Expression and Purification of LigA Antigen, a Surface Lipoprotein in the Pathogenic *Leptospira interrogans*

Aida Chalesh1, Pejvak Khaki2*, Soheila Moradi Bidhendi3, Majid Tebianian 4, and Morteza Taghizadeh Tarnabi 5

1 Department of Microbiology, North Tehran Branch, Islamic Azad University, Tehran, Iran. , ORCID: https://orcid.org/0009-0001-5850-0812

2 Department of Microbiology, Razi Vaccine and Serum Research Institute, Agricultural Research, Education and Extension Organization (AREEO), Karaj, Iran, ORCID: https://orcid.org/0000-0001-8839-1023

3 Department of Microbiology, Razi Vaccine and Serum Research Institute, Agricultural Research, Education and Extension Organization (AREEO), Karaj, Iran, ORCID: https://orcid.org/0000-0003-0069-0585

4 Department of Immunology, Razi Vaccine and Serum Research Institute, Agricultural Research, Education and Extension Organization (AREEO), Karaj, Iran, ORCID: https://orcid.org/0000-0002-9778-5048

5 Department of Medical Vaccine, Razi Vaccine and Serum Research Institute, Agricultural Research, Education and Extension Organization (AREEO), Karaj, Iran, ORCID: https://orcid.org/0000-0001-5456-4889

Abstract

Although considerable progress has been made in leptospiral vaccine development, their use is limited because of short-term and serovar-specific immunity. Thus far, many approaches have been used to identify heterologous and cost-effective antigen(s) against the leptospirosis. Recent advances have identified leptospiral immunoglobulin-like (Lig) proteins as promising candidates for vaccine development. Hence, in this study, the recombinant LigA subunit consists of the ligA9, ligA10, ligA11, and ligA12 domains, were selected as conserved regions of the LigA protein. Immunoinformatics approaches including I-TASSER, ProSA, DiscoTope v2.0, and Molprobity were exploited to check the conformational accuracy.

Furthermore, 10 of the most efficient peptides for MHC-I and II grooves were predicted by the ElliPro servers, NetMHCpan 4.1 EL and NetMHCIIpan 4.1 EL. The Ramachandran plot showed acceptable conformations of the selected recombinant protein amino acid residues. The results showed that selected epitopes elicit both humoral- and cell-mediated immune responses. Hence, the selected epitopes were constructed in the pET41a^{+} plasmid and synthesized by General Biosystems. Recombinant plasmids were transferred to *E. coli* Top10-DH5α and BL21 StarTM (DE3) competent cells for cloning and expression, respectively. Plasmid transformation and purification were confirmed using polymerase chain reaction and enzymatic digestion. Recombinant LigA (r-LigA) was expressed in the presence of 0.5 M IPTG at 30°C for 16 hours. The SDS-PAGE result revealed the production of 38-kDa protein, which accumulated mostly in inclusion bodies and was purified using the urea method and dialysis. r-LigA protein dot blotting confirmed a high degree of accuracy of immunogenicity. The present study revealed that r-LigA is a promising candidate for developing diagnostic and subunit leptospirosis vaccines.

Keywords: Leptospirosis, Immunoinformatics approaches, Recombinant protein, LigA epitope

1. Introduction

Leptospirosis or Weil's disease is a zoonotic and a potentially serious bacterial disease caused by pathogenic spirochetes of the genus *Leptospira*. It is transmitted to humans primarily through contact with water or soil contaminated with the urine of infected animals, particularly rodents. Sporadic outbreaks have occurred throughout the world, and it is particularly endemic in tropical and subtropical regions. Thus, leptospirosis is an important public health concern in many countries (1).

One of the most effective strategies for combating leptospirosis is vaccination, which stimulates the immune system by producing antibodies against the bacteria. The first successful vaccine against Leptospirosis was reported in 1919, and it dramatically decreased the number of leptospirosis cases in Japan (2). Heat-killed whole-cell vaccines (becterin) against leptospirosis successfully protected coal miners in Kyushu, where the disease was endemic in 1933. Live-attenuated and lipopolysaccharide (LPS) vaccine were another alternative introduced vaccine. Major issues like lack of immunological memory, incomplete immunity, reactogenicity, and serovar-restricted protection have led to a decrease in the use of these types of vaccines.The drawbacks of these types of vaccines highlight the necessity of novel vaccine strategies (2-4). In addition, with recent advances in molecular techniques and the availability of the complete Leptospira genome sequence, the development of novel vaccines like recombinant proteins, using reverse vaccinology has shown promising outcomes. Recent studies have highlighted the key role of outer membrane proteins (OMPs) in pathogenesis through the communication between the leptospira and host (5).

Highly immunogenic OMPs such as *Leptospira* subsurface lipoproteins (LipL32, LipL41), trans-membrane outer membrane protein L1 (OmpL1), *leptospira* immunoglobulin-like protein (LigA, LigB, and LigC), and leptospira OmpA-like protein (Loa22) as the main skeleton of recombinant vaccines have played a major role in the development of new vaccines. Nevertheless, variation in serovar distribution in different regions hinders the development of effective leptospira vaccines (4-6).

Structurally, LigA contains 13 domains of the bacterial immunoglobulin (Big) -like fold, and numerous studies support its role in virulence. This protein promotes interactions with a range of specific host proteins that mediate adhesion of *L. interrogans* to the extracellular matrix, inhibit hemostasis, and inactivate key complement proteins to establish a systemic infection (7).

The development of vaccines against bacterial pathogens can be complex and time-consuming, and there is always a risk of unforeseen complications or side effects. In some studies, in silico approaches were used to identify potential epitopes on the LigA protein that could be used to develop vaccines or diagnostic tools for leptospirosis. Moreover, by providing targeted protection against specific serovars of Leptospira, the vaccine can reduce the incidence of leptospirosis and improve public health outcomes.

In terms of successful vaccine production, recombinant-protein production and purification present challenges, particularly in terms of cost effectiveness, expression yield, and stability. On the other hand, Lig proteins have been identified as suitable markers for the serodiagnosis of acute leptospirosis; thus, they could be implemented in immunodiagnostic kits to address leptospirosis

(8). Accordingly, this study aimed to design and produce a recombinant LigA antigen in the prokaryotic system to develop a recombinant leptospirosis vaccine.

2. Materials and Methods

This study was conducted from March 2021 to December 2021 at Razi Serum and Vaccine Research Institute (Karaj, Iran).

