In Silico Cloning and Analysis of Divalent Subunit OMP31-SODc Proteins As A Prophylaxis Vaccine Against *Brucella melitensis* Infection

Sri Kartika Wijaya*, Arizah Kusumawati, Andri Wardiana, Yana Rubiyana, Ulfatul Husnaa, and Adi Santoso

Research Center for Biotechnology, Indonesian Institute of Science (LIPI), Indonesia

Abstract

The urgency to develop a new protein subunit based vaccine candidate against *Brucella* was provoked by its frequent infection to human and livestock. Since *Brucella melitensis* is found as the most frequently isolated Brucella species from human, thus the outer membrane of *B. melitensis* becomes a prominent subcellular localization to search for promising antigen to be developed as vaccine candidate due to its interaction with host cell. Among outer membrane proteins suggested by Vaxign program, OMP31 was found as the most promising candidate. Moreover, analysis on other subcellular localization led our interest to SODc protein, which was expected to support OMP31 in triggering immune response. The OMP31-SODc divalent vaccine candidate was analysed *in silico* to predict its stable three-dimensional structure, cloning process and expectation on the ease during expression, purification and vaccine delivery to elicit the expected immune response.

Keywords: *Brucella melitensis*, outer membrane proteins, OMP31, SODc, divalent vaccine structure

---

*Corresponding author:
Cibinong Science Center, Jl. Raya Bogor Km. 46, Cibinong 16911, Indonesia
Tel. +62-21-8754587, Fax. +62-21-87754588
E-mail. sri.kartika.wijaya@lipi.go.id

Introduction

Brucellosis is a disease caused by gram-negative intracellular bacteria called *Brucella* (Siadat *et al*., 2012). The bacteria has a range of hosts including goat and sheep (*Brucella melitensis*) (Cloeckaert *et al*., 1992), cattle (*Brucella abortus*), swine (*Brucella suis*), dog (*Brucella canis*), sea water animals like dolphin and whale (*Brucella ceti*), and also desert wood rat (*Brucella neotomae*) (Gomez *et al*., 2013). Despite the differences on its host preferences, their physiological abilities and structure of the cell surface are the main factors of *Brucella* classification (Siadat *et al*., 2012). Some of *Brucella* species are the major cause of abortion and reproductive disorders in animals, whereas other species are found to cause mastitis, epididymitis, arthritis and abscesses in various organs. Cross infection to human is usually occurred while maintaining a direct contact to infected animals, which includes consumption of contaminated raw meats, non-pasteurized milk and dairy products (Young, 1983; Young, 1995) with the possibility of human to human transmission (Mesner *et al*., 2007). Earlier, *B. melitensis* causes the most severe disease to human being, followed by *B. suis*, *B. abortus* and *B. canis* (Khan *et al.*, 2001; Lucero *et al.*, 2008). However, recent finding confirmed the presence of human *Brucella inopinata* (Jubier-Maurin *et al*., 2001).

Protection to *Brucella* infection could be achieved through the development of both killed and live-attenuated vaccines. Live attenuated *Brucella* vaccine has been proven to be able to prompt higher immunity response compare to killed vaccines. However, the drawbacks of live-attenuated vaccines, which can cause late abortion in female animals during pregnancy (Siadat *et al*., 2012), could not be tolerated. Thus, the development of other vaccine groups to combat *Brucella* infection is necessary. A closer study on brucellosis, particularly on its virulence strategies, is a major requirement to understand the pathway of *Brucella* bacteria in manipulating host defence mechanism. By understanding the immune response routes, the assessment and construction of desired vaccines can be done precisely and right on target.

Early study on *Brucella* molecular movement in animal model confirmed that macrophages are the target of its infection (Porte *et al*., 1999; Roop *et al*., 2004). Recently, *Brucella* infection was also found to inhibit DCs (dendritic cells) maturation with the involvement of Btp1 protein, which had consequences on inappropriate cytokine secretion and antigen presentation (Salcedo *et al*., 2008). This scheme was highlighted as the major motion for *Brucella* escaping the host defence mechanism. Moreover, the implications of *Brucella* outer membrane proteins (OMPs) to play the role on DCs maturation, via TNF-alpha blockage (Salcedo *et al*., 2008), generate an intensive interest on the outer membrane proteins (Thompson *et al*., 2013) of *Brucella*. The OMPs contain lipopolysaccharide (LPS), protein and phospholipids. The OMP members are one of suggested antigens that fulfils the criteria to be developed as vaccine against brucellosis along with other nine proteins located in cytoplasm, periplasm and cytoplasmic membrane.

Recently, an advanced progress on bioinformatics tools makes a great improvement on vaccines development. Numerous free and proprietary softwares are available to help researchers sharpen their goals on constructing a desired vaccine. Bioinformatics serves as a filter of a bulk of information available to make a prediction of the vaccines target, and moreover, analysis of immune epitopes becomes essential in the case of infectious diseases. Here in this study, we present our interest to design a divalent OMP31-SODc protein vaccine against Brucella infection, extended with some analysis through several bioinformatics softwares in the purpose of making a prediction on its characteristics, confirmation of selected protein criteria and its structure prior to wet laboratory cloning. Selection of OMP31 as a vaccine candidate was based on Vaxign web analysis for *B. melitensis* infection with the outer membrane as one of the preferred parameter. We understand that there are other OMPs that were also recommended to be developed as vaccine candidate. However, we were primarily interested in OMP31 based on its contribution to the virulence of *B. melitensis* that has the highest infection rate to human, and also due to the urgency to develop a new subunit protein-based vaccine candidate, since the current available vaccine of OMP31 is DNA-based vaccine. At last, *in silico* cloning analysis was also discussed, as a confirmation of in-frame insertion.

