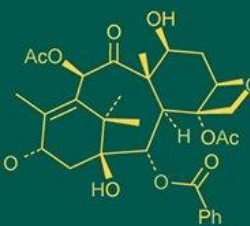
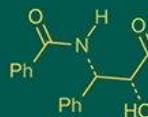


International Journal of Advanced Biochemistry Research



ISSN Print: 2617-4693
ISSN Online: 2617-4707
NAAS Rating (2026): 5.29
IJABR 2026; 10(8): 434-442
www.biochemjournal.com
Received: 02-05-2026
Accepted: 08-06-2026

D Jenita Rani
Ph.D. Scholar, Department of
Animal Biotechnology, MVC,
Chennai, Tamil Nadu, India

TMA Senthilkumar
Professor and In-charge, ZRL,
CAHS, Chennai, Tamil Nadu,
India

P Raja
Assistant Professor,
Department of Animal
Biotechnology, MVC, Chennai,
Tamil Nadu, India

K Manimaran
Professor and In-charge,
Vaccine Research Centre-Viral
Vaccine, CAHS, Chennai,
Tamil Nadu, India

M Parthiban
Professor and Head,
Department of Animal
Biotechnology, MVC, Chennai,
Tamil Nadu, India

R Mythili
Young Professional-I, ICAR
Scheme, ZRL, CAHS, Chennai,
Tamil Nadu, India

Corresponding Author:
D Jenita Rani
Ph.D. Scholar, Department of
Animal Biotechnology, MVC,
Chennai, Tamil Nadu, India

Immunoinformatics Approaches to Leptospiral Outer Membrane Proteins for Designing a Multi-Epitope Diagnostic Antigen

D Jenita Rani, TMA Senthilkumar, P Raja, K Manimaran, M Parthiban and R Mythili

DOI: <https://www.doi.org/10.33545/26174693.2026.v10.i8f.9577>

Abstract

Leptospirosis is a globally distributed, widespread, neglected, re-emerging zoonotic disease caused by pathogenic spirochetes of the genus *Leptospira*, affecting humans and animals. The diagnosis of leptospirosis remained challenging because conventional serological assays were affected by cross-reactivity, variable sensitivity and limited applicability to rapid screening in endemic settings. Therefore, improved immunodiagnostic approaches are required. *In silico* methods provide a valuable strategy for the development of a diagnostic antigen. In this study, an immunoinformatics-based approach was employed to identify and select conserved, antigenic linear B-cell epitopes from the outer membrane proteins (OMPs) of pathogenic *Leptospira* including LipL21, LipL32, LipL41, OmpL1 and Loa22. The predicted epitopes were selected and assembled along with suitable peptide linkers to design a synthetic gene construct for the production of a recombinant five-epitope leptospiral chimeric protein (rLep5ECP). The physicochemical properties of rLep5ECP were further analysed *in silico* to assess its preliminary suitability as a diagnostic antigen. The designed rLep5ECP provides a rational framework for developing a recombinant antigen-based immunodiagnostic assay for leptospirosis. Although rLep5ECP exhibited promising *in silico* characteristics, its diagnostic performance requires confirmation through experimental validation.

Keywords: Leptospirosis, immunoinformatics, reverse vaccinology, linear B-cell epitopes, multi-epitope antigen, immunodiagnostics

Introduction

Leptospirosis is an emerging neglected, zoonotic disease caused by pathogenic *Leptospira* spp. The estimated leptospirosis cases are approximately 1 million human cases and 60,000 deaths annually (Costa *et al.*, 2015) [1]. *Leptospira* sp. belongs to the family Leptospiraceae and order Spirochaetales. They are non spore forming, gram negative, tightly coiled and spiral-shaped bacteria. They are obligate aerobes and motile organisms. *Leptospira* sp. can be categorised into pathogenic and saprophytic groups. Serological classification allows the discrimination of more than 300 serovars that can be grouped into 30 serogroups based on the expression of surface exposed lipopolysaccharides (Amran - *et al.*, 2023) [2].

Leptospirae survive in moist environments for prolonged periods and are transmitted through direct contact with the urine of infected animals or indirectly through contaminated water, soil, food and fomites. Rodents are considered as the major reservoir hosts, while domestic animals such as dogs, cattle, pigs and horses act as incidental or maintenance hosts depending on the infecting serovar. Because of its zoonotic nature, significant economic and public health impact, leptospirosis remains a major concern under the One Health concept (Bharti *et al.*, 2003) [3].

The diagnosis of leptospirosis remains challenging due to limitations in currently available tests. Culture isolation is rarely used in routine diagnosis because leptospirae are slow-growing organisms that may require specialised media and prolonged incubation periods. This method is expensive, has low sensitivity and is labour-intensive, particularly after antibiotic therapy (Levett, 2001) [4]. The microscopic agglutination test (MAT) continues to be regarded as the gold standard, but it requires paired sera, specialised laboratories and interpretation challenges due to cross-reactivity. The MAT also shows low sensitivity during