2.1 Recombinant LigA (r-LigA) immunogenic epitope sequence and properties prediction and validations

Previously designed conserved amino acid regions of LigA that stimulate the immune system were used in this study (unpublished data). The sequence of the selected r- LigA protein consisted of amino acid positions

852-937 (ligA9: BID_2), 943-1029 (ligA10: Big_2), 1034-1119 (ligA11: Big_2), and 1125-1210 (ligA12: Big_2).

The three-dimensional structure of the r-LigA protein was predicted by submitting the recombinant protein sequence in I-TASSER (Iterative Threading ASSEmbly Refinement) (https://zhanglab.ccmb.med.umich.edu/I-TASSER/) (9). Invasin protein of *Yersinia pseudotuberculosis* is a well-characterized protein which belongs to immunoglobulin superfamily in structure and topology, and its Big domain shares common themes to those observed in LigA (10). Hence, during threading for the conformational annotation Invasin of *Yersinia pseudotuberculosis* protein (PDB ID: 1CWV) was used as reference template for protein homology modeling (11). The structure with the minimum score in the previous step was selected and downloaded in protein database (pdb) compatible file format, visualized by Pymol, and used for further analysis. The prediction model quality was assessed using ProSA (https://prosa.services.came.sbg.ac.at). The Ramachandran's plot was predicted using Molprobity webserver (http://molprobity.biochem.duke.edu/).

2.2 Linear and conformational B-Cell epitopes presence in r-LigA prediction

To predict the linear (continuous) B-cell-epitopes present on r-LigA using the Ellipro module on the IEDB server, the threshold was adjusted to 0.5. Conformational (discontinuous) B-cell epitopes were predicted by another module, DiscoTope 3.0 (https://services.healthtech.dtu.dk/services/DiscoTope-3.0/), while the input structure type set as default (Solved structure) and Calibrated score set as default (Moderate confidence;0.90, recall up to ~50 %) (11).

2.3 MHC class I and II peptides T-Cell epitopes presence in r-LigA prediction

T-cell epitopes were evaluated by their binding capacity to major histocompatibility complex (MHC) I and II peptides on the http://tools.iedb.org/ server, and their locations were mapped on structural models. The NetMHCpan 4.1 EL (recommended epitope predictor- 2023.09) method was used for MHC-I peptides. The human-specific alleles HLA-A01:01, HLA-A02:03, HLA-A03:01, HLA-A11:01, HLA-A23:01, HLA-

B35:01, and HLA-B44:03 with a peptide length of 10 were selected (12). MHC-II peptides were predicted using the NetMHCIIpan 4.1 EL (recommended epitope predictor- 2023.09) method. Human HLA-DR was selected for prediction with alleles DRB1*01:01, DQA1*01:01/DQB1*02:01, and DPA1*01:03/DPB1*04:01. A peptide length of 15 was selected (11).

2.4 Cloning and expression of r-LigA

r-LigA sequence synthesized between EcoRI and BamHI restriction sites of the pET41a^{+} plasmid (carrying glutathione-S-transferase (GST-Tag), 6xHis, thrombin site, S-tag, enterokinase cleavage site and 8x*His tag* coding sequence as a *fusion partner* beside having the kanamycin resistance gene) by General Biosystems (Anhui, USA). r-LigA – pET41a^{+} cloning vector transformed into the *Escherichia coli*Top10-DH5α competent cells prepared by the conventional $MgCl_2$ protocol. Transformations were performed using either standard 42 $^{\circ}C$ heat shock protocol at 42°C and cultured on 2xYT agar (Merck, Germany) with the addition of kanamycin (50 μg mL^{-1}, K1377 sigma).

Random colonies were first screened by PCR using reverse T7 universal (Gene Link, Cat# 26-3000-13) and forward T7 universal (Gene Link, Cat# 26-3000-13) primers. Bacterial amplified plasmids were extracted using BioFact™ Plasmid Mini Prep Kit (Cat# PM105-100, South Korea) according to the manufacturer's instructions, followed by restriction enzyme digestion with EcoRI and BamHI (Fermentas Life Sciences, Thermo Fischer Scientific, Waltham, Massachusetts, USA) performed to confirm the presence of the insert r-LigA gene in the pET41a^{+}-LigA cassette before downstream transformation. To achieve enhanced and better expression, extracted plasmids (500 ng) were transformed into *E. coli* BL21 Star™ (DE3) competent cells were screened using the heat shock method described above.

The induction-expression analysis of recombinant *E. coli* was conducted in the presence of 0.5 mM IPTG in 2xYT broth at 30$^{\circ}C$ for 16 hours with constant agitation. *E. coli* BL21 Star™ (DE3) without the plasmid was used as a control strain.

2.5 Purification of r-LigA and Blotting analysis of the r-LigA protein

Transformed cells were harvested by centrifugation (5500 rpm for 5 min at 4°C), and the resulting bacterial pellet was re-suspended in 10 mL of Phosphate Buffered Saline (PBS 1X), containing 10 µL of Phenylmethanesulfonyl Fluoride (PMSF), and lysed with the aid of a sonicator (UP200St, Hielscher Ultrasonics, Germany). Sonication was conducted 12 times by 1 min pulses with 1 min intervals in an ice bath. The samples were then centrifuged at 15000 rpm for 15 min at 4°C, supernatants and pellets were collected separately and then subjected to SDS-PAGE to compare their solubility and yield.

As r-LigA was expressed in the form of inclusion bodies, the protein was further purified under denaturation conditions using serial concentrations of 1–8 M urea solutions in PBS buffer at 37°C. To this end, the pellets were suspended in 5 mL of 1 M urea solution and shaken for 1 h, then it was centrifuged (5000rpm, 10 min), and supernatant discarded. This step was repeated with 2, 3, 4, and 5 M urea solutions. For 6–8 M urea, the incubation time was increased overnight, and the supernatant was collected. In addition, after each step, the efficiency of the purification was assessed using 10% SDS-PAGE. Fractions containing the recombinant proteins were mixed and extensively dialyzed (MWCO 100 kDa, Thermo Fisher Scientific, MA) against PBS (1:1000) overnight at 4°C, for 24 h. The Protein concentration was quantified using the Bradford assay (13).

Purified r-LigA protein (approximately 200 ng) was used to charge the polyvinylidene fluoride membrane (PVDF) (Hybond-P, Amersham Biosciences, UK) and allowed to dry. The membranes were then blocked with 5% BSA in Tris-Buffered Saline-Tween 20 (TBST) for 40 min at RT. Consequently, the membrane was washed in 1X TBST three times for 10 minutes, each with gentle rocking. The proteins were reacted with conjugated anti-His tags antibodies (Thermo Fisher Scientific, USA) at RT for 1.5 hr. Finally, after extensive washing with TBST once for 15 min and twice for 5 min, the reaction was visualized using 5-thio-2-nitrobenzoic acid (TNB).

