chimeric designed antigen were predicted by IEDB Analysis Resource (http://tools.immuneepitope.org/tools/), Vaccine Design server (http://www.violinet.org/vaxign/index.php) and the Proteome Binders Epitope Choice Resource (http://bioware.ucd.ie/epic/). Chimeric antigen physicochemical parameters: total number of residues, solvent accessibility, aliphatic index, theoretical isoelectric point (pI), extinction coefficient, half-life, molecular weight, grand average hydropathy and instability index, were computed using Expasy’s Protparam (http://us.expasy.org/tools/protparam.html), Protein Calculator v3.4 (http://protcalc.sourceforge.net/) and Recombinant Protein Solubility Prediction (http://www.biotech.ou.edu/). Protein solubility of different residues was predicted by DSSP (http://www.cmbi.ru.nl/dssp.html) and VADAR (http://vadar.wishartlab.com).

Antigenic and Allergenic Epitopes Prediction

T-cell epitopes prediction parameters, including binding sites of both MHC class I and MHC class II, were analyzed by GPS-MBA Prediction of MHC-binding system Version 1.0 online software (http://mba.biocuckoo.org/links.php) and Immune Epitope Database-IEDB-Analysis Resource online software (http://tools.immuneepitope.org/). Chimeric antigens were analyzed for continuous B-cell epitopes using Bcepred (http://www.imtech.res.in/s of -raghava/bcepred/). The discontinuous B-cell epitopes were predicted with DiscoTope server (http://www.cbs.dtu.dk/services/DiscoTope/). Conformational B-cell epitope was predicted with web server CUTOPE (http://www.imtech.res.in/raghava/cbtope/) and the PSIPRED Protein Sequence Analysis Workbench (http://bioinf.cs.ucl.ac.uk/psipred/). Presence of possible allergenic sites, based on the similarity of known epitopes with any region of antigen, was predicted using AlgPred (http://www.imtech.res.in/raghava/algpred/) and SDAP-Structural Database of Allergenic Proteins (https://fermi.utmb.edu/).

RNA Secondary Structure

Analysis of the secondary structure of messenger RNA of the designed chimera was predicted using online software, including Themfold Web Server (http://mfold.rit.albany.edu/?q=mfold/RNA-Folding), RNAfold Web Server (http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi) and RNA Structure Web Servers for RNA Secondary Structure Prediction (http://rna.urmc.rochester.edu/RNAstructureWeb/).
**Secondary and Tertiary Structure Protein**

Secondary structure prediction to show and analyze the chimeric recombinant protein structure, and also computing of protein sequence analysis and functional parameters of protein, such as, molecular weight, protein half-life, number of β-turns and random coils, were performed with GOR IV secondary structure prediction method (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_gor4.html) and Predict Protein server (https://www.predictprotein.org/). Tertiary structure -3D and stability prediction of protein in different situation were performed by DeepView-Swiss-PdbViewer (http://spdbv.vital-it.ch/). 3D structure prediction to show the final construction was simulated and modeled by using RasMol-Molecular Graphics Visualisation Tool (http://rasmol.org/). Recombinant chimeric protein modeling to predict binding sites and 3D structure was performed using I-TASSER server (http://zhanglab.ccmb.med.umich.edu/I-TASSER). The C-score (confidence score for estimating the quality of designed model and structure) and TM-score (measuring scale of the structural similarity with original structure and other similar proteins) were given in the I-TASSER result page. Tertiary structure to recognize faults in the generated models, energy plot, Z-score (overall model quality) and 3D structures were validated by using Ramachandran plot (Z-score) in PROCHECK (http://www.ebi.ac.uk/thornton-srv/software/PROCHECK/).

**In Vitro Experiments and Optimization**

After *in silico* design and prediction, the chimeric gene was synthesized (Biomatik, Canada) and subsequently cloned into pET-28a (+) to construct pET-chimeric protein (pET-CP) plasmids (Novagen, Madison, Wisconsin, USA).

**Gene Expression and Purification of Recombinant Protein**

The recombinant gene was synthesized and subsequently cloned (Biomatik, Ontario, Canada) into pET-28a (+) to construct pET-recombinant protein (pET-rTF/Bp26/Omp31); restriction enzyme sites were added at 5’ and 3’ ends. The pET-rTF/Bp26/Omp31 was transformed into *E. coli* BL21 (DE3) strain (Novagen-Merck KGaA, Darmstadt, Germany). The transformed clones were inoculated into Luria Bertani (LB) broth medium (Merck KGaA, Darmstadt, Germany), containing 50 µg/mL kanamycin (Sigma-Aldrich, Taufkirchen, Germany). The incubation was continued by agitation (300 rpm) to 0.5 OD values at 600 nm isopropyl-β-D-thiogalactopyranoside (IPTG) (Sigma-Aldrich, Taufkirchen, Germany) was added to induce the gene expression (IPTG concentration 700 µg/mL, 4 h).