Materials and Methods

Vaccine target prediction using Vaxign software. Free Vaxign website accommodates analysis of 35 genome groups, including *Brucella* genome group that covers 31 *Brucella* strains. Each protein of the genome can be assessed based on its subcellular localization. However, for vaccine candidates, proteins located on the cell surface (*e.g* outer membrane, cytoplasmic membrane), are preferable due to its interaction with the host cell, which can accelerate required immune response. In this study, *B. melitensis* 16M strain and the outer membrane were selected as the target genome and subcellular target location, respectively. Other filtering options available on the website include adhesin value and number of transmembrane helices. Adhesin is important for microbial adherence to the host cell and has a cut-off value of 0.5. A cut-off value lower than 0.5 is considered to have a less pathogenic activity, which is not preferable on the selection of a vaccine candidate (Gallagher & Buchmeier, 2001). On the other hand, the number of transmembrane helices was set to be less or not more than 1. Proteins that were identified to have more than one transmembrane helices were excluded due to the difficulties during cloning, expression and purification of proteins exhibiting multiple transmembrane spanning regions. The option to analyse ortholog was ignored since we were not searching for conserved vaccine gene target between the strains but specifically targeting *B. melitensis* instead.

Binding prediction to MHC class I and class II are the second step analysis offered by Vaxign. However, epitope analysis of OMP31 protein was found not satisfying without selection of host preferences. Moreover, the option of host selection on Vaxign is limited for only human, mouse, chimpanzee and macaque. Therefore, we used Uniprot online web resources (http://www.uniprot.org) to obtain more information of OMP31 immunogenic region. Similar strategy was applied to analyse the other preferred protein as vaccine candidate, which is SODc protein.

Prediction of three-dimensional structure. The genes of both selected proteins were retrieved from NCBI website. The synthetic gene construct was built with the consideration that OMP31 is the main target of protein vaccine. Therefore, OMP31 was settled at the N-terminal end of the construct. The OMP31 and SODc proteins were joined together via a flexible linker (Gly4Ser). The amino acids sequence of the construct was submitted to online prediction tool on Expasy (http://swissmodel.expasy.org), followed by 3D prediction on Pymol software. The amino acids sequence was also submitted to I-Tasser website (http://zhanglab.ccmb.med.umich.edu). Finally, a comparison was made between the results. The prediction of the structure was useful to examine the composition of divalent protein vaccine along with confirmation of its exposed OMP31 epitopes.

Codon optimization. Since *Pichia pastoris* was selected as the host for protein expression, both OMP31 and SODc codons were optimized using selected tools to improve gene expression. There are several online programs freely available to optimize preferred gene. Here, we ran the optimization using both free and proprietary programs and compared some of the suggested results of codon optimization. A free JCat program (http://www.jcat.de) was used to optimize the gene. The optimized gene result was then compared to optimization analysis by Geneious program.

***In silico* cloning analysis with Geneious software.** Geneious is powerful software that supports a rapid and reliable molecular biology analysis, including synthetic biology. The features include codon optimization, cloning option through four cloning strategies, primer design and many more. In this study, *in silico* ligation between synthetic DNA insert and pPIC9K plasmid was completed following traditional cloning methods, which included digestion method with selected restriction enzymes (*Eco*RI and *Not*I), virtual confirmation of gel agarose, and cloning the insert into pPIC9K plasmid. Optional Geneious element of protein translation will confirm the correct insertion by accurate protein reading.

Results

Vaxign Web results analysis.

The first analysis of selecting vaccine target was done using Vaxign website. This website is provided with some tools to narrow highly potential protein candidates. Vaxign web components include prediction of subcellular localization, transmembrane domain prediction, calculation of adhesion probability, protein conservations among different genomes, comparison analysis between the protein and hosts, and also binding prediction of the proteins to MHC class I and class II. Among all of those characteristics, the easiest selection parameter of protein characteristic to be considered is the value of adhesin probability. Taking into consideration of their adhesin value, the OMPs family along with lipoprotein was suggested as protein candidates to be developed as prophylaxis vaccine to combat *B. melitensis* infection.

As seen on Table 1, some of *B. melitensis* OMPs were recognized as the prominent vaccine candidates, namely OMP19, OMP31, OMP22, and OMP25, along with outer membrane lipoproteins. Moreover, none of these proteins have similarity to host proteins (human or mouse), which means that they are predicted to be able to prompt immune response while unintentionally or deliberately introduced to hosts. The OMP31 is found to be one of the promising candidates and the most immune-dominant (Azad *et al*., 2013). We recognized that there are two OMP31 proteins presented on Vaxign results that are different on their sequences and length (BMEI0402 and BMEII0844, respectively) as shown in Table 1. Both are member of the brucella-rhizobium porin (brp) family and developed by the same strain (strain 16M/ ATCC 23456/ NCTC 10094), which means that they might share similar molecular function and holding similar biological process. However, one of them has been known to have higher adhesion value and exhibits epitope region for monoclonal antibody binding (Figure 1A) that is situated on the most hydrophilic part of the protein (Vizcaíno *et al*., 1996), which is shown on white-grey colour of the OMP31 ribbon structure depicted on Figure 1B. Based on these results, BMEII0844 gene was selected as vaccine candidate against *B. melitensis*.