3. Results

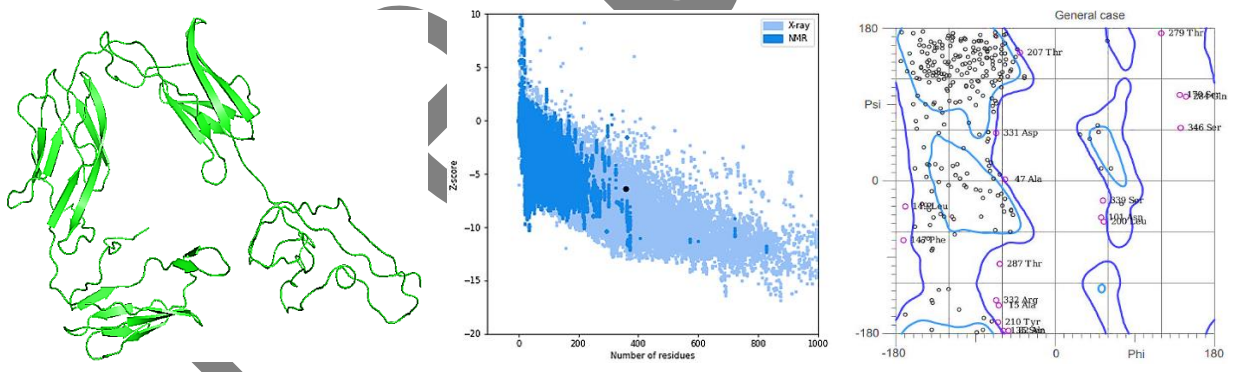
3.1 Recombinant LigA (r-LigA) immunogenic epitope structure and properties prediction and validations

The selected sequence of r-LigA consists of 359 amino acids (approximately 38 KDa). The 3D structure of the protein was predicted using the I-TASSER server, which are represented in Table 1.

Table 1. Properties of modeled protein based on the ITASSER results

Parameters	Values
Number of amino acids	359
Molecular weight	37897.12= ˷38 kDa
C-score	-0.62
Estimated TM-score	0.63±0.13
Estimated RMSD	8.0±4.4Å

The predicted three-dimensional structure of the r-LigA protein visualized using Chimera software (v1.13.1, University of California, San Francisco, CA, USA) is depicted in Figure 1A. The majority of the amino acid residues present in the protein may adopt β-sheet conformation. The overall quality of the predicted structure of r-LigA using ProSA had a z-score of −6.37 (Figure 1B), which is well within the range of native conformations, indicating that the predicted structure is reliable. According to MolProbity scores, 92.2% (329/357) of all residues were Ramachandran-favored (i.e., without any steric clashes) (Figure 1C).



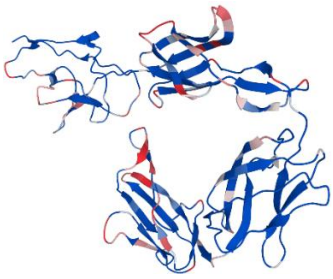
(A) **(B)** **(C)**

Figure 1. Prediction and validation results of r-LigA protein. (A) Predicted 3D structure of r-LigA protein visualized by Pymol software. (B) Overall model quality of the predicted structure given in a graphical form where X & Y axes represents residue position & predicted

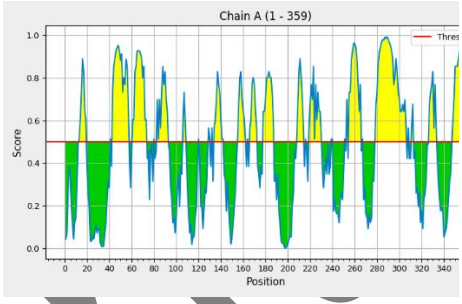
3.2 Linear and conformational B-Cell epitopes presence in r-LigA prediction

The graphical representation of amino acid scores as propensities for involvement in linear (continuous) and conformational (discontinuous) B-cell epitopes is presented in Figure 2. According to the DiscoTope

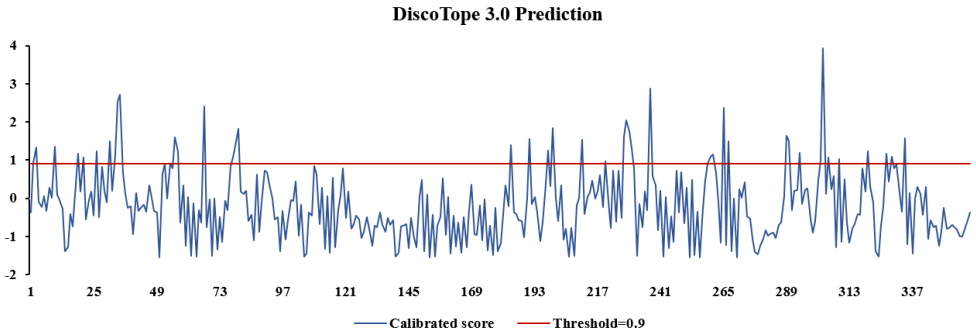
calibrated score with the default threshold (0.9), only 44 amino acids tended to be functional discontinuous B-cell epitopes (Figure 2A and 2B). According to the Bepipred Linear Epitope Prediction 2.0 module; there were 175 amino acid sequences above the threshold value (0.5) that are highlighted in yellow and considered continuous B-cell epitope residues (Figure 2C).



(A)



(C)



(B)

Figure 2. (A) Continuous epitope by DiscoTope: Residues with higher epitope propensity are colored in a deeper red, while residues with lower epitope propensity are colored in a deeper blue. (B) Residues interacting as continuous epitope calculated by DiscoTope (C) Linear epitope: Amino acid residue with a score above the threshold Bepipred Linear Epitope Prediction is considered to be part of an epitope & coloured in yellow.

The ElliPro prediction server identified 13 linear and three conformational potential peptide sequences as B-cell-epitopes (Figures 3 and 4). The linear B-cell epitopes range in length from 5 to 31 amino acids, whereas the three potential conformational B-cell epitopes comprise 57 and 87 amino acids, respectively.