early infection and cannot differentiate between infection-induced antibodies and vaccine-induced antibodies, especially in vaccinated dogs (Sykes *et al.*, 2011 and WOA, 2021) ^[5, 6]. These limitations have created the need for improved diagnostic approaches that are rapid, sensitive, specific, safe, economical and suitable for routine clinical use. Molecular techniques like polymerase chain reaction provide higher sensitivity for early detection of *Leptospiral* deoxyribonucleic acid (DNA) in samples, but limit the on-site resource availability and point-of-care diagnostics (Pinto *et al.*, 2022) ^[7]. Recently, the improved serological diagnosis of leptospirosis has been targeted with recombinant proteins (Flannery *et al.*, 2001) ^[8]. Recombinant protein-based serologic tests may achieve high sensitivity and specificity because of the high concentration of immunoreactive antigens. The desired antigen would be a principal target of the host immune response, expressed only in pathogenic *Leptospira* spp. and conserved among the more than 200 serovars (Faine *et al.*, 1999) ^[9]. Many rapid immunodiagnostic assays such as latex agglutination tests, enzyme-linked immunosorbent assay and lateral flow assays provide point-of-care options for leptospirosis detection, targeting antibodies against pathogenic *Leptospira* antigens. Synthetic biology has emerged as a promising approach for the production of recombinant antigens for diagnostic applications. It offers a safe alternative for the synthesis of antigenic components derived from organisms that pose a high biological risk (Yaakob *et al.*, 2017) ^[10]. Immunoinformatics, also known as computational immunology, integrates experimental immunology with computer science to generate novel insights into immune responses. By combining well-curated immunological databases with sophisticated computational prediction methods, it enables a deeper understanding of the molecular interactions that govern immune function (Lefranc, 2014) ^[11]. Emerging approaches to vaccine design and engineering are based on recent insights into structural biology, immunology and computational biology. These approaches are broadly termed as reverse vaccinology (Donati and Rappuoli, 2013) ^[12]. With the availability of whole-genome sequences, reverse vaccinology can be applied more comprehensively and identify potential antigens coded by a genome irrespective of their abundance, immunogenicity and phase of expression (Masignani *et al.*, 2002) ^[13]. Recent advances in bioinformatics have facilitated the engineering of an epitope-based serological assay for leptospirosis (Validi *et al.*, 2018) ^[14]. Hundreds of genes encoding surface-exposed lipoproteins and OMPs have been identified as candidates for the development of rapid diagnostics for leptospirosis. They are used for recombinant antigen development because they are highly conserved, strongly immunogenic and directly exposed to the host immune system during infection, making them readily recognised by host antibodies (Haake *et al.*, 2000) ^[15]. Outer membrane proteins play a vital role in the virulence mechanisms of pathogens, attachment to various substrates, importing nutrients, maintaining the bacterial cell structure, exporting bactericidal and toxic agents and they can also serve as putative diagnostic candidates (Palaniappan *et al.*, 2007) ^[16]. The advent of recombinant DNA technology has enabled the use of several OMPs such as *LipL32* (Sumarningsih *et al.*, 2016) ^[17], *LipL21* (Joseph *et al.*, 2012) ^[18], *OmpL1* (Natarajaseenivasan *et al.*, 2008) ^[19], *LipL41* (Senthilkumar *et al.*, 2007) ^[20] and *Loa22* (Ye *et al.*, 2014)

^[21] which are found in pathogenic leptospires. They serve as promising targets for molecular diagnostic assays. Recombinant OMPs overcome several limitations associated with whole-cell *Leptospira* antigen-based assays since they lack the non-specific moieties present in whole-cell preparations (Lin *et al.*, 2008) ^[22]. This enhances the specificity and reproducibility of diagnostic assays while minimising cross-reactivity. The B-cells recognise solvent-exposed antigens through antigen receptors called B-cell receptors, consisting of membrane-bound immunoglobulins. Upon activation, B-cells differentiate and secrete soluble forms of the immunoglobulins (antibodies) which mediate humoral adaptive immunity (Paul, 2012) ^[23]. A B-cell epitope is the antigen portion binding to the immunoglobulin or antibody. B-cell epitope prediction aims to identify immunogenic B-cell epitopes. B-cell epitopes are divided in two main groups such as linear and conformational. Linear or continuous B-cell epitopes consist of sequential residues or peptides, whereas conformational or discontinuous B-cell epitopes consist of patches of solvent-exposed atoms from residues that are not necessarily sequential. Antibodies recognising linear B-cell epitopes can recognise denatured antigens, while denaturing the antigen results in loss of recognition for conformational B-cell epitopes. Most B-cell epitopes are conformational (approximately 90%) and only a few of the native antigens contain linear B-cell epitopes. Despite their relatively low abundance, the prediction of linear B-cell epitopes has received considerable attention. Most antigens are proteins and those are the subjects for epitope prediction methods (Van Regenmortel, 2009) ^[24]. The ability to identify the binding areas in the antigen's sequence or structure is important for the development of diagnostic tests or assays. Epitope mapping refers to the systematic identification and characterisation of potential antigenic determinants (epitopes) (Jespersen *et al.*, 2019) ^[25]. It is a fundamental step in immunological research, as it provides critical insights into antigens that aid in the development of diagnostic assays. Earlier linear B-cell epitopes were predicted from the primary sequence of antigens using sequence-based methods. Early computational approaches for B-cell epitope prediction relied on amino acid propensity scales that utilised the physicochemical properties associated with B-cell epitopes. Hopp and Wood applied residue hydrophilicity calculations for B-cell epitope prediction (Hopp and Wood, 1981) ^[26] on the assumption that hydrophilic regions are predominantly located on the protein surface and are potentially antigenic. Now it is known that protein surfaces contain roughly the same number of hydrophilic and hydrophobic residues (Lins *et al.*, 2003) ^[27]. Other amino acid propensity scales introduced for B-cell epitope prediction are based on flexibility (Karplus and Schulz, 1985) ^[28], surface accessibility (Emeni *et al.*, 1985) ^[29] and β -turn propensity (Pellequer *et al.*, 1993) ^[30]. A related method to predict B-cell epitopes was introduced by Kolaskar and Tongaonkar (Kolaskar and Tongaonkar, 1990) ^[31], consisting on a simple antigenicity scale derived from physicochemical properties and frequencies of amino acids in experimentally determined B-cell epitopes. This index was the most popular antigenic scale for B-cell epitope prediction. The limitations in the performance of amino acid propensity scales for the prediction of linear B-cell epitopes led to the development of machine learning (ML)-based methods. These methods

are developed by training ML algorithms to distinguish experimental B-cell epitopes from non-B-cell epitopes. Prior to training, B-cell epitopes are translated into feature vectors capturing selected properties, such as those given by different propensity scales. Relevant examples of B-cell epitope prediction methods based on ML include BepiPred. The datasets, training features and algorithms used for developing these methods differ. BepiPred is based on random forests trained on B-cell epitopes obtained from 3D-structures of antigen-antibody complexes (Jespersen *et al.*, 2017) [32]. In general, B-cell epitope prediction methods employing ML algorithms are reported to outperform those based on amino acid propensity scales. Some studies have reported that ML based algorithms demonstrate only marginal improvement over traditional single-scale-based methods (Greenbaum *et al.*, 2007) [33]. The Molecular Evolutionary Genetics Analysis (MEGA) software was designed for comparative analysis of homologous sequences (Kumar *et al.*, 2008) [34]. An Immune Epitope Database (IEDB) - epitope conservancy analysis tool was developed to analyse the variability or conservation of epitopes. The tool is user-friendly and aids in the design of epitope-based vaccines and diagnostics (Bui *et al.*, 2007) [35]. VaxiJen server is used for the antigenicity prediction of candidate diagnostic antigens. It is an alignment-free bioinformatics tool for *in silico* identification of immunogens/antigens. Its latest version (VaxiJen 3.0) uses machine learning to predict the antigenicity of a protein using E-descriptors (Miles *et al.*, 2025) [36]. The physicochemical properties of the putative antigen can be assessed to evaluate its potential as a diagnostic antigen using the ProtParam tool, accessible on the ExPASy server (Omidian *et al.*, 2024) [37]. It computes various physical and chemical parameters for a given protein sequence. Also, the solubility of the putative diagnostic antigen can be evaluated using the Protein-Sol server based on the observation of a bimodal distribution of protein solubilities for *E.coli* proteins in cell-free expression (Hebditch *et al.*, 2017) [38].

