Designing of Multiepitope Subunit Vaccine Against Leptospirosis Using Immunoinformatics Approach

Kopal Kapoor[1], Satyamitra Shekh[1], Chaitanya G. Joshi[1], Amrutlal Patel[1*]

[1] *Gujarat Biotechnology Research Centre (GBRC), Department of Science and Technology, Government of Gujarat, Gandhinagar, Gujarat, India – 382001*

***Correspondence:**

Dr. Amrutlal Patel

Scientist D & Joint Director

Gujarat Biotechnology Research Centre, Gujarat, India

E-Mail: jd2@gbrc.res.in

Abstract

The most widely spread and neglected zoonotic disease – Leptospirosis, with an estimated 1.03 million cases and 58,900 deaths occurring annually, is a concern and can be prevented only by effective immunization. Due to the urgent need for a novel vaccine, various research organizations have been evaluating the protective immune response elicited by recombinant vaccines. In this study amino-acid sequences of Outermembrane protein (OMPL) 1, LipL 32, 41 and 46 from ten pathogenic serovars of *Leptospira* were selected to obtain epitopes that induce CD4+ and CD8+ T cell responses by binding to the MHC molecules. These B cell and T cell epitopes were used to design multi-epitope vaccine to which an adjuvant sequence was added at the N-terminal end and appropriate linkers were used to join the different epitopes to increase the efficiency of the vaccine constructs. Using an immuno-informatics approach, constructs were analyzed for the physiochemical properties, secondary and tertiary structure

 doi:10.20944/preprints202212.0299.v1

and its validation, docking with different receptors and prediction of binding affinity for each docked complex, followed by molecular dynamics simulation. Taking in consideration all the results, multi-epitope constructs designed may prove to be promising vaccine candidates against Leptospirosis and warrants experimental validation.

Keywords: Leptospirosis, Multi-epitope vaccine, immuno-informatics, T-cell epitope, B-cell epitope, docking, molecular dynamic simulation.

1. Introduction

Leptospirosis is a zoonosis caused by infection with pathogenic *Leptospira* species, which has a worldwide distribution. Leptospirosis is endemic in developing countries such as Malaysia, India, Sri Lanka, and Brazil where thousands of cases are reported annually [1]. The illness range caused by *Leptospira* species in humans is highly broad, ranging from asymptomatic infection to a severe multiorgan infection with substantial mortality [2]. Leptospires are tightly coiled spirochetes that share a typical double membrane structure with other spirochetes, in which the cytoplasmic membrane and peptidoglycan cell wall are closely connected and are overlain by an outer membrane. Furthermore, leptospirosis has emerged as a health threat in new settings due to the influence of globalization and climate [3]. Humans are usually infected through direct or indirect contact with the urine of an infected animal. Warm-climate countries have a much higher prevalence than temperate countries, owing to leptospires prolonged survival in the environment under warm, humid circumstances.

Recent advances in immunology, structural biology, and computational biology have revolutionized vaccine design and engineering. These methods are collectively known as reverse vaccinology [4]. More than 250 putative outer-membrane proteins have been identified using genomic research [5]–[8] which are expected to be involved in host-pathogen interaction due to their location. Inactivated whole cell or membrane preparations of pathogenic

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0299.v1

leptospires are used in veterinary and human vaccines; however, these vaccines fail to provide cross-protective immunity against other serovars not included in the vaccine preparation, and thus do not provide long-term protection against infection [8] – [11]. Inactivated whole cell vaccines, commonly known as bacterins, are currently accessible licensed vaccinations for humans and animals. The use of these vaccines (bacterins) is limited due to a number of drawbacks, including a lack of cross-reactivity or protection against only closely related serovars [12], a lack of long-term immunological memory, the need for regular boosters due to short duration of immunity, and sometimes high reactogenicity [13] – [16]. High production costs and a lack of consistency or repeatability