The culture was harvested by centrifugation at (10000× g, 10 min, 4°C), then re-suspended in lysis buffer (8 M urea, 0.1 M NaH₂PO₄, and 0.01 M Tris, pH8.0) containing protease inhibitors (Sigma-Aldrich, Germany). Recombinant chimeric protein was purified using Nickel-nitrilotriacetic acid (Ni-NTA) resin (Qiagen, Manchester, UK). Proteins were eluted in 1mL buffer containing 200 mM imidazole (Sigma-Aldrich, Germany) and 500 µL MES buffer (20 mM) (Sigma-Aldrich, Germany). The purified protein was monitored by SDS-PAGE (Bio-Rad, USA) and its concentration was estimated by using Nanodrop-Biowave II analyzer (Biochrom, UK) and Bradford protein method measurement. The protein elution was dialyzed against 0.1 M phosphate buffered saline (PBS, pH 7.4) for 72 h in cold room to remove urea and then stored at -70°C for future use.

Molecular weight (MW) of protein was determined by electrophoresis with using pre-stained protein ladder marker (SM0671), with 10 bands (approx. 10, 15, 25, 35, 40, 55, 70, 100, 130, 170 kDa) (Fermentas, USA). To evaluate of accuracy of *in silico* data, protein solubility were performed by culturing in two different temperatures 37°C and RT, with above procedures.

**Western Blot Analysis**

SDS-PAGE protein bands were transferred into nitrocellulose membrane (Sigma-Aldrich, Germany). The membrane was then blocked in 3% bovine serum albumin (BSA), phosphate buffered saline, overnight at 4°C. The membrane was then washed three times in phosphate buffer saline, Tween-20 (PBST) and was incubated with anti His-Tag antibody (Sigma-Aldrich, Germany), 1 h at 37°C. Membrane was then washed three times with PBST and incubated with anti-mouse IgG-peroxidase (Sigma-Aldrich, Germany) for 1 h at 37°C. The membrane was then washed three times and developed in diaminobenzidine (DAB) solution (Sigma-Aldrich, Germany). By visualizing the protein band, the reaction was stopped adding distilled water. This procedure was repeated to patient pulled serum (with brucellosis) to confirm the reactivity of antibody with recombinant protein.
Circular Dichroism (CD) Analysis
To evaluate and determine the secondary structures of predicted (in silico) recombinant protein rTF/Bp26/Omp31 (0.25 mg/mL in PBS) CD was recorded on a JASCO J-810 spectrometer (Jasco, Inc., Easton, MD, USA).

Results
Bioinformatics Analysis
BLAST and alignment sequence comparison illustrated the highly conserved sequences among chimera amino acid sequences and strains of Brucella spp. Final construction of chimera, 1-485aa (TF), 486-495aa (EAAKEAAAK-Linker), 496-520aa (Bp26), 521-525aa (EAAAK-Linker) and 526-552aa (Omp31), was made by fusing the C terminal of TF, middle portion of Bp26 87-111 and N terminal of Omp31 48-74, and use of two hydrophobic-rigid amino acid linkers.

Gene optimization to expression in E. coli was improved by changing the GC count from 51% (in native form) to 55% in reforming nucleotide to get the best expression in E. coli as per the results of data bases analysis. Prediction of antigenicity and linear epitope of antigen showed the antigenic determinant and epitopes in several different sequences in chimeric antigen. MHC I and II classes binding sites were determined in several positions in the protein structure as described previously. According to databank analysis, based on the similarity of known epitopes with any region of antigen, no possible allergenic sites were present. Prediction of physicochemical parameters of protein was computed (approx): mol wt, ~65 kDa; number of amino acids, 552; theoretical pI, ~5.0, extinction coefficient, ~24780 M⁻¹ cm⁻¹; estimated half-life, >10 h (E. coli, in vivo). N-terminal of the sequence considered, M (Met); instability index, ~45.00 (regarding that chimeric protein was stable); aliphatic index, ~ 80.00 (a positive factor for the increase of thermostability); grand average of hydropathicity-GRAVY (sum of hydropathy values), ~-0.680; and ~100% chance of solubility when over expressed in E. coli.

RNA secondary structure analysis indicated no disorder in mRNA conformational structure and normal formation of folding. Optimal secondary structure with a minimum free energy of ~-480 kcal/mol prepared a suitable ΔG in nucleotides of mRNA and it had no hairpin or pseudo knot in the first nucleotides (Fig. 1).

Protein secondary structure analysis showed 58.70, 7.07 and 34.24% of protein sequences were α-helix, extended strand and random coil, respectively. As we expected, two helixes were present in positions 485-495 and 520-525 that correlated with the position of linkers. No signal peptide cleavage site was present in protein sequence as described previously. Tertiary structure of the protein showed a construction with three determined domains (Fig. 2), which linked together with two linkers. Comparison of chimera protein with native domain structures illustrated that the chimera protein had acceptable stability.