Table 1. The Vaxign web analysis result
Genome: Brucella melitensis 16M
Subcellular Localization: OuterMembrane
Maximum Number of Transmembrane Helices: 1
Minimum Adhesin Probability: 0.51

#	Protein Accession	Protein GI	Gene Symbol	Locus Tag	Gene ID	Protein Note	Localization	Probability	Adhesin Probability
1	NP_539053.1	17986419	omp19	BMEI0135	1195847	OUTER MEMBRANE LIPOPROTEIN	Outer Membrane	1	0.679
2	NP_539319.1	17986685	BMEI0402	BMEI0402	1196113	31 KDA OUTER-MEMBRANE IMMUNOGENIC PROTEIN PRECURSOR	Outer Membrane	1	0.621
3	NP_539321.1	17986687	BMEI0404	BMEI0404	1196115	hypothetical protein BMEI0404	Outer Membrane	0.949	0.544
4	NP_539371.1	17986737	BMEI0454	BMEI0454	1196165	OUTER MEMBRANE PROTEIN W PRECURSOR	Outer Membrane	1	0.586
5	NP_539634.1	17987000	BMEI0717	BMEI0717	1196428	22 KDA OUTER MEMBRANE PROTEIN PRECURSOR	Outer Membrane	0.992	0.609
6	NP_539924.1	17987290	BMEI1007	BMEI1007	1196718	25 KDA OUTER-MEMBRANE IMMUNOGENIC PROTEIN PRECURSOR	Outer Membrane	1	0.562
7	NP_539942.1	17987308	BMEI1025	BMEI1025	1196736	OUTER MEMBRANE PROTEIN E	Outer Membrane	0.949	0.595
8	NP_539996.1	17987362	BMEI1079	BMEI1079	1196790	LIPOPROTEIN NLPD	Outer Membrane	0.993	0.608
9	NP_540166.1	17987532	omp25	BMEI1249	1196960	25 KDA OUTER-MEMBRANE IMMUNOGENIC PROTEIN PRECURSOR	Outer Membrane	1	0.551
10	NP_540411.1	17987777	BMEI1494	BMEI1494	1197205	hypothetical protein BMEI1494	Outer Membrane	0.993	0.543
11	NP_540747.1	17988113	BMEI1830	BMEI1830	1197541	25 KDA OUTER-MEMBRANE IMMUNOGENIC PROTEIN PRECURSOR	Outer Membrane	1	0.592
12	NP_540789.1	17988155	BMEI1872	BMEI1872	1197583	CELL SURFACE PROTEIN	Outer Membrane	0.992	0.656
13	NP_540994.1	17988361	omp10	BMEII0017	1197788	OUTER MEMBRANE LIPOPROTEIN	Outer Membrane	1	0.627
14	NP_541082.1	17988449	BMEII0105	BMEII0105	1197876	IRON-REGULATED OUTER MEMBRANE PROTEIN FRPB	Outer Membrane	1	0.521
15	NP_541354.1	17988721	BMEII0376	BMEII0376	1198148	HEAT RESISTANT AGGLUTININ 1 PRECURSOR	Outer Membrane	1	0.546
16	NP_541822.1	17989189	BMEII0844	BMEII0844	1198616	31 kDa outer-membrane immunogenic protein precursor	Outer Membrane	1	0.671

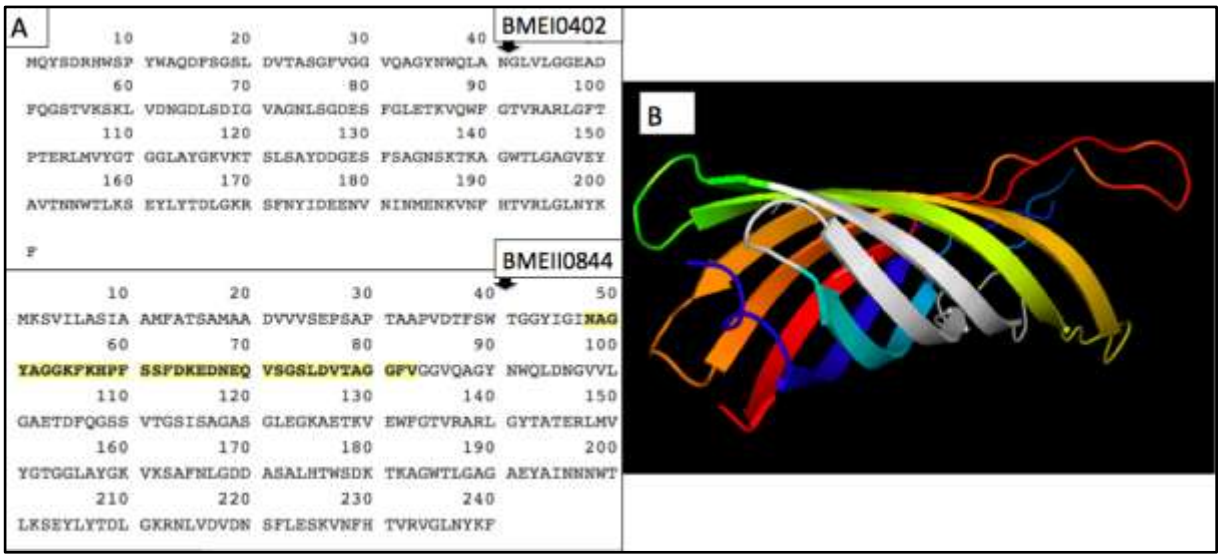

Figure 1. Figure 1A shows the highlighted section on BMEII0844 retrieved from Uniprot webpage, which is the antibody recognition site that is considered as immunogenic site of the protein. Even though both OMP31 are different on their amino acid sequences, the three dimensional structures of both OMPs share the same 3D template model (PDB entry: 1Qj8A) as shown in Figure 1B, which is the ribbon presentation of OMP31 structure from Protein Model Portal (http://www.proteinmodelportal.org) on Pymol program.

Table 1 shows that as many as 16 proteins were presented on result pages if the selected parameters were outer membrane subcellular location, adhesin probability value of 0.5 and number of transmembrane helices protein < 1 for *Brucella melitensis* 16M strain. Two OMP31 strains were found, BMEI0402 and BMEII0844, respectively. Both share the same cellular location but differ in protein length as shown in Figure 1A.

Selection of the second protein was focused on those directed to other subcellular localization. While an outer membrane protein was selected during the search for the major vaccine candidates, we took a chance to select for other candidates from the bacterial periplasm. We interested in SODc protein that

was suggested to support Brucella infection in a host cell by protecting the bacteria from macrophages oxidative killing (Gee *et al*., 2005; Piddington *et al*., 2001). As formerly explained that outer membrane is the primary location for selecting vaccine candidates, Vaxign web report on periplasm antigen candidates confirmed that some suggested proteins have a range of adhesin values between 0.1-0.59 (Figure 2A), which are lower than those of proteins located on the outer membrane. However, the decision to select SODc is not only because of its contribution to *Brucella* survival, but also due to the small size of SODc protein. Therefore, it came into prediction that its attachment to the C–terminus of OMP31 would not interfere with the exposed immunogenic site of OMP31, which is important for prompting the immune response.