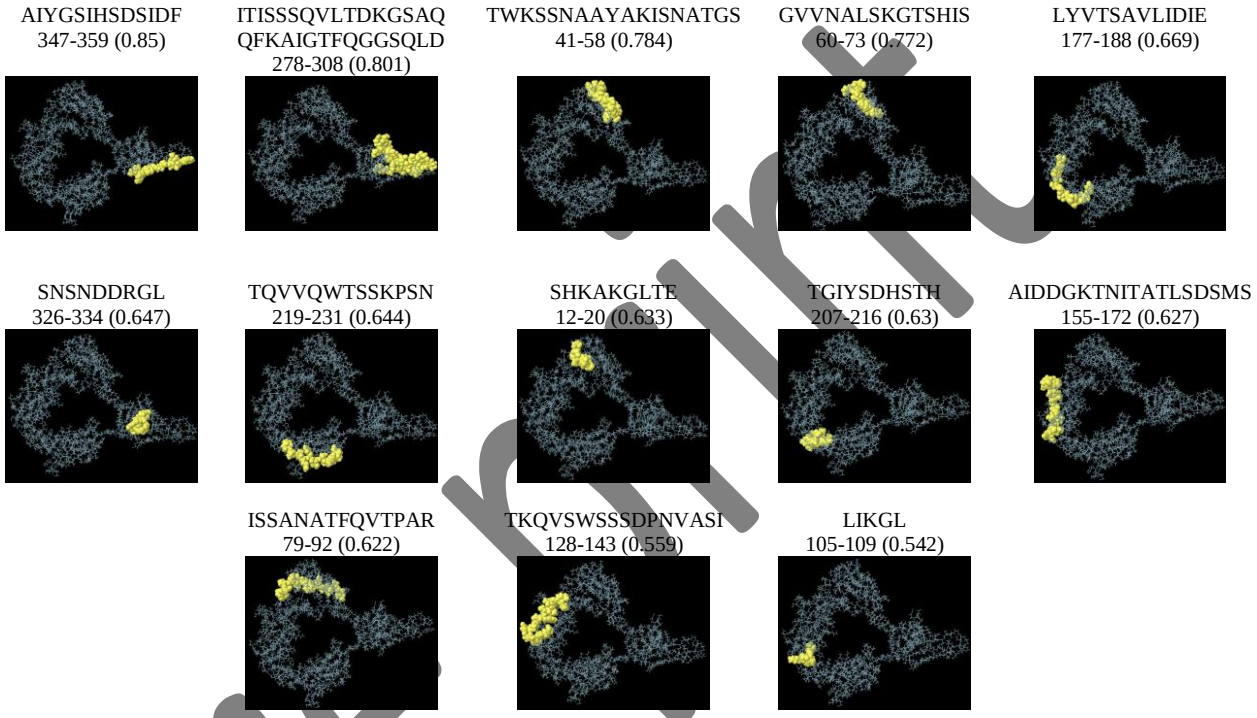


Figure 3. Linear B cell epitopes predicted using Ellipro prediction server. Yellow highlight regions capable to stimulate B-cells. Peptide sequences, residue positions and predicted scores (in parenthesis) are given are given above each pictures.

4.	Residues	No of residues	Score	3-D structure

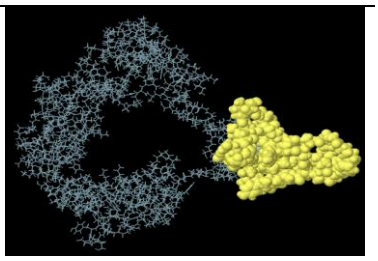
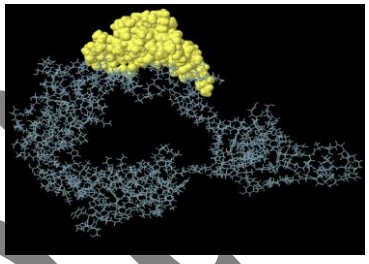
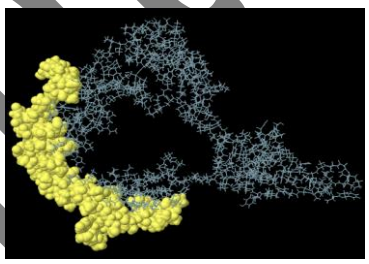
I278, T279, I280, S281, S282, S283, Q284, V285, L286, T287, D288, K289, G290, S291, A292, Q293, Q294, F295, K296, A297, I298, G299, T300, F301, Q302, G303, G304, S305, Q306, L307, D308, D311, L312, V313, S326, N327, S328, N329, D330, D331, R332, G333, L334, T335, A347, I348, Y349, G350, S351, I352, H353, S354, D355, S356, I357, D358, F359	57	0.766	
S11, S12, H13, K14, A15, K16, G17, L18, T19, E20, T41, K43, S44, S45, N46, A47, A48, Y49, A50, K51, I52, S53, N54, A55, T56, G57, S58, G60, V61, V62, N63, A64, L65, S66, K67, G68, T69, S70, H71, I72, S73, I79, S80, S81, A82, N83, A84, T85, F86, Q87, V88, T89, P90, A91, R92, T120, D121	57	0.704	
F104, L105, I106, K107, G108, L109, Q130, V131, S132, S134, S135, S136, D137, P138, N139, V140, A141, S142, I143, A155, I156, D157, D158, G159, K160, T161, N162, T164, T166, L167, S168, D169, S170, M171, S172, L177, Y178, V179, T180, S181, A182, V183, L184, I185, D186, I187, V189, A206, T207, G208, I209, Y210, S211, D212, H213, S214, H216, T219, Q220, V221, V222, Q223, W224, T225, S226, S227, K228, P229, S230, N231, G239, K240, A248, D251, E253, T255, A256, T257, Y258, D259, S260, I261, K262, S263, N264, R265, W267	87	0.647	

Figure 4. Conformational B cell epitopes predicted using Ellipro prediction server. Yellow highlight regions capable to stimulate B cells.

4.1 MHC class I and II peptides T-Cell epitopes presence in r-LigA prediction

Based on the IEDB analysis prediction results, there were various amino acid residues in the shape of the MHC-I and MHC-II grooves, and their most efficient peptides are presented with their scores in Tables 2 and 3, respectively.

Table 2. MHC-I binding peptides predicted by the NetMHCpan EL 4.1.

No	Peptide Sequence	Allele	Position	Score
1	IVLNPTSSHK	HLA-A*03:01	5-14	0.892
2	ATYDSIKSNR	HLA-A*11:01	247-256	0.877
3	IVLNPTSSHK	HLA-A*11:01	5-14	0.758
4	FGDSEFTATY	HLA-A*01:01	240-249	0.747
5	SIEVTPNNFF	HLA-B*44:03	90-99	0.739
6	ATYDSIKSNR	HLA-A*03:01	247-256	0.738
7	TTALSVGSSK	HLA-A*11:01	321-330	0.732
8	VLSEGLTLQL	HLA-A*02:03	186-195	0.713
9	DSMSASTTLY	HLA-A*01:01	164-173	0.689
10	LTDKGSAQQF	HLA-A*01:01	272-261	0.892

Table 3. MHC-II binding peptides predicted by the IEDB recommended 2.22 method.