2. Methodology

2.1 Selection of leptospiral outer membrane proteins and sequence analysis

The Leptospiral OMPs that have high copy number, that are immunodominant and that are highly conserved across the majority of the serovars in pathogenic *Leptospira* were selected for the *in-silico* analysis. The FASTA sequences of the targeted OMPs from different species of *Leptospira* were retrieved from the UniProt/Swiss-prot (<https://uniprot.org>) and National Center for Biotechnology Information (NCBI) protein database (<https://www.ncbi.nlm.nih.gov/>). Protein sequences of the selected OMPs of the pathogenic *Leptospira* species were subjected to multiple sequence alignment (MSA) performed using the Alignment Explorer of MEGA version 12 to assess the homologous protein sequences from different species.

2.2 Linear B-cell epitope prediction of leptospiral outer membrane proteins

The selected sequence was further subjected to immunoinformatics analysis for the prediction of linear B-cell epitopes using the IEDB server (<https://tools.iedb.org/bcell/>) (Vita *et al.*, 2019) [39]. The FASTA sequences of the proteins were submitted to the server and epitopes were predicted using the Bepipred linear

epitope prediction 2.0, Chou & Fasman beta-turn prediction and Karplus & Schulz flexibility prediction methods, with the default threshold value.

2.3 Conservancy analysis

The IEDB conservancy analysis tool for epitopes (<https://tools.iedb.org/conservancy/>) was used to select conserved epitope sequences. The tool analyses the conservancy or variability of amino acid sequences within a given set of protein sequences. The sequence identity threshold was set at the default ($\geq 100\%$) (Bui *et al.*, 2007) [35]. In order to confirm the homologs of the selected proteins within different serovars of pathogenic *Leptospira* species, Protein Basic Local Alignment Search Tool (BLASTp) (<https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins>) was performed against the proteome of pathogenic *Leptospira* species.

2.4 Antigenicity analysis

Antigenicity refers to the ability of an antigen to induce an immune response. The antigenicity of the predicted epitopes and the *rLep5ECP* was determined using the VaxiJen v2.0 (<https://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html>) (Doytchinova and Flower, 2007) [40] and VaxiJen v3.0 (<http://ddg-pharmfac.net/vaxijen3/home/>) (Miles *et al.*, 2025) [36] tools using a bacterial model at a threshold of 0.4.

2.5 Engineering of the synthetic multi-epitope construct for the production of rLep5ECP

A synthetic multi-epitope construct was designed by assembling the five selected predicted linear B-cell epitope sequences. The selected epitopes were linked using the flexible peptide linker GSGSG, which served as a spacer between adjacent epitope sequences. The resulting chimeric construct was designated for subsequent recombinant expression and characterisation of Lep5ECP, which was subjected to further analysis to assess its antigenicity, physicochemical parameters and solubility.

2.6 Analysis of physicochemical properties of the rLep5ECP

The physicochemical properties of the *rLep5ECP* such as theoretical isoelectric point (pI), molecular weight, amino acid composition, atomic composition, extinction coefficient, estimated half-life, instability index, aliphatic index, total number of positive and negative residues and grand average of hydropathicity (GRAVY) were determined using the online server ExPASy - ProtParam (<https://web.expasy.org/protparam/>) (Aathmanathan *et al.*, 2018) [41] to predict its suitability for recombinant protein expression and downstream applications.

2.7 Solubility prediction of the rLep5ECP

The prediction of solubility of the *rLep5ECP* was performed using the Protein-Sol server (<https://protein-sol.manchester.ac.uk/>). This server uses available data for *Escherichia coli* protein solubility in a cell-free expression system, to predict the solubility of a protein (Hebditch *et al.*, 2017) [38].

3. Results

3.1 Selection and multiple sequence alignment of leptospiral outer membrane proteins

The five OMPs selected from pathogenic leptospires such as *L. interrogans*, *L. kirschneri*, *L. weilii*, *L. borgpetersenii*, *L.*

santarosai and *L. noguchii* were *LipL21*, *LipL32*, *LipL41*, *OmpL1* and *Loa22*. Protein sequences showing high sequence homology among the selected OMPs of pathogenic *Leptospira* species determined by multiple sequence alignment (Table 1), were selected for subsequent linear B-cell epitope prediction.

Table 1: Multiple sequence alignment of selected leptospiral outer membrane protein sequences using alignment explorer of MEGA 12 showing sequence similarities

LipL21	
Species/Abbrv	* * * * *
1. INT	G V G V Y E C K A T G S G S D P K D V S K D N W E C C Q C V I Y A K F P G G
2. KIR	G V G V Y E C K A T G S G S D P K D V S K D N W E C C Q C V I Y A K F P G G
3. SAN	G V G V Y E C K A T G S G S D P K D V S K D N W E C C Q C V I Y A K F P G G
4. WEI	G V G V Y E C K A T G S G S D P K D V S K D N W E C C Q C V I Y A K F P G G
5. Sequence 1	
LipL41	
Species/Abbrv	* * * * *
1. INT	A F A A S M A T G K D V H G N E P V S K P T G V R M M L I P L D A
2. KIR	A F A A S M A T G K D V H G N E P V S K P T G V R M M L I P L D A
3. BOR	A F A A S M A T G K D V H G N E P V S K P T G V R M M L I P L D A
4. WEI	A F A A S M A T G K D V H G N E P V S K P T G V R M M L I P L D A
5. NOG	
LipL32	
1. INT	K A A T P E E K S P H W F D T W I R V E R M S A I M P D Q I A K A A K A K P V Q K L D D D D G D D T Y K E E R H N K Y
2. BOR	K A A T P E E K S P H W F D T W I R V E R M S A I M P D Q I A K A A K A K P V Q K L D D D D G D D T Y K E E R H N K Y
3. WEI	K A A T P E E K S P H W F D T W I R V E R M S A I M P D Q I A K A A K A K P V Q K L D D D D G D D T Y K E E R H N K Y
4. SAN	K A A T P E E K S P H W F D T W I R V E R M S A I M P D Q I A K A A K A K P V Q K L D D D D G D D T Y K E E R H N K Y
5. KIR	K A A T P E E K S P H W F D T W I R V E R M S A I M P D Q I A K A A K A K P V Q K L D D D D G D D T Y K E E R H N K Y
6. NOG	
OmpL1	
Species/Abbrv	* * * * *
1. INT	V A N L A D G T D P I T T R E H V R F R T S G I A
2. SAN	V A N L A D G T D P I T T R E H I R F R A S G I A
3. BOR	V A N L A D G T D P I T T R E H I R F R A S G I A
4. KIR	V A N L A D G T D P I T T R E H V R F R T S G I A
5. NOG	
6. WEI	
Loa22	
Species/Abbrv	* * * * *
1. INT	K E G L G K L P D S A L E I T G H T D A I G P E A E G A K K G N I F Y S E L R A N A V
2. KIR	K E G L G K L P D S A L E I T G H T D A I G P E A E G A K K G N I F Y S E L R A N A V
3. SAN	K E G L G K L P D S A L E I T G H T D A I G P E A E G A K K G N I F Y S E L R A N A V
4. BOR	K E G L G K L P D S A L E I T G H T D A V G P E A E G A K K G N I F Y S E L R A N A V
5. WEI	K E G L G K L P D S A L E I T G H T D A V G P E A E G P K K G N I F Y S E L R A N A V
6. NOG	