Vaxign web results indicated that SODc protein exhibits similarity to human protein (Figure 2A). The similarity is found at 2 sites with sequence motif G-HGFHVHE and G-R-ACGVI, respectively. The G-HGFHVHE motif was shown in green, whereas the G-R-ACGVI motif, which is also the C–terminus part of SODc protein, was shown in yellow (Figure 2B). The adhesin value of SODc is lower than 0.5, which implies that SODc does not have contribution to bacterial adherence to host cell, which might also has a lower capability to induce host immune response. Nevertheless, a study confirmed that immunization of mice with SODc DNA vaccine was able to elicit a partial immune response against *B. abortus* (Onate *et al*., 2003).

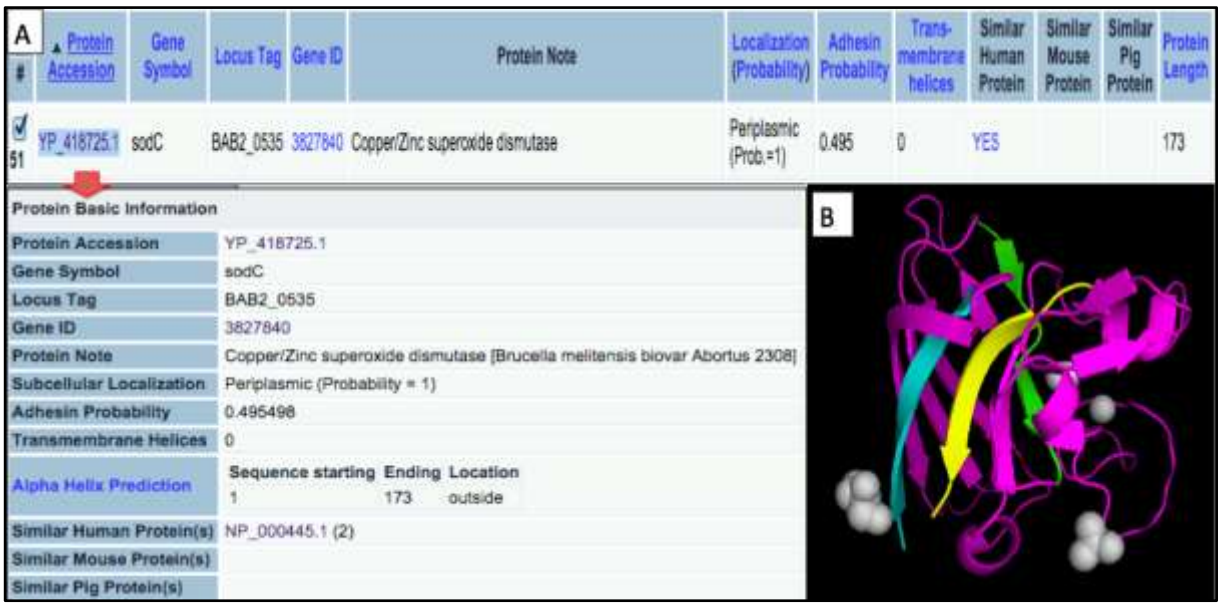

Figure 2. SODc protein is one of the recommended proteins located in periplasm for vaccine development along with other 85 suggested proteins (not shown). Figure 2A is a part of Vaxign webpage result that confirmed the presence of SODc as the suggested protein. When the highlighted protein accession button was selected (shown by the red arrow), the Vaxign will present similar basic information of designated protein but in different format view. Figure 2B is the 3D-structure of SODc protein obtained from PDB and Pymol software. It shows the N-terminal part (cyan), the C-terminal part (yellow), as well as the metal ion constituted on SODc protein (silver spheres).

Gene Construction and structure homology analysis.

As previously explained, the OMP31 protein was placed at the N-terminus of the fusion protein, followed by Gly4Ser-linker and SODc protein at the C-terminus. Examination of each OMP31 and SODc single protein was carried out with Pymol program. Once examination of each independent protein was completed, the prediction of the structure of divalent OMP31-SODc protein vaccine was carried out. Moreover, the structure of OMP31-SODc divalent vaccine was also examined via I-Tasser and Swiss-Model online program. However, the suggested results of I-Tasser were not satisfying which might be caused by the lack of information of 3D structure of OMP31 protein (data not shown).

However, I-Tasser online software recognized SODc protein and showed a correct 3D modelling. We are well informed that the crystal structure of OMP31 protein is not appropriately determined yet. Nevertheless, with similarities around 25-30 % of its sequences on the Protein Model Portal software, the results suggested that OMP31 structure has a ß-barrel shape, similar to OMPX protein in general. As depicted in Figure 3, the three-dimensional structure model of the divalent vaccine was constructed manually with the support of Pymol software.

Codon optimization.

After designing the gene construct, the codons of OMP31 and SODc proteins were optimized to improve heterologous gene expression in the yeast *Pichia pastoris.* First attempt of codon optimization was done at Java Codon Adaptation Tool website (JCat). The DNA sequence of OMP31-linker-SODc was adapted into the codon preference for *Saccharomyces cerevisae*, because the *P. pastoris* codon usage is considered to be the same as that of *S. cerevisae*. The Codon Adaptation Index (CAI) value of the original OMP31-SODc proteins gene was found to be 0.1, whereas after adaptation it was dramatically increased to 0.975, very close to maximum CAI value of 1.0 (data not shown). A higher CAI value is expected to give an impact on higher expression rate in *P. pastoris*.