No	Peptide Sequence	Allele	Position	Adjusted Rank *	Method Used
1	KSNRAWIFVNDEKLV	HLA-DQA1*01:01/DQB1*02:01	262-276	0.77	NetMHCIIpan
2	NRAWIFVNDEKLVNI	HLA-DQA1*01:01/DQB1*02:01	264-278	0.98	NetMHCIIpan
3	SNRAWIFVNDEKLVN	HLA-DQA1*01:01/DQB1*02:01	263-277	1.1	NetMHCIIpan
4	TTLYVTSAVLIDIEV	HLA-DQA1*01:01/DQB1*02:01	175-189	1.1	NetMHCIIpan
5	DNTFSLAGSATAIDD	HLA-DRB1*01:01	144-158	1.3	Consensus (comb.lib./smm/nn)
6	IDNTFSLAGSATAID	HLA-DRB1*01:01	143-157	1.3	Consensus (comb.lib./smm/nn)
7	NTFSLAGSATAIDDG	HLA-DRB1*01:01	145-159	1.3	Consensus (comb.lib./smm/nn)
8	TLYVTSAVLIDIEVK	HLA-DQA1*01:01/DQB1*02:01	176-190	1.4	NetMHCIIpan
9	IKSNRAWIFVNDEKL	HLA-DQA1*01:01/DQB1*02:01	261-275	1.6	NetMHCIIpan
10	SIDNTFSLAGSATAI	HLA-DRB1*01:01	142-156	1.8	Consensus (comb.lib./smm/nn)

*Lower adjusted rank indicates a better binding.

4.2 Cloning and expression of r-LigA protein

Transformation of the r-LigA pET41a^{+} plasmid into the *E. coli* Top10-DH5α (cloning vector) and *E. coli* BL21 Star^{TM} (DE3) (expression vector) produced several colonies in the presence of ampicillin. This was confirmed by amplification of a 2000-bp segment in PCR of the target gene using T7 universal primers on multiple random clones (Figure 5A). As shown in Figure 5B, EcoRI and BamHI digestion of the purified r-LigA- pET41a^{+} plasmid yielded 5.9-kb and 1.1-kb fragments that can be attributed to the plasmid and *r-LigA*, respectively.

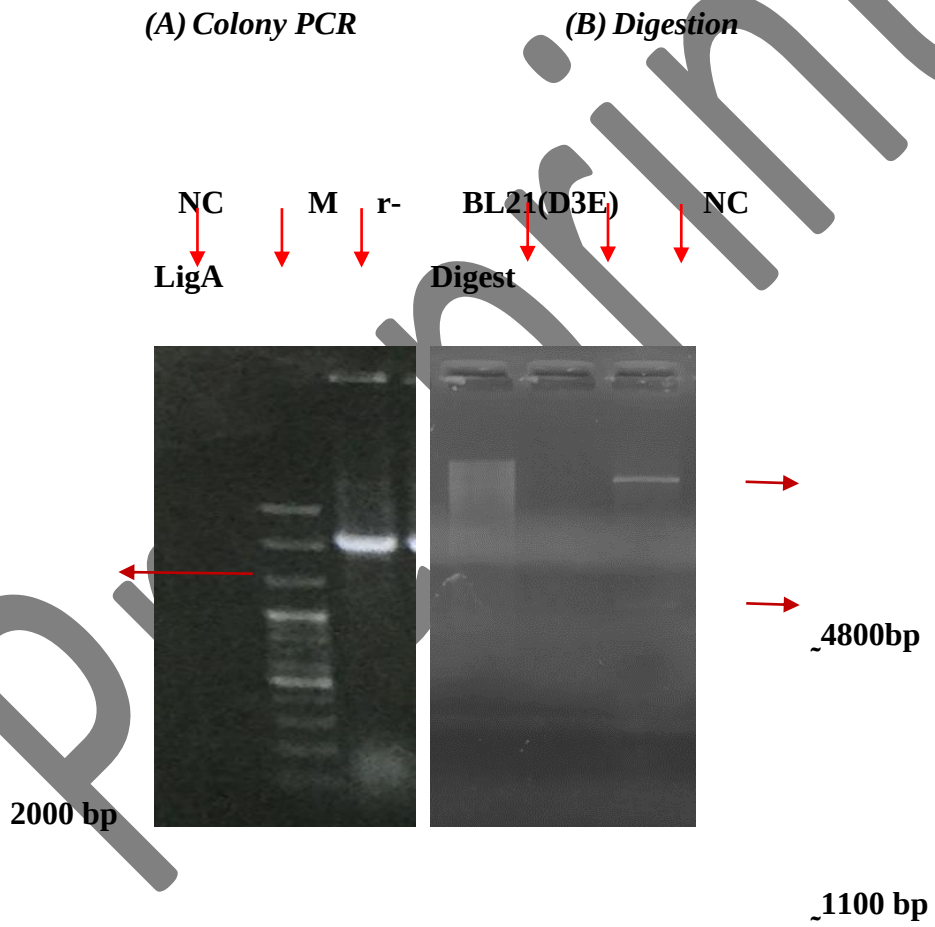


Figure 5. Transformation confirmation electrophoresis on 1% agarose gel. (M: Marker, NC: Negative control)

To achieve high-level expression of the peptide plasmid transformed into the *E. coli* BL21 Star^{TM} (DE3) induced with 0.5 mM IPTG at 30°C for 16 hours which led to the expression of recombinant r-LigA protein

with approximately 71 KDa confirmed on SDS-PAGE. *E. coli* BL21 Star™ (DE3) without the plasmid was subjected to the same experiment and did not express the protein under any conditions.

4.3 Purification and Blotting analysis of the r-LigA protein

Analysis of the cell lysate post-sonication revealed that the r-LigA r protein expressed using the pET41a+ vector in the E. coli BL21 system mostly accumulates as an insoluble form in the pellet, probably in inclusion bodies. The molecular weight of r-LigA is approximately 38 KDa, whereas the different fused partner proteins of the pET41a+ plasmid are approximately 32.29 KDa. Thus it is speculated that the molecular weight of the expressed protein of this construct to be approximately 71 KDa. There was an obvious band at 71 KDa in the purified protein after washing in serial concentrations of urea in PBS (1 to 5M) followed by solubilization (6 to 8M) at 37 °C. The outcome of this procedure on SDS-PAGE revealed that r-LigA was successfully induced. Expressed and purified. The protein content gradually increased and then decreased. The highest level of r-LigA gained in 6-M urea (Figure 6A).