3.2 Linear B-cell epitope prediction

In total, five epitopes were predicted to be immunogenic by the computational analysis of the linear B-cell epitopes done using the IEDB methods and are presented in Table 2, while the list of selected linear B-cell epitope sequences is provided in Table 3.

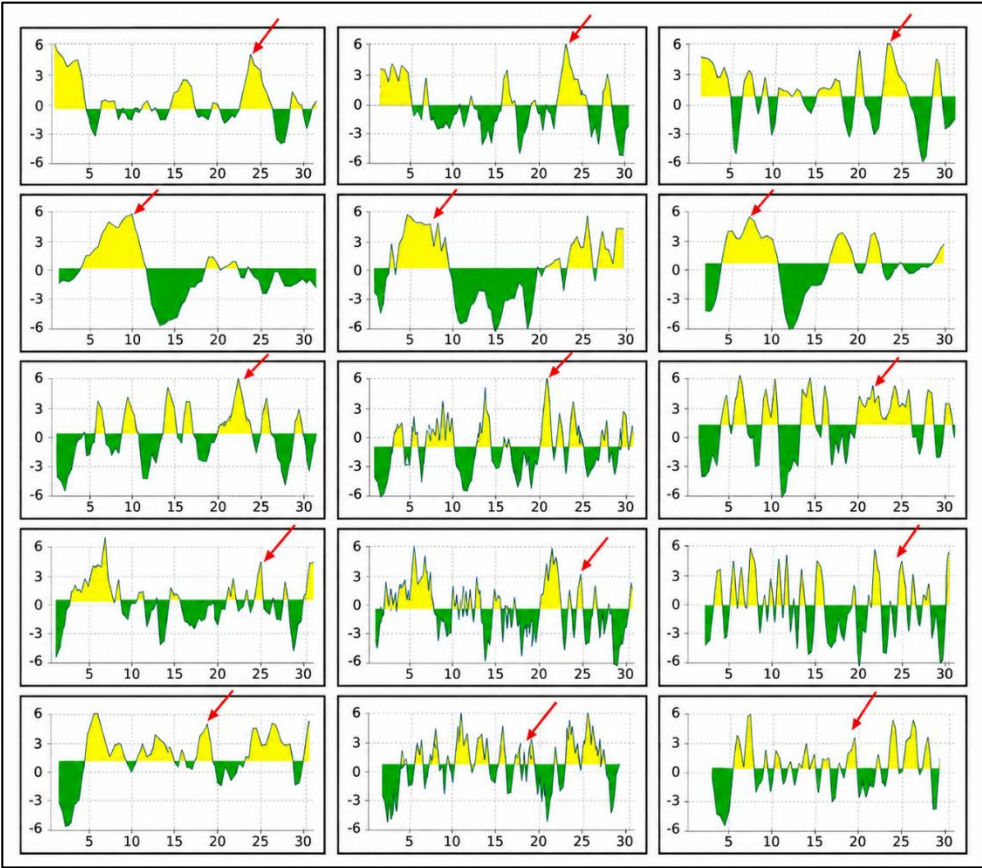


Fig 1: Linear B-cell epitope prediction by IEDB using the Bepipred linear epitope prediction 2.0, Chou & Fasman beta-turn prediction and Karplus & Schulz flexibility prediction methods, showing potentially antigenic epitopes (yellow peaks) above the threshold

Table 2: List of predicted linear B-cell epitope sequences from leptospiral outer membrane proteins

OMP	Predicted linear B-cell epitope	No. of amino acids
LipL21	CKATGSGSDPKDVSNDWE	19
LipL41	VNTGNEPVSKPTG	13
LipL32	IAKAAKAKPVQKLDDDDGDDTYKEERHNK	30
OmpL1	ADGTDPIIT	9
Loa 22	HTDAIGPEQAEGAKK	15

3.4 Prediction of antigenicity

All five predicted OMP epitopes were predicted as probable antigen/immunogen. Antigenicity values varied from 0.5159 to 1.2378. Epitopes that had the highest antigenicity value and probability of 100% were selected. The detailed antigenicity analysis results are presented in Table 2.

The antigenicity score of the *rLep5ECP* predicted by the VaxiJen 2.0 server was 1.4268, while VaxiJen 3.0 predicted the recombinant antigen to be immunogenic with a probability of 100%. The results indicate that the generated *rLep5ECP* was antigenic in nature.

3.4 Prediction of antigenicity

All five predicted OMP epitopes were predicted as probable antigen/immunogen. Antigenicity values varied from 0.5159 to 1.2378. Epitopes that had the highest antigenicity

value and probability of 100% were selected. The detailed antigenicity analysis results are presented in Table 4.

The antigenicity score of the *rLep5ECP* predicted by the VaxiJen 2.0 server was 1.4268, while VaxiJen 3.0 predicted the recombinant antigen to be immunogenic with a probability of 100%. The results indicate that the generated *rLep5ECP* was antigenic in nature.

Table 3: Results of antigenicity analysis using VaxiJen 2.0 and VaxiJen 3.0 server

VaxiJen v2.0	VaxiJen v3.0
1.0428 (Probable antigen)	Probable immunogen with 100% probability
1.1997 (Probable antigen)	Probable immunogen with 100% probability
0.5159 (Probable antigen)	Probable immunogen with 100% probability
1.2378 (Probable antigen)	Probable immunogen with 100% probability
0.5376 (Probable antigen)	Probable immunogen with 100% probability

3.5 Physicochemical analysis of *rLep5ECP*

The physicochemical characteristics of *rLep5ECP* were predicted using the ExPASy ProtParam tool and the results are presented in Table 4. The predicted parameters provide a comprehensive overview of the protein's basic physicochemical features.