As a comparison, an optimization by means of Geneious software was also done, and the resulted optimized codons from both JCat and Geneious were aligned to observe differences. Alignment between non-optimized and optimized codon-using JCat and Geneious revealed that optimized codon via Geneious showed less modifications compare to optimized codon by JCat software.

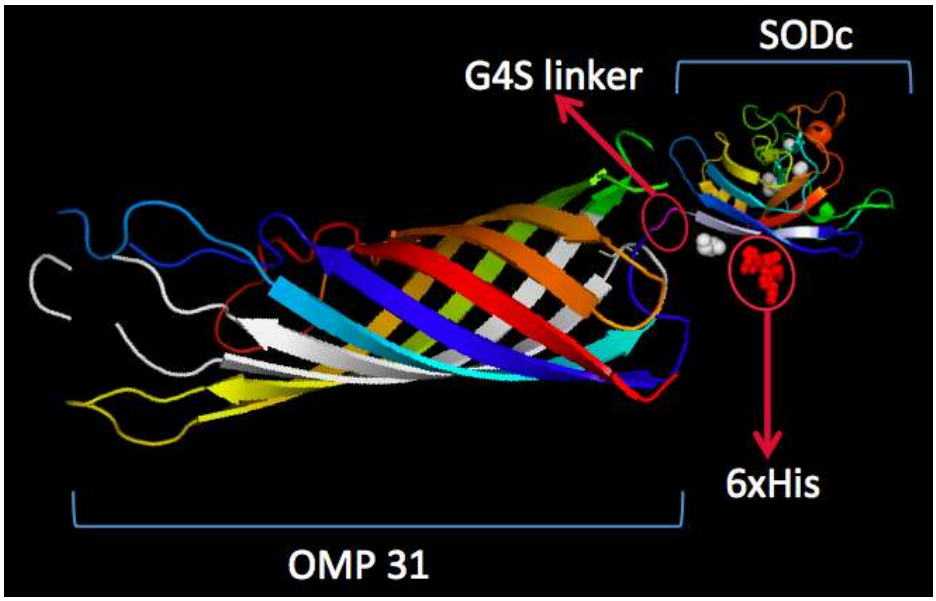

Figure 3. Schematic assembly of OMP 31-Linker-SODc-6xHis-tag divalent vaccine construct. A flexible linker was chosen to join both proteins. The use of more than one linker was avoided to minimize the possibility of SODc protein covers the OMP31 immunogenic site (shown by white-grey part on OMP 31 structure). At the C-terminal end of the fusion protein, the 6×His are shown to be exposed, which is important during the purification process.

However, since CAI calculation of the optimized codon using Geneious was lower than that using JCat (data not shown), thus the optimized codon obtained by JCat was used for further experiments.

***In silico* cloning using Geneious Software.**

Confirmation of in frame insertion was finalised via *in silico* analysis. The synthetic DNA construct comprises 2 restriction sites, a linker, 2 protein target genes, stop-codon and 6×His-tag (Figure 4). Geneious software was used to perform *in silico* ligation process. This tool enables the creation of a gene manually or in imported form from other sources in a matching format. Both vector (pPIC9K) and insert (synthetic DNA construct) were digested virtually with *Eco*RI and *Not*I, respectively. Geneious will automatically keep the restriction site annotations of the gene, thus a

virtual ligation result of vector and insert will prove the correct or incorrect translational reading of the gene.

As seen on the Figure 4, the translated reading of the gene after ligation process showed a correct protein translation. Both ends of the ligation site, shown in light green at N- and C-terminus, pointed by the red arrows, these sites determine the location of insertion. At the N-terminus, *Eco*RI restriction site was recognized as well as the start-codon of the OMP31 protein. On the other hand, at the C-terminus, 6×His-tag was translated correctly that represented an additional proof of a correct protein translation.

Discussion

With an increased concern over *Brucella* infection, vaccination is considered to be a strategic approach to deal with economic loss and human infection severely caused by *Brucella* bacteria. The requirements for a safe and effective subunit vaccine against *Brucella* infections become urgent, taking into account that live attenuated vaccines were prohibited in some countries due to their side effects. In this study, investigation on some protective antigen candidates of Brucella's outer membrane and other subcellular location was done with the support of Vaxign web site. The analysis results showed that the OMP31 protein has been found to be a dominant candidate along with other outer membrane proteins.