To ensure high confidence, the Dot blot-binding assay was performed, and its result confirmed the interaction of r-LigA with conjugated anti-His tags antibodies (Figure 6B).

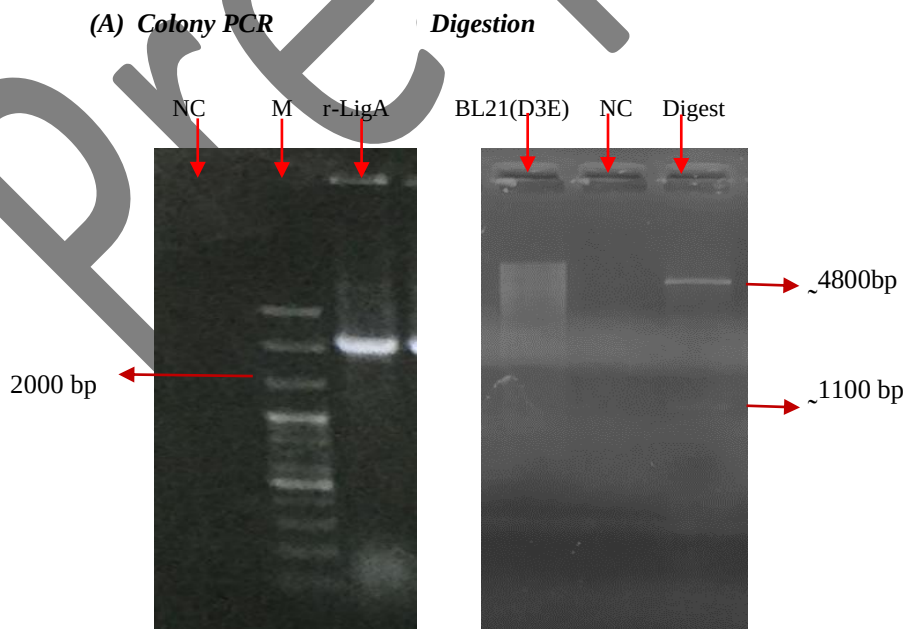


Figure 5. Transformation confirmation electrophoresis on 1% agarose gel. (M: Marker, NC: Negative control)

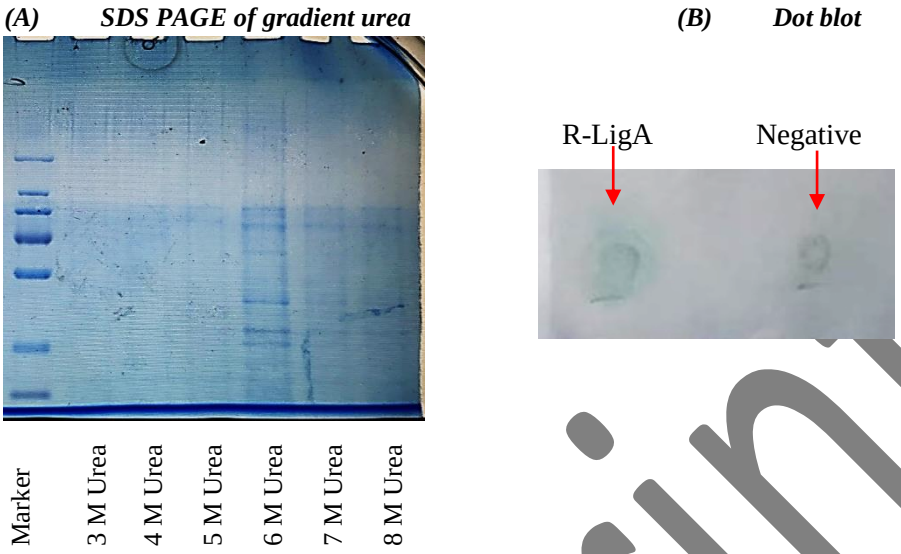


Figure 6. SDS-PAGE of expressed r-LigA in *E. coli* (BL21 Star™ (DE3)) under different concentration of Urea after dialyze (Marker: PM2500)

The quantity of the yield of purified recombinant protein was estimated to be 834 µg/mL in Bradford assay.

5. Discussion

The global distribution of leptospirosis and its association with autoimmune disease have motivated researchers to develop effective and safe vaccines. Commercial Leptospiral vaccines are available in many countries for use in dogs, pigs, and livestock, most of which have limited efficacy because of a lack of serovar specificity and their adverse effects, such as pain and fever (14, 15).

In fact, most surface proteins of pathogens are involved in virulence and even host immune response stimulation, making them good targets for vaccine development (2). Matsunaga *et al.* (2003) expressed surface-exposed proteins belonging to the bacterial immunoglobulin superfamily by pathogenic Leptospira species and reported a new family of Big domain proteins called immunoglobulin-like proteins (Lig) in

pathogenic Leptospira (10). LigA, as a member of the Leptospiral outer membrane proteins, possess highly conserved domains, which make it a proper candidate in antigen-based vaccine development (2, 16).

Kozumi *et al.* determined that recombinant full-length LigA and truncated LigB proteins can induce immunoprotection against Leptospiral lethal infection in C3H/HeJ mice (17). Silva *et al.* expressed carboxy-terminal region unique to the LigA protein (LigANI), LigBNI, and LigBrep from *Leptospira interrogans* serovar Copenhageni in *E. coli* BL21(DE3)Star as recombinant proteins and in the hamster infection model found that their potent activity of immunoprotection against mortality (18).

As it has been suggested that the presence of at least three Big domains (especially ligA11, and ligA12 specific Big) is critical, a construct of r-LigA protein carrying ligA9, 10, 11, and 12 domains inserted in pET41a^{+} plasmid transformed into the *E .coli* Top10-DH5α and *E. coli* Star^{TM} (DE3) cells in this study (19).

I-TASSER is an automated for protein tertiary structure prediction in which uses PSSpred for second structure prediction, LOMETS for template detection, replica-exchange Monte Carlo simulation for fragment structure assembly, SPICKER for model selection by clustering structure decoys, fragment-guided molecular dynamics simulation (FG-MD) or ModRefiner for atomic-level structure refinement and COACH for structure-based biology function annotation (9, 20). The I-TASSER server calculates the confidence score (C-score), which estimates the quality of the predicted models based on the significance of threading template alignments and the convergence parameters of the structure assembly simulations, revealing high confidence in prediction. Normally, C-score values lie in between [−5 to 2], where a value denotes a model with a higher confidence and correct topology (9). Then, a C-score value of -0.62 indicates good topology of the modeled structure. The topological similarity of protein structures is reported as TM-score in I-TASSER. TM-score <0.17 indicates a random similarity, while TM-score > 0.5 indicates a model of correct topology (9). Based on these results, the TM-score 0.63±0.13 confirms a proper topology.