Table 4: Physicochemical properties of the *rLep5ECP*

Physicochemical properties		Values
Number of amino acids		127
Formula		C ₅₂₅ H ₈₂₆ N ₁₇₀ O ₁₉₈ S ₃
Molecular weight		12783.53
Theoretical isoelectric point		5.94
Total number of negatively charged residues (Asp + Glu)		19
Total number of positively charged residues (Arg + Lys)		14
Total number of atoms		1722
Ext. coefficient		6990
Estimated half-life		30 hours (mammalian reticulocytes, <i>in vitro</i>) >20 hours (yeast, <i>in vivo</i>) >10 hours (<i>Escherichia coli</i> , <i>in vivo</i>).
Instability index		17.62 (classified as stable)
Aliphatic index		33.86
Grand average of hydropathicity (GRAVY)		-1.176 (protein is hydrophilic in nature)
Amino acid	Number of residues	Composition (%)
Ala (A)	9	7.1%
Arg (R)	2	1.6%
Asn (N)	4	3.1%
Asp (D)	13	10.2%
Cys (C)	1	0.8%
Gln (Q)	2	1.6%
Glu (E)	6	4.7%
Gly (G)	23	18.1%
His (H)	9	7.1%
Ile (I)	3	2.4%
Leu (L)	2	1.6%
Lys (K)	12	9.4%
Met (M)	2	1.6%
Phe (F)	0	0.0%
Pro (P)	7	5.5%
Ser (S)	17	13.4%
Thr (T)	8	6.3%
Trp (W)	1	0.8%
Tyr (Y)	1	0.8%
Val (V)	5	3.9%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%

3.6 Solubility prediction of the rLep5ECP

The Protein-Sol server predicted the rLep5ECP to be soluble upon expression with a score of 0.859.

4. Discussion

Leptospirosis is a globally re-emerging zoonotic disease that affects both humans and animals. Therefore, there is a critical need for an early and reliable diagnostic test to facilitate timely diagnosis and effective disease management.

Understanding the nature and identification of leptospiral antigens is crucial for both the pathophysiology and diagnostics of leptospirosis (Chang *et al.*, 2016) [42]. The identification and characterisation of immunogenic leptospiral proteins expressed during infection can be done with suitable tools and methods (Samrot *et al.*, 2021) [43]. Identifying epitopes in antigens is of great interest for several practical reasons, such as designing and developing epitope-based diagnostic assays.

The application of multi-epitope antigens has significantly improved diagnostic accuracy, particularly for pathogens with high genetic variability, where single-epitope strategies often face challenges such as cross-reactivity. To address these limitations, innovative approaches such as mimotopes and fusion antigens have been developed, in which multiple antigenic domains or epitopes are combined within a single polypeptide to improve the specificity and sensitivity of immunodiagnostic assays (Bulla *et al.*, 2025) [44]. Efforts to improve serological diagnostics are increasingly focused on improving accuracy and sensitivity. Currently, epitope selection for diagnostic purposes prioritises those capable of inducing rapid, specific and measurable immune responses. Consequently, epitope mapping has become essential to identifying antigenic regions with the greatest diagnostic potential (Ahmad *et al.*, 2016) [45].

The reverse vaccinology *in-silico* approach provides preliminary predictions about potential candidates by using genome sequences that can be translated into proteins. The utilisation of the reverse vaccinology *in-silico* approach is crucial because it offers predictions of antigenicity, B-cell epitope regions and other parameters such as the solubility of the targeted proteins (Can *et al.*, 2020) [46]. Reverse vaccinology leverages genomic information to predict specific antigenic regions that interact with antibodies. B-cell epitopes are recognised by B-cell receptors or antibodies in their native structure (Soria *et al.*, 2015) [47].

Knowledge of B-cell epitope binding regions in the antigen's sequence or structure is important in the development of diagnostic tests (Schellekens *et al.*, 2000) [48]. Therefore, identification of putative immunogenic epitopes with the help of *in-silico* methods can help in minimising the time, cost and resources associated with the experimental mapping of epitopes. However, the advancement of machine learning has improved the performance of *in-silico* prediction methods for mapping of B-cell epitopes (Kringelum *et al.*, 2012) [49]. *In-silico* linear B-cell epitope prediction is achieved by correlating B-cell epitope localisation with physicochemical properties, including hydrophilicity, solvent accessibility, flexibility, turns, polarity, antigenicity and surface exposure. The analysis resource offers tools for predicting linear epitopes using several approaches, including amino acid scales, machine learning techniques and molecular geometrical properties.

Computational analysis of the consensus sequences of the identified genes to map potential B-cell epitopes was done using the IEDB BepiPred 1.0 prediction server for the prediction of the location of linear B-cell epitopes based on a combination of a hidden markov model and a propensity scale method (Larsen *et al.*, 2006) [50]. One of the latest additions is the second version of the BepiPred-2.0 method, a sequence-based B-cell epitope prediction method that was developed in 2017. This method is based on a random forest algorithm trained on epitopes annotated from antibody-antigen protein structures (Jespersen *et al.*, 2017) [32]. VaxiJen server for the identification of immunogenic antigens for the determination of highly antigenic proteins. VaxiJen 2.0 uses an alignment-independent method based on auto- and cross-covariance (ACC) transformation of protein sequences into uniform equal-length vectors and a partial least squares algorithm (PLS) for antigen prediction (Doytchinova and Flower, 2007) [40]. The amino acids in protein sequences are presented by z-descriptors, representing their main physicochemical properties. The latest version VaxiJen 3.0 uses machine learning to predict the antigenicity of a protein using machine learning based on E-descriptors which are the principal component analysis of the physicochemical properties of amino acids. Each protein is transformed into a uniform vector, which is then used to train three machine learning algorithms that employ XGBoost, random forest and random subspace Method-k nearest neighbour (RSM-kNN) that classify proteins as probable immunogens (antigens) or non-immunogens (Miles *et al.*, 2025) [36]. The physicochemical and structural properties of the proteins are well understood with the use of computational tools. The statistics about a protein sequence such as the number of amino acids, sequence length and the physicochemical properties of a protein, can be computed by ExPASy - ProtParam (Sivakumar *et al.*, 2007) [51].