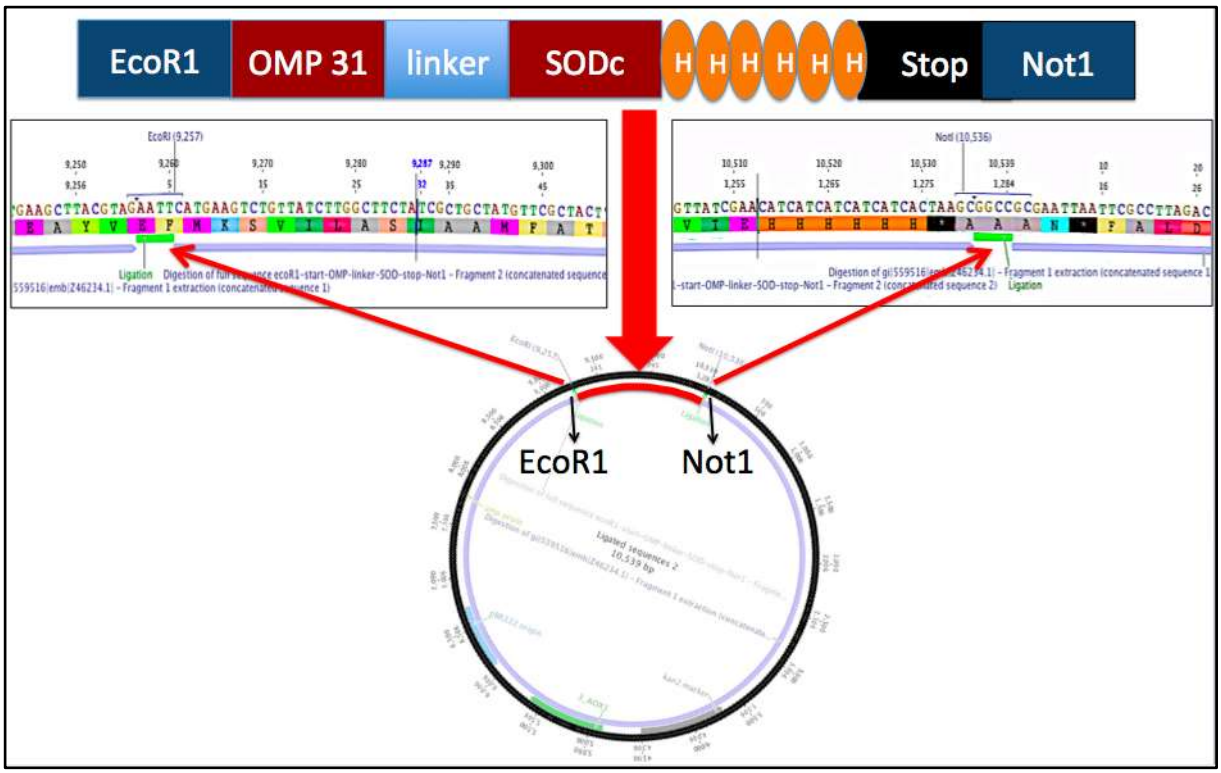

Figure 4. A complete virtual ligation result of DNA synthetic gene (as insert) and plasmid pPIC9K (as the vector) by using Geneious as the confirmation of the in frame cloning. Prior to *in silico* analysis of the ligation process, the sequence of the plasmid vector was retrieved from NCBI in a FASTA format. On the other hand, the insert was created manually on Geneious. The insert is the optimized gene as described previously on codon optimization section.

Brucella is divided into smooth and rough strains, whereas smooth *Brucella* strain is considered to play principal roles during infection and host attachment (Cloeckaert *et al*., 1990). Early study of outer membrane of smooth-*Brucella* using monoclonal antibody binding confirmed the presence of 10 kDa, 16 kDa, 19 kDa, and 98 kDa proteins exposed on both outer membrane of *B. melitensis* and *B. abortus*. However, it was only the 31 kDa protein that is absence on smooth *B. abortus* (Cloeckaert *et al.*, 1990). Therefore, since protection against *B. melitensis* infection is the major objective of this research, the absence of

OMP31 on smooth *B. abortus* becomes one of supporting evidence to make conclusion that OMP31 is rather specific to differentiate initiated protection between *B. abortus* and *B. melitensis*. This aspect is important considering that SODc is sometimes found capable of developing a protective immunity against *B. abortus* (Araya *et al.*, 1989*;* Onate *et al.*, 2003), while some other previous researches experienced failure in generating SODc-specific protection in mice (Latimer *et al*., 1992). However, there is a strong evidence that SODc is required to improve the bacterial living through its catalytic activity, and newest study on single SODc protein expression confirms that SODc has the ability to induce immune response as well as it has potential to be developed as LPS-free protein for diagnostic use (Sung *et al.*, 2014).

The OMP 31 protein is known to have molecular function with porin activity (Vizcaíno *et al*., 1996), which is categorized as a member of transporters available in nature. Porins are specifically described to have ß-barrel form and its structure allows energy-independent passage of solutes across the porin tunnel (Saier, 2000). Others stated that porin is also crucial for pathogenesis of bacteria due to contact with host tissue (Achouak *et al*., 2001; Galdiero *et al*., 2003; Galdiero *et al*., 2012). Moreover, the entire porin group members are similar on their abundance in native membrane along with the consistency in their ß-barrels structure (Zeth *et al*., 2000). Therefore, even though 3D structure of OMP31 protein was not perfectly depicted yet, the structure of other members of porin group can be used as a reference to predict the OMP31 structure. Basically, we used homology-protein modelling approach to define a prediction on OMP31 structure. The basic knowledge is based on common premise that about 50 % identical alignment result between protein sequences will be considered to be close-related sequences. Those closely related proteins will have higher probability to share similar conformation. Finally, similar protein conformation and 3D structure will determine similar protein function. In this study, a general OMP-X structure as suggested by Protein Model Portal software was used as a parental model. It was also possible to use OMP25 as a model structure, because early study on OMP31 revealed that OMP31 protein sequence is similar to OMP25 and RopB sequence of *R. leguminosarum* (Vizcaíno *et al*., 1996) and the OMP25 as a parental was also suggested by Swiss-Model software. However, OMP25 covers only 18-19 % of the OMP31 sequences. This identity value is lower than OMP-X structure in general which exhibits around 25 % identical sequence to OMP31. Here, we recognized the valuable exploitation of the available online softwares, so that discrepancies and comparable results shown by the Protein Model Portal and Swiss-Model resulted in a better OMP31 protein modelling based on the percentage of identical sequences, even though both softwares aligned the target and parental sequence through local alignment method.

On the other hand, 3D structure of SODc protein is well defined. The SODc protein is classified into a group of metalloprotein by the existence of four metals. Comparison of of SODc three-dimensional structures obtained from both I-Tasser and Swiss-Model confirmed similar protein models. However, a structure for the fusion protein consisted of OMP31 and SODc proteins was not successfully determined by I-Tasser software. At this point, a decision to make a manual prediction was made based on independent structures of both proteins as shown on Figure 3. We understood that the predicted structure exhibited simplification on some of the features. For example, this 3D analysis study of OMP31-SODc fusion protein has not consider about the interaction between residues and side chains, because the prediction of fusion protein structure was created based on the curve of amino acids main chains. However, the contribution of side chains was credited on protein hydrophobicity analysis, in which will be explained further. Moreover, the utilisation of flexible linker, which connects both proteins, enables orientations in different angles.