As a structural validation, the statistical distribution of the combinations of the backbone dihedral angles ϕ and ψ of a protein can be determine using the Ramachandran plot. In good-quality homology models, very

few dihedral angles are found in forbidden regions (11). In this study, the three-dimensional structure predicted for the r-LigA protein was reliable because only 7.8% of the residues were Ramachandran outliers.

Epitopes are small specific segments of antigens that usually affect the specificity of cellular and humoral immune responses. In linear epitopes, the primary amino acid sequence is recognized by the antibodies, whereas in conformational epitopes, amino acids are not neighboring and are brought into close proximity in the folded protein and directly bind to a receptor of the immune system(21, 22). Interestingly, the linear epitopes with high scores as follow AIYGSIHSDSIDF (347-359), ITISSSQVLTDKGSAQQFKAIGTFQGGSQLD (278-308), and SNSNDDRGL (326-334) composing a conformational epitope with the highest score as well.

Binding to MHC class I molecules (MHC-I) plays a pivotal role in antigen presentation and induction of cytotoxic T-cell responses. MHC class I and II genes in humans are called "*H*uman *L*eukocyte *A*ntigen" (HLA). NetMHCpan 4.1 uses artificial neural networks (ANNs) and has been trained on a combination of more than 850,000 quantitative binding affinity (BA) and mass spectrometry, Eluted Ligands (EL) peptides that make it as one of the most reliable prediction servers(11, 23). Limited amount of data available for most MHC class II molecules and considerable differences in sequence polymorphism and corresponding details in the molecular structures across the different MHC class II loci hinder the development of cross-species training strategies and is known as main limitation of pan-specific or cross species approach for MHC class II prediction (11). The results for MHC-I binding peptides showed that epitope binding with HLA-A*03:01allele had the highest binding affinity scores at 0.892.

Behera et al. applied similar approaches and found that, three out of the most efficient peptides for MHC-I grooves predicted by NetMHCpan 4.1 server were presented by MHC Class II molecules, indicating that both CMI and humoral immune response can be induced by those peptides (11).

The NetMHCIIpan 4.0 server was used to determine Helper T lymphocyte (HTL) epitopes (23). The Number of epitope binding with the HLA-DQA1*01:01/DQB1*02:01 allele was higher than other alleles

in MHC-II binding peptides. Interestingly, NRAWIFVNDEKLV core sequence binding with HLA-DQA1*01:01/DQB1*02:01 allele was present in top three NetMHCIIpan 4.0 prediction server, followed by the NTFSLAGSATAIDD sequence binding with HLA-DRB1*01:01 in Consensus method.

For subunit vaccines, several expression systems were used, including bacteria, yeast, mammalian or plant cells (24). In this study, the r-LigA construct transferred into *E. coli* Top10-DH5α and *E. coli* BL21 Star™ (DE3). r-LigA was expressed in *E. coli* BL21 Star™ (DE3), it was observed that it is mostly in the form of inclusion bodies. It has been reported that LigA and LigB normally contain a lipobox, whereas LigA encodes a lipoprotein signal peptide distinct from the motifs of *E. coli* and other gram-negative bacteria. Due to the presence of the lipobox, there is a *sec*-dependent export pathway for LigA exists across the cytoplasmic membrane. Expression of LigA in *leptospira sp*. results in a surface protein with a signal peptide subject to proteolytic removal of its lipid anchor. Therefore, a portion of LigA protein can be recovered from the culture supernatant in the normal system(25).

Palaniappan et al. demonstrated that intact ligA without its signal sequence expression in *E. coli* was toxic to *E. coli*, whereas the expression of a 90-kDa truncated LigA was not toxic to E. coli cells. They found that complete LigA protein can be expressed only in leptospira-infected hamster kidney (26). Interestingly, in a study LigA lipoproteins were expressed and exposed on the surface of the saprophyte *L. biflexa* cells and suggested that *L. biflexa* is an appropriate surrogate host for the expression of at least some *L. interrogans* outer membrane proteins (27).

Due to differences in leptospiral lipobox sequences, it is anticipated that leptospiral lipoproteins to not be processed correctly in *E. coli* (27). To minimize protein inclusion bodies formation, it is recommended to take some strategies such as changing vectors, host strains, production of endogenous chaperones or /chaperone co-expression, low temperature induction, and using the target protein to a soluble protein or peptide tags (28).

High levels of r-ligA (834 µg/mL) were purified from inclusion bodies with high efficiency by serial washing with 6–8 molar urea. It can be concluded that r-LigA protein could be recovered from inclusion bodies, using the urea method, which is a simple and low-cost technique. Hartwig *et al.*, (2010) expressed rLigANI (61 kDa) in the eukaryotic expression system of *Pichia pastoris*, resulting in a significantly lower protein yield (24).

In this study, the r-LigA construct produced an approximately 38-kDa recombinant protein, and high yield of protein obtained at 6m urea solution. Urea purification method is a low-cost purification strategy in comparison to Ni-NTA affinity chromatography purification. The optimum urea concentration in purification procedure of this study were in accordance with a study focused on Lipl41 recombinant protein expression in *E. coli* BL21 (DE3) carrying pET32a+ expression vector (13)

r-LigA protein production was confirmed by immunoblotting analysis with HRP-conjugated Anti-6x His-Tag antibodies. The performed dot-blot technique showed promising results in discriminating between positive and negative serum samples.

Although expression of r-LigA was successful, due to the limited commercial availability of immunological reagents for use in hamsters, we were not able to determine the cell-mediated immune response of hamsters to r-LigA protein immunizations. Although this study focused on a comprehensive mmunoinformatics, it have not yet been assessed as vaccine candidates; and hence, could be worthy of further investigation as novel vaccine candidates.

In conclusion, these findings suggest this immunoinformatics study represents novel vaccine candidates that will further aid in the development of improved vaccines for leptospirosis. However, further refinement of this technique is required before it can be used for leptospirosiss diagnosis or vaccine development.

Acknowledgment

The authors would like to thank the Razi Vaccine and Serum Research Institute, Agricultural Research, Education and Extension Organization (AREEO), Karaj, Iran, for providing laboratory space, facilities and funding from the project source No. 12-18-18-106-96045-961025 on developing Leptospira vaccines.

Authors' Contribution

Conceptualization, Methodology, and Supervision: P. KH.

Formal analysis and investigation: All authors.

Drafting of the manuscript: A. C.

Writing - review and editing: All authors.

Funding acquisition: P. KH. And S. MB.