In this study, five immunogenic linear B-cell epitopes from the pathogenic *Leptospira* OMPs with high copy number, that are immunodominant and highly conserved across the majority of the serovars were selected and identified using the IEDB analysis resource. Candidate linear B-cell epitopes were predicted using a combination of the machine learning-based BepiPred-2.0 algorithm and classical amino acid propensity scale methods, including Chou & Fasman β -turn prediction and Karplus & Schulz flexibility prediction. Residues with prediction scores above the respective threshold values were considered more likely to constitute linear B-cell epitopes, as higher scores indicate a greater probability of antigenicity and antibody recognition.

The identification of conserved residues following the similarity analysis enabled the selection of epitopes with high levels of sequence conservation by the IEDB Epitope Conservancy Analysis Tool. Epitopes exhibiting conservancy ranging from 93.33% to 100% among pathogenic *Leptospira* species were considered highly conserved were selected for this study. In the development of a diagnostic antigen, the use of highly conserved epitopes is expected to enable broad and reliable detection of antibodies against multiple strains or species by targeting antigenic regions with minimal sequence variation, thereby improving the sensitivity and diagnostic applicability of the recombinant antigen. Also, the selected epitopes were analysed for their similarity to *Leptospira* strains using the BLASTp server. The epitope sequences showing higher sequence identity and query coverage were selected.

The design of a synthetic multi-epitope construct is a critical step in the production of a functional recombinant chimeric protein. By fusing two or more protein domains, a fusion protein usually obtains the biological activities from each component. However, the direct fusion of proteins often results in impaired biological activity (Gokhale and Khosla, 2000) [52]. Therefore, the use of appropriate peptide linkers between adjacent epitopes helps maintain their structural integrity and facilitates proper folding of the recombinant protein, thereby enhancing its functional and immunological properties. Flexible linkers rich in amino acids such as glycine and serine provide good flexibility and solubility. They are suitable choices when certain movements or interactions are required for fusion protein domains. The length of the flexible linkers can be adjusted to allow for proper folding or to achieve optimal biological activity of the fusion proteins (Chen *et al.*, 2013) [53].

The antigenicity of the five predicted OMP-derived epitopes was evaluated using the VaxiJen servers. The VaxiJen 2.0 scores ranged from 0.5159 to 1.2378, while VaxiJen 3.0 classified all five epitopes as 100% probable immunogens. The antigenicity score of *rLep5ECP* predicted by the VaxiJen 2.0 server was 1.4268, while VaxiJen 3.0 predicted the recombinant antigen to be immunogenic with a probability of 100%. These findings indicate that *rLep5ECP* possesses strong antigenic potential and is a promising candidate for diagnostic applications.

Various physicochemical properties of the *rLep5EFP* were predicted by the ExPASy - ProtParam online server to assess its suitability for downstream applications such as recombinant protein expression and protein purification. The protein consisted of 127 amino acids (5 epitope sequences along with linkers and a 6x histidine tag) with a theoretical pI of 5.94, indicating that it is slightly acidic in nature. Since protein solubility is influenced by the relationship between environmental pH and pI, this property may facilitate solubility and aid in purification strategy for optimisation of an appropriate buffer. The molecular weight was predicted to be 12.8 kDa. Analysis of the amino acid composition revealed that glycine (18.1%) and serine (13.4%) were the most abundant residues, followed by aspartic acid (10.2%) and lysine (9.4%). The relatively high glycine content may contribute to structural flexibility, while the hydroxyl group of serine residues enables hydrogen bond formation, which may influence protein stability. Aspartic acid and glutamic acid together constituted 19 negatively charged residues, while arginine and lysine accounted for 14 positively charged residues, consistent with the acidic theoretical pI of the recombinant protein. The atomic composition of *rLep5EFP* was determined to be $C_{525}H_{826}N_{170}O_{198}S_3$, comprising a total of 1,722 atoms. This molecular formula provides the elemental composition of the recombinant protein and serves as a reference for computational characterisation and molecular mass determination. The predicted extinction coefficient of $6,990 \text{ M}^{-1} \text{ cm}^{-1}$ at 280 nm, with an absorbance value of 0.547 (0.1%, 1 g L⁻¹), indicates that the recombinant protein can be reliably quantified using UV spectrophotometry.

The estimated half-life of the recombinant protein was predicted to be 30 hours in mammalian reticulocytes (*in vitro*), greater than 20 hours in yeast (*in vivo*) and greater than 10 hours in *Escherichia coli* (*in vivo*). The predicted half-life suggests that the recombinant protein is unlikely to

undergo rapid intracellular degradation, thereby favouring efficient expression, accumulation and purification. The aliphatic index of the recombinant protein was calculated to be 33.86. Protein stability was further supported by an instability index of 17.62, which is well below the threshold value of 40 that classifies the recombinant protein as stable (Shams *et al.*, 2021) [54]. Stable proteins are generally more suitable for recombinant protein expression because they are less susceptible to proteolytic degradation during production and purification, thereby improving protein yield and reproducibility. The grand average of hydropathicity value was -1.176, indicating that the recombinant protein is predominantly hydrophilic in nature and can interact with water molecules (Ali *et al.*, 2017) [55]. Proteins with a hydrophilic character have important advantages for inducing effective immunologic responses. Hydrophilic amino acid residues are typically exposed on the protein surface, resulting in readily accessible B-cell epitopes that enhance B-cell recognition and antibody production (Franjinc, 2025) [56]. The Protein-Sol server predicted the *rLep5ECP* to be soluble upon heterologous expression with a score of 0.859. Any scaled solubility value greater than 0.45 is predicted to have a higher solubility than the average soluble *E. coli* protein from the experimental solubility dataset (Niwa *et al.*, 2009) [57] and any protein with a lower scaled solubility value is predicted to be less soluble in nature.

Recent technological advances have transformed epitope prediction through the integration of computational analysis, structural modelling and experimental validation enabling the identification of precise diagnostic candidates. The development of artificial intelligence-based tools has further strengthened the role of bioinformatics in epitope mapping. Nevertheless, experimental validation remains essential to confirm computational predictions and ensure epitope specificity and diagnostic accuracy. The combined use of advanced computational approaches and experimental validation is expected to improve the discovery of reliable candidates (Bulla *et al.*, 2025) [44].