We took into account that side chains of protein may contribute to polarity, electric charge and in general may generate different folding pattern. Incorrect protein folding is unfavourable due to the functional loss. Moreover, the amino acid side chains may alter protein preference of aqueous environment. Therefore, hydrophobic effect of side chain residues was analysed via GPMAWlite software (http://www.alphalyse.com), which gave a hydrophobicity index of -0.19. Thus, the

protein could be considered as a neutral protein. The synthetic gene construct stability index was obtained by analysis using Expasy-ProtParam (http://web.expasy.org/protparam/). Both hydrophobic value and instability index validated the common premise that most of proteins are marginally stable, which is important to facilitate protein turnover.

In majority, porins have overall dimensional size around 35 Å (3.5 nm) with the maximum length at 50 Å (5 nm) (Galdiero *et al*., 2012). The OMP31-SODc fusion protein will approximately have a similar size to overall size of porin proteins. Antigen Presenting Cells (APCs) are able to take up particles with both large and small size (Gamvrellis *et al*., 2004a), where macrophages ingest substances in bacterial size (500-2000 nm) and dendritic cells (DCs) seem to phagocyte any substances in smaller size (more in viral size which is around 20-100 nm) (Fifis *et al*., 2004; Gamvrellis *et al*., 2004b). Therefore, the OMP31-SODc vaccine is expected to specifically target dendritic cells *in vivo*, thus immuno-stimulatory carriers might be needed in the future to deliver a 'danger' signal to DCs due to the incredible small size of the divalent vaccine. The carrier, in particular, has to interact specifically with DCs or macrophages but not to other cell types.

Having informed with adequate results of divalent vaccine structure and its characteristic, the expectation to produce OMP31-SODc proteins in active state, correct protein folding and highly soluble was further supported by codon optimization of the synthetic gene. The yeast *P. pastoris* and pPIC9K plasmid were selected as host for high multiple expressions of OMP31-SODc proteins. As it is well known that codon distribution between species is not random, especially with the existence of synonymous codons, adoption of high prevalence codon of the selected host becomes one of important strategy to improve protein expression due to its influence during protein translation. Herein, the synthetic gene codon was optimized with the help of various software tools. Disparities between the optimized codons provide more options that could be adjusted to the research goals whether the codon will be exceedingly adopted with most of the codons were changed, or it will just be partially or slightly changed. Optimized codon that has highest CAI value was selected, with expectation that *P. pastoris* will be able to express the protein in a high manner. However, we realized that codon optimization is not always successful to force higher protein expression, because codon optimization method was built by calculating the codon frequency of the codon usage and neglecting the influence of environmental growth condition and the cell cycles. Finally, a future wet laboratory work will be a judgement of these bioinformatic prediction frameworks. Initial laboratory scale expression with some modifications on growth conditions might be needed to obtain abundant level of desired divalent OMP31-SOD subunit vaccine, along with efficient purification strategies. Nonetheless, this bioinformatic study of OMP31-SODc divalent protein vaccine will be a reference for future works, in particular in minimizing the possibility of misconducted research due to the lack of information of core protein knowledge.

Acknowledgements

The authors would like to thank the *Kegiatan Unggulan* LIPI 2015 project financed by Indonesian Institute of Sciences. We would also like to thank the anonymous referees for the valuable comments and suggestions that have significantly improved the manuscript.

References

Achouak, W., Heulin, T., & Pages, J. (2001). Multiple facets of bacterial porins. *FEMS Microbiology Letters*, *199*, 1-7.

Araya, L. N., Elzer, P. H., Rowe, G. E., Enright, F. M., & Winter, A. J. (1989). Temporal development of protective cell-mediated and humoral immunity in BALB/c mice infected with *Brucella abortus*. *Journal of immunology*, *143*, 3330-3337.

Azad, A. K., Hasan, M. M., Hossain, M. S., Rahman, M. R., & Chowdhury, P. A. (2013). In silico analysis of outer membrane protein 31 of *Brucella* app. to identify and characterize the potential T cell epitope. *International Journal of Pharma Medicine and Biological Sciences*, *2*(3), 28-45.

Cloeckaert, A., de Wergifosse, P., Dubray, G., & Limet, J. N. (1990). Identification of

seven surface-exposed *Brucella* outer membrane proteins by use of monoclonal antibodies: immunogold labeling for electron microscopy and enzyme-linked immunosorbent assay. *Infection and Immunity*, *58*, 3980-3987.

Cloeckaert, A., Jacques, I., de Wergifosse, P., Dubray, G., & Limet, J. N. (1992). Protection against Brucella melitensis or Brucella abortus in mice with immunoglobulin G (IgG), IgA, and IgM monoclonal antibodies specific for a common epitope shared by the Brucella A and M smooth lipopolysaccharides. *Infection and Immunity*, *60*, 312-315.

Fifis, T., Gamvrellis, A., Crimeen-Irwin, B., Pietersz, G. A., Li, J., Mottram, P. L., McKenzie, I. F. C., & Plebanski, M. (2004). Size-dependent immunogenicity: therapeutic and protective properties of nano-vaccines against tumors. *Journal of Immunology*, *173*, 3148-3154.

Galdiero, M., Vitiello, M., & Galdiero, S. (2003). Eukaryotic cell signaling and transcriptional activation induced by bacterial porins. *FEMS Microbiology Letters*, *226*, 57-64.

Galdiero, S., Falanga, A., Cantisani, M., Tarallo, R., Pepa, M. E. D., D'Oriano, V. & Galdiero, M. (2012). Microbe-host interactions: Structure and role of Gram-negative bacterial porins. *Current Protein & Peptide Science*, *13*, 843-854.

Gallagher, T. M., & Buchmeier, M. J. (2001). Coronavirus spike proteins in viral entry and pathogenesis. *Virology*, *279*, 371-374.