All authors checked and approved the final version of the manuscript.

Ethics

All the procedures for animal experiments were approved by the Research Ethics Committees of Islamic Azad University- North Tehran Branch, under IR.IAU.TNB.REC.1401.049 Approval ID and performed in accordance to the ethical principles and the national norms and standards for conducting Medical Research in Iran guidelines.

Conflict of Interest

The authors declare that they have no known conflicts of interest or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author upon reasonable request.

References

1. Breda LC, Hsieh C-L, Castiblanco Valencia MM, da Silva LB, Barbosa AS, Blom AM, et al. Fine mapping of the interaction between C4b-binding protein and outer membrane proteins LigA and LigB of pathogenic *Leptospira interrogans*. PLoS neglected tropical diseases. 2015;9(10):e0004192.
2. Bashiru G, Bahaman ARJIJoMR. Advances & challenges in leptospiral vaccine development. 2018;147(1):15-22.
3. Dellagostin OA, Grassmann AA, Hartwig DD, Félix SR, da Silva EF, McBride AJ. Recombinant vaccines against leptospirosis. Human vaccines. 2011;7(11):1215-24.
4. Koizumi N, Watanabe H. Leptospirosis vaccines: past, present, and future. Journal of postgraduate medicine. 2005;51(3):210-4.
5. Haake DA, Zückert WRJL, Leptospirosis. The leptospiral outer membrane. 2015:187-221.
6. Barazzone GC, Teixeira AF, Azevedo BO, Damiano DK, Oliveira MP, Nascimento AL, et al. Revisiting the development of vaccines against pathogenic *Leptospira*: innovative approaches, present challenges, and future perspectives. Frontiers in Immunology. 2022;12:760291.
7. Haake DA, Matsunaga J. Leptospiral immunoglobulin-like domain proteins: Roles in virulence and immunity. Frontiers in Immunology. 2021;11:579907.
8. Croda J, Ramos JG, Matsunaga J, Queiroz A, Homma A, Riley LW, et al. Leptospira immunoglobulin-like proteins as a serodiagnostic marker for acute leptospirosis. 2007;45(5):1528-34.
9. Zhang YJBb. I-TASSER server for protein 3D structure prediction. 2008;9:1-8.
10. Matsunaga J, Barocchi MA, Croda J, Young TA, Sanchez Y, Siqueira I, et al. Pathogenic Leptospira species express surface-exposed proteins belonging to the bacterial immunoglobulin superfamily. 2003;49(4):929-46.
11. Behera SK, Sabarinath T, Mishra PKK, Deneke Y, Kumar A, ChandraSekar S, et al. Immunoinformatic study of recombinant liga/bcon1-5 antigen and evaluation of its diagnostic potential in primary and secondary binding tests for serodiagnosis of porcine leptospirosis. 2021;10(9):1082.
12. Andreatta M, Nielsen MJB. Gapped sequence alignment using artificial neural networks: application to the MHC class I system. 2016;32(4):511-7.
13. Golab N, Khaki P, Harzandi N, Esmaelizad M, Tebianian M. Expression and purification of the Lipl41, a surface-exposed lipoprotein antigen of pathogenic *Leptospira* spp. Veterinary Archives. 2020; 90((3)):297-305.
14. Srivastava SJIJoMM. Supplement-Prospects of developing leptospiral vaccines for animals. 2006;24(4):331-6.
15. Techawiwattanaboon T, Patarakul KJAB. Update on molecular diagnosis of human leptospirosis. 2019;13(6):207-16.
16. Varma VP, Kadivella M, Kavela S, Faisal SM. Leptospira Lipid A Is a Potent Adjuvant That Induces Sterilizing Immunity against Leptospirosis. Vaccines. 2023;11(12):1824.
17. Koizumi N, Watanabe H. Leptospiral immunoglobulin-like proteins elicit protective immunity. Vaccine. 2004;22 (11-12):1545-52.
18. Silva EF, Medeiros MA, McBride AJ, Matsunaga J, Esteves GS, Ramos JG, et al. The terminal portion of leptospiral immunoglobulin-like protein LigA confers protective immunity against lethal infection in the hamster model of leptospirosis. Vaccine. 2007;25(33):6277-86.
19. McBride AJ, Cerqueira GM, Suchard MA, Moreira AN, Zuerner RL, Reis MG, et al. Genetic diversity of the Leptospiral immunoglobulin-like (Lig) genes in pathogenic *Leptospira spp*. Infection, Genetics and Evolution. 2009;9(2):196-205.

20. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang YJNm. The I-TASSER Suite: protein structure and function prediction. 2015;12(1):7-8.
21. Lin Xa, Sun A, Ruan P, Zhang Z, Yan JJBm. Characterization of conserved combined T and B cell epitopes in Leptospira interrogans major outer membrane proteins OmpL1 and LipL41. 2011;11:1-9.
22. Fookolaee SP, Talebishelimaki S, Rad MTS, Rokni MAJJoCBR. In silico prediction of B cell epitopes of the hemolysis-associated protein 1 for vaccine design against leptospirosis. 2020;1(1):32-7.
23. Kapoor K, Patel A, Shekh S, Joshi CG. Designing of Multiepitope Subunit Vaccine Against Leptospirosis Using Immunoinformatics Approach. 2022.
24. Hartwig DD, Oliveira TL, Seixas FK, Forster KM, Rizzi C, Hartleben CP, et al. High yield expression of leptospirosis vaccine candidates LigA and LipL32 in the methylotrophic yeast *Pichia pastoris*. Microbial cell factories. 2010;9:1-7.
25. Matsunaga J, Sanchez Y, Xu X, Haake DAJI, immunity. Osmolarity, a key environmental signal controlling expression of leptospiral proteins LigA and LigB and the extracellular release of LigA. 2005;73(1):70-8.
26. Palaniappan RU, Chang Y-F, Jusuf S, Artiushin S, Timoney JF, McDonough SP, et al. Cloning and molecular characterization of an immunogenic LigA protein of *Leptospira interrogans*. Infection immunity. 2002;70(11):5924-30.
27. Figueira CP, Croda J, Choy HA, Haake DA, Reis MG, Ko AI, et al. Heterologous expression of pathogen-specific genes *ligA* and *ligB* in the saprophyte *Leptospira biflexa* confers enhanced adhesion to cultured cells and fibronectin. BMC microbiology. 2011;11:1-9.
28. Bhatwa A, Wang W, Hassan YI, Abraham N, Li X-Z, Zhou TJFiB, et al. Challenges associated with the formation of recombinant protein inclusion bodies in Escherichia coli and strategies to address them for industrial applications. 2021;9:630551.