5. Conclusion

Leptospirosis remains a significant public health concern due to its increasing global incidence and frequent outbreaks in both developing and developed countries. Immunoinformatics-based identification of diagnostic targets represents a promising strategy for leptospirosis diagnosis. In the present study, an immunoinformatics approach was employed to identify highly immunogenic, conserved, linear B-cell epitopes from leptospiral OMPs for the production of a recombinant multi-epitope diagnostic antigen (*rLep5EFP*). In addition, the physicochemical properties of the designed antigen were evaluated to assess its suitability for recombinant protein expression and purification. The *in silico* analyses indicated that the designed antigen possesses favourable antigenic and physicochemical characteristics, supporting its potential as a candidate for the serodiagnosis of leptospirosis. Nevertheless, computational predictions have several limitations. Therefore, comprehensive experimental validation is essential to confirm the sensitivity, specificity and overall diagnostic utility of the proposed multi-epitope antigen. Overall, this study provides a rational framework for the development of improved recombinant multi-epitope antigen-based diagnostic assays for leptospirosis and

highlights the value of immunoinformatics in accelerating diagnostic antigen discovery.

6. Acknowledgement

We gratefully acknowledge Zoonoses Research Laboratory, Centre for Animal Health Studies, Madhavaram Milk Colony and Department of Animal Biotechnology, Madras Veterinary College, TANUVAS for their constant support for the successful completion of this study. We acknowledge the ICAR-All India Network Project on “One Health Approach to Zoonotic Diseases” for providing the necessary funding support to carry out the present study.

References

1. Aathmanathan VS, Jothi N, Prajapati VK, Krishnan M. Investigation of immunogenic properties of Hemolin from silkworm, *Bombyx mori*, as carrier protein: an immunoinformatic approach. *Sci Rep*. 2018;8(1):6957.
2. Ahmad TA, Eweida AE, Sheweita SA. B-cell epitope mapping for the design of vaccines and effective diagnostics. *Trials Vaccinol*. 2016;5:71–83.
3. Ali M, Pandey RK, Khatoon N, Narula A, Mishra A, *et al*. Exploring dengue genome to construct a multi-epitope based subunit vaccine by utilizing immunoinformatics approach to battle against dengue infection. *Sci Rep*. 2017;7:9232.
4. Amran F, Noor Halim NA, Muhammad AH, Mohd Khalid MKN, Dasiman NM, *et al*. Application of multilocus sequence typing for the characterization of *Leptospira* strains in Malaysia. *Trop Med Infect Dis*. 2023;8(2):69.
5. Bharti AR, Nally JE, Ricaldi JN, Matthias MA, Diaz MM, *et al*. Leptospirosis: a zoonotic disease of global importance. *Lancet Infect Dis*. 2003;3(12):757–771.
6. Bui HH, Sidney J, Li W, Fusseder N, Sette A. Development of an epitope conservancy analysis tool to facilitate the design of epitope-based diagnostics and vaccines. *BMC Bioinformatics*. 2007;8:361.
7. Bulla ACS, Sbrano da Silva A, Prado Sereno B, Dias MFR, Leal da Silva M. Computational methods in immunoinformatics: epitope discovery and diagnostic applications. *ACS Omega*. 2025;10(39):44816–44839.
8. Can H, Alak SE, Köseoglu AE, Döşkaya M, Ün C. Do *Toxoplasma gondii* apicoplast proteins have antigenic potential? An *in silico* study. *Comput Biol Chem*. 2020;84:107158.
9. Chang MY, Cheng YC, Hsu SH, Ma TL, Chou LF, *et al*. Leptospiral outer membrane protein LipL32 induces inflammation and kidney injury in zebrafish larvae. *Sci Rep*. 2016;6:27838.
10. Chen X, Zaro JL, Shen WC. Fusion protein linkers: property, design and functionality. *Adv Drug Deliv Rev*. 2013;65(10):1357–1369.
11. Costa F, Hagan JE, Calcagno J, Kane M, Torgerson P, *et al*. Global morbidity and mortality of leptospirosis: a systematic review. *PLoS Negl Trop Dis*. 2015;9(9):e0003898.
12. Donati C, Rappuoli R. Reverse vaccinology in the 21st century: improvements over the original design. *Ann N Y Acad Sci*. 2013;1285(1):115–132.
13. Doytchinova IA, Flower DR. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics*. 2007;8(1):4.
14. Emini EA, Hughes JV, Perlow DS, Boger J. Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. *J Virol*. 1985;55(3):836–839.
15. Faine S, Adler B, Perolat P, Bolin CA. *Leptospira* and leptospirosis. 2nd ed. Melbourne: MediSci; 1999.
16. Flannery B, Costa D, Carvalho FP, Guerreiro H, Matsunaga J, *et al*. Evaluation of recombinant *Leptospira* antigen-based enzyme-linked immunosorbent assays for the serodiagnosis of leptospirosis. *J Clin Microbiol*. 2001;39(9):3303–3310.
17. Franjinc S. Solubility and immunity. *Mathews J Immunol Allergy*. 2025;9(1).
18. Gokhale RS, Khosla C. Role of linkers in communication between protein modules. *Curr Opin Chem Biol*. 2000;4(1):22–27.
19. Greenbaum JA, Andersen PH, Blythe M, Bui HH, Cachau RE, *et al*. Towards a consensus on datasets and evaluation metrics for developing B-cell epitope prediction tools. *J Mol Recognit*. 2007;20(2):75–82.
20. Haake DA, Chao G, Zuerner RL, Barnett JK, Barnett D, *et al*. The leptospiral major outer membrane protein LipL32 is a lipoprotein expressed during mammalian infection. *Infect Immun*. 2000;68(4):2276–2285.
21. Hebditch M, Carballo-Amador MA, Charonis S, Curtis R, Warwicker J. Protein-Sol: a web tool for predicting protein solubility from sequence. *Bioinformatics*. 2017;33(19):3098–3100.
22. Hopp TP, Woods KR. Prediction of protein antigenic determinants from amino acid sequences. *Proc Natl Acad Sci U S A*. 1981;78(6):3824–3828.
23. Jespersen MC, Mahajan S, Peters B, Nielsen M, Marcatili P. Antibody specific B-cell epitope predictions: leveraging information from antibody-antigen protein complexes. *Front Immunol*. 2019;10:298.
24. Jespersen MC, Peters B, Nielsen M, Marcatili P. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Res*. 2017;45(W1):W24–W29.
25. Joseph S, Thomas N, Thangapandian E, Singh VP, Verma R, *et al*. Evaluation and comparison of native and recombinant LipL21 protein-based ELISAs for diagnosis of bovine leptospirosis. *J Vet Sci*. 2012;13(1):99–101.
26. Karplus PA, Schulz GE. Prediction of chain flexibility in proteins. *Naturwissenschaften*. 1985;72(4):212–213.
27. Kolaskar AS, Tongaonkar PC. A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS Lett*. 1990;276(1-2):172–174.
28. Kringelum JV, Lundegaard C, Lund O, Nielsen M. Reliable B cell epitope predictions: impacts of method development and improved benchmarking. *PLoS Comput Biol*. 2012;8(12):e1002829.
29. Kumar S, Dudley J, Nei M, Tamura K. MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. *Brief Bioinform*. 2008;9(4):299–306.
30. Larsen JEP, Lund O, Nielsen M. Improved method for predicting linear B-cell epitopes. *Immunome Res*. 2006;2:2.
31. Lefranc MP. Immunoglobulin and T cell receptor genes: IMGT and the birth and rise of immunoinformatics. *Front Immunol*. 2014;5:22.