Gamvrellis, A., Leong, D., Hanley, J. C., Xiang, S. D, Mottram, P., & Plebanski, M. (2004). Vaccines that facilitate antigen entry into dendritic cells. *Immunology and Cell Biology*, *82*, 506-516.

Gee, J. M., Valderas, M. W., Kovach, M. E., Grippe, V. K., Robertson, G. T., Ng, W.-L, Richardson, J. M., Winkler, M. E., & Roop, R. M. (2005). The *Brucella abortus* Cu, Zn superoxide dismutase is required for optimal resistance to oxidative killing by murine macrophages and wild-type virulence in experimentally infected mice. *Infection and immunity*, *73*, 2873-2880.

Gomez, G., Adams, L. G., Rice-Ficht, A., & Ficht, T. A. (2013). Host-Brucella interactions and the Brucella genome as tools for subunit antigen discovery and immunization against brucellosis. *Frontiers in Cellular and Infection Microbiology*, *3*, 17-32.

Jubier-Maurin, V., Boigegrain, R. A., Cloeckaert, A., Gross, A., Alvarez-Martinez, M. T., Terraza, A., Liautard, J., Kohler, S., Rouot, B., Dornand, J., & Liautard, J. P. (2001). Major outer membrane protein Omp25 of *Brucella suis* is involved in inhibition of tumor necrosis factor alpha production during infection of human macrophages. *Infection and Immunity*, *69*, 4823-4830.

Khan, M. Y., Mah, M. W., & Memish, Z. A. (2001). Brucellosis in pregnant women. *Clinical Infectious Diseases*, *32*, 1171-1177.

Latimer, E., Simmers, J., Sriranganathan, N., Roop Ii, R. M., Schurig, G. G., & Boyle, S. M. (1992). *Brucella abortus* deficient in copper/zinc superoxide dismutase is virulent in BALB/c mice. *Microbial Pathogenesis*, *12*, 105-113.

Lucero, N. E., Ayala, S. M., Escobar, G. I., & Jacob, N. R. (2008). Brucella isolated in humans and animals in Latin America from 1968 to 2006. *Epidemiology and Infection*, *136*, 496-503.

Mesner, O., Riesenberg, K., Biliar, N., Borstein, E., Bouhnik, L., Peled, N., & Yagupsky, P. (2007). The many faces of human-to-human transmission of Bbucellosis: Congenital Iifection and outbreak of nosocomial disease related to an unrecognized clinical case. *Clinical Infectious Diseases*, *45*, e135-e140.

Onate, A. A., Cespedes, S., Cabrera, A., Rivers, R., Gonzalez, A., Munoz, C., Folch, H. & Andrews, E. (2003). A DNA vaccine encoding Cu, Zn superoxide dismutase of *Brucella abortus* induces protective immunity in BALB/c mice. *Infection and Immunity*, *71*, 4857-4861.

Piddington, D. L., Fang, F. C., Laessig, T., Cooper, A. M., Orme, I. M., & Buchmeier, N. A. (2001). Cu, Zn superoxide dismutase of *Mycobacterium tuberculosis* contributes to survival in activated macrophages that are generating an oxidative burst. *Infection and Immunity*, *69*, 4980-4987.

Porte, F., Liautard, J. P., & Kohler, S. (1999). Early acidification of phagosomes containing *Brucella suis* is essential for intracellular survival in murine macrophages. *Infection and Immunity*, *67*, 4041-4047.

Roop, R. M., Bellaire, B. H., Valderas, M. W., & Cardelli, J. A. (2004). Adaptation of the Brucellae to their intracellular niche. *Molecular Microbiology*, *52*, 621-630.

Saier, M. (2000). A functional-phylogenetic classification system for transmembrane solute transporters. *Microbiology and Molecular Biology Reviews, 64*, 354-411.

Salcedo, S. P., Marchesini, M. I., Lelouard, H., Fugier, E., Jolly, G., Balor, S., Muller, A., Lapaque, N., Demaria, O., Alexopoulou, L., Comerci, D. J., Ugalde, R. A., Pierre, P. & Gorvel, J. P. (2008). Brucella Control of dendritic cell maturation is dependent on the TIR-containing protein Btp1. *PLoS Pathogens*, *4*(2):e21.

Siadat, S., Salmani, A., & Aghasadeghi, M. (2012). Brucellosis vaccines: An overview, zoonosis. Lorenzo-Morales, J. (Ed.). ISBN: 978-953-51-0479-7, InTech. Available from: http://www.intechopen.com/books/zoonosis/brucellosis-vaccines.

Sung, K. Y., Jung, M., Shin, M. K., Park, H. E., Lee, J. J., & Kim, S. (2014) Induction of immune responses by two recombinant proteins of *Brucella abortus*, outer membrane proteins 2b porin and Cu/Zn superoxide dismutase, in mouse model. *Journal of Microbiology and Biotechnology*, Jun 28;*24*(6):854-61. PubMed PMID: 24608566. Epub 2014/03/13. eng

Thompson, C.M., Petiot, E., Lennaertz, A., Henry, O., & Kamen, A. A. (2013). Analytical technologies for influenza virus-like particle candidate vaccines: challenges and emerging approaches. *Virology Journal*, *10*, 141-141.

Vizcaíno, N., Cloeckaert, A., Zygmunt, M. S., & Dubray, G. (1996). Cloning, nucleotide sequence, and expression of the Brucella melitensis omp31 gene coding for an immunogenic major outer membrane protein. *Infection and Immunity*, *64*, 3744-3751.

Young, E. J. (1983). Human brucellosis. *Reviews of Infectious Diseases*, *5*, 821-842.

Young, E. J. (1995). An overview of human brucellosis. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, *21*, 283-289; quiz 290.

Zeth, K., Diederichs, K., Welte, W., & Engelhardt, H. (2000). Crystal structure of Omp32, the anion-selective porin from *Comamonas acidovorans*, in complex with a periplasmic peptide at 2.1 Å resolution. *Structure*, *8*, 981-992.