Fig. 1 — Prediction and analysis of mRNA secondary structure. [Analysis was performed with Secondary Structure Prediction software (http://rna.urmc.rochester.edu/RNAstructureWeb/). RNA secondary structure analysis indicated no disorder in mRNA conformational structure and normal formation of folding; mRNA with a minimum free energy of ~-480 kcal/mol; no hairpin or pseudo knot in first nucleotides; provide a suitable structure (demonstrated with red line)].
(~−14000 Kcal/mol). This data was confirmed by Ramachandran plot.

**Expression and Purification of Recombinant Protein**

Expression condition was optimized by using gradient change in temperature, IPTG concentration and growth time. Transformed pET over expressed into *E. coli* BL21 (DE3) strain in the presence of 700 µg/mL IPTG at 37°C after 4 h of growth (Fig. 3). Solubility of protein was adjusted by culturing the bacterium at two different temperatures (37°C & RT), and then running the supernatant and precipitant of bacteria lysate by SDS-PAGE to demonstrate the existence of protein (Fig. 4). The purified protein was absolutely soluble and illustrated in the supernatant of bacteria lysate gel electrophoresis of both temperatures. The result of *in vitro* solubility was similar to *in silico* prediction according to 100% solubility prediction of recombinant protein. Recombinant chimeric protein elution was purified by using Ni-NTA resin. The purified protein was monitored by SDS-PAGE and its concentration was estimated. The concentration of eluted protein after dialysis against 0.1 M phosphate buffered saline (PBS, pH=7.4) for 72 h in cold room was 700 µg/mL. Mol wt of protein was ~70 kDa, which was higher compared to the predicted mol wt (~65kDa) because of addition the amino acid residues of pET28a (~5kDa) (Figs 4 & 5).

**Western Blot Analysis**

The results of Western blot assay using anti-His Tag antibody confirmed that major band observed in SDS-PAGE (~70 kDa) was the recombinant protein (Fig. 5).
Fig. 5 — Western blotting results of the recombinant protein expression and purification. [Recombinant protein reactivity by Western blotting using anti-His Tag antibody and pulled sera from infected patients for identification. The purified protein was monitored by SDS-PAGE and its concentration was estimated; purification of protein was performed by using Ni-NTA resin. Lane 1, Protein size marker; Lane 2, Anti His-Tag antibody; & Lane 3, Pulled sera from infected patients.]

Circular Dichroism (CD)

CD analysis of protein showed that the physicochemical parameters, α-helix (55.8%), extended strand (9.9%) and random coil (34.3% of protein), were approx similar to their predicted value.

Discussion

Safe elimination of infectious diseases by using recombinant subunit vaccines, which are well defined, avirulent, noninfectious and nonviable, has an important advantage in comparison to live vaccines because of the hazards of remnant virulence and infectious potential of viable microorganisms. Concurrence of results obtained through bioinformatics approach in laboratory experiments demonstrates that in silico analysis can be utilized for vaccine design in a safe manner, in contrast with live vaccines. In Brucella infection, the ability of induction of both B/T-cell responses is important in a new vaccine candidate. Therefore, mapping of B/T-cell antigenic determinants by in silico approaches is an important method for designing a successful vaccine.

There are several components in Brucella with immune stimulation activity. Among these antigens, immune response of B/T-cells to TF (acts as a chaperone by maintaining the newly synthesized protein in an open conformation), Bp26 (26 kDa periplasmic immunogenic protein) and Omp31 (major outer membrane protein associated with peptidoglycans) immune-determinant epitopes has been described by other researchers. Although the immunity validation of TF, Omp31 and Bp26 has been well defined, the immunogenicity property of these three components in combination was not studied previously. Because of the known proper immune stimulation of these antigenic determinants individually, we constructed a recombinant subunit chimera to increase the immune response against Brucella spp. In the present study, in silico prediction of chimera indicates that B/T-cell epitopes from each protein caused immune stimulation.

In this study, we evaluated the immunogenicity properties of a chimeric protein with in silico prediction by using bioinformatics software. In silico data and results showed induction of both B- and T-cell mediated immune responses, which is important for the design of a protective vaccine. However, in our additional studies, it is an ongoing project and further studies focusing on enhancing the efficacy of TF, Bp26, and Omp31 recombinant subunit based vaccine using different adjuvants or vaccine strategies are underway. Estimation of the in silico and in vitro data accuracy and reliability were necessary; and in vitro experiments showed the accurate in silico designing of recombinant protein in physicochemical parameters, over expression and stability. Although different bioinformatics databases can result in different prediction by using different software, they do not affect the design of a new recombinant vaccine enormously, as we report in this and previous study. Our results indicate that this chimeric protein could be a potential immunogenic candidate for development of new subunit vaccines against Brucella. Moreover, other studies focusing on enhancing protective activity of Th1/2 response of this recombinant protein are underway.

Acknowledgements

The present study was financially supported by the Kerman University of Medical Sciences, Kerman, Iran (PhD Grant No. 93-1499).

References