32. Levett PN. Leptospirosis. Clin Microbiol Rev. 2001;14(2):296–326.
33. Lin X, Chen Y, Yan J. Recombinant multiepitope protein for diagnosis of leptospirosis. Clin Vaccine Immunol. 2008;15(11):1711–1714.
34. Lins L, Thomas A, Brasseur R. Analysis of accessible surface of residues in proteins. Protein Sci. 2003;12(7):1406–1417.
35. Masignani V, Rappuoli R, Pizza M. Reverse vaccinology: a genome-based approach for vaccine development. Expert Opin Biol Ther. 2002;2(8):895–905.
36. Miles S, Menafrá G, Iriarte A, Chabalgoity JA. IAPred: a versatile open-source tool for predicting protein antigenicity across diverse pathogens. Immunoinformatics. 2025;20:100061.
37. Natarajaseenivasan K, Vijayachari P, Sharma S, Sugunan AP, Selvin J, *et al.* Serodiagnosis of severe leptospirosis: evaluation of ELISA based on the recombinant OmpL1 or LipL41 antigens of *Leptospira interrogans* serovar Autumnalis. Ann Trop Med Parasitol. 2008;102(8):699–708.
38. Niwa T, Ying BW, Saito K, Jin W, Takada S, *et al.* Bimodal protein solubility distribution revealed by an aggregation analysis of the entire ensemble of *Escherichia coli* proteins. Proc Natl Acad Sci U S A. 2009;106(11):4201–4206.
39. Omidian M, Mostafavi-Pour Z, Asadi M, Sharifdini M, Nezafat N, *et al.* Design and expression of a chimeric recombinant antigen (SsIR-Ss1a) for the serodiagnosis of human strongyloidiasis: evaluation of performance, sensitivity, and specificity. PLoS Negl Trop Dis. 2024;18(7):e0012320.
40. Palaniappan RU, Ramanujam S, Chang YF. Leptospirosis: pathogenesis, immunity, and diagnosis. Curr Opin Infect Dis. 2007;20(3):284–292.
41. Paul WE. Fundamental immunology. 7th ed. Philadelphia: Lippincott Williams & Wilkins; 2012.
42. Pellequer JL, Westhof E, Van Regenmortel MHV. Correlation between the location of antigenic sites and the prediction of turns in proteins. Immunol Lett. 1993;36(1):83–99.
43. Pinto GV, Senthilkumar K, Rai P, Kabekkodu SP, Karunasagar I, *et al.* Current methods for the diagnosis of leptospirosis: issues and challenges. J Microbiol Methods. 2022;195:106438.
44. Samrot AV, Sean TC, Bhavya KS, Sahithya CS, Chandrasekaran S, *et al.* Leptospiral infection, pathogenesis and its diagnosis: a review. Pathogens. 2021;10(2):145.
45. Schellekens GA, Visser H, de Jong BAW, van den Hoogen FHJ, Hazes JMW, *et al.* The diagnostic properties of rheumatoid arthritis antibodies recognizing a cyclic citrullinated peptide. Arthritis Rheum. 2000;43(1):155–163.
46. Senthilkumar TMA, Subathra M, Ramadass P. Evaluation of recombinant leptospiral antigen LipL41 in enzyme-linked immunosorbent assay and latex agglutination test for serodiagnosis of canine leptospirosis. Vet Arhiv. 2007;77(6):475–484.
47. Shams M, Nourmohammadi H, Asghari A, Basati G, Majidiani H, *et al.* Construction of a multi-epitope protein for human *Toxocara canis* detection: immunoinformatics approach multi-epitope construct for *T. canis* serodiagnosis. Inform Med Unlocked. 2021;26:100732.
48. Sivakumar K, Balaji S, Gangaradhakrishnan. *In silico* characterization of antifreeze proteins using computational tools and servers. J Chem Sci. 2007;119(5):571–579.
49. Soria-Guerra RE, Nieto-Gomez R, Govea-Alonso DO, Rosales-Mendoza S. An overview of bioinformatics tools for epitope prediction: implications on vaccine development. J Biomed Inform. 2015;53:405–414.
50. Sumarningsih S, Tarigan S, Susanti, Kusmiyati. Recombinant LipL32 protein for leptospirosis detection in Indonesia. Procedia Chem. 2016;18:18–25.
51. Sykes JE, Hartmann K, Lunn KF, Moore GE, Stoddard RA, *et al.* 2010 ACVIM Small Animal Consensus Statement on leptospirosis: diagnosis, epidemiology, treatment, and prevention. J Vet Intern Med. 2011;25(1):1–13.
52. Validi M, Karkhah A, Prajapati VK, Nouri HR. Immuno-informatics based approaches to design a novel multi epitope-based vaccine for immune response reinforcement against leptospirosis. Mol Immunol. 2018;104:128–138.
53. Van Regenmortel MHV. What is a B-cell epitope? In: Morris GE, editor. Epitope mapping protocols. 2nd ed. Totowa (NJ): Humana Press; 2009. p. 3–20.
54. Vita R, Mahajan S, Overton JA, Dhanda SK, Martini S, *et al.* The Immune Epitope Database (IEDB): 2018 update. Nucleic Acids Res. 2019;47(D1):D339–D343.
55. World Organisation for Animal Health. Manual of diagnostic tests and vaccines for terrestrial animals (Terrestrial manual). Paris: World Organisation for Animal Health; 2021.
56. Yaakob Y, Rodrigues KF, Opook F, William T, John DV. Recombinant LipL32 protein developed using a synthetic gene detects *Leptospira*-specific antibodies in human serum samples. Malays J Med Sci. 2017;24(5):44–51.
57. Ye C, Yan W, Xiang H, He H, Yang M, *et al.* Recombinant antigens rLipL21, rLoa22, rLipL32 and rLigACon4-8 for serological diagnosis of leptospirosis by enzyme-linked immunosorbent assays in dogs. PLoS One. 2014;9(12):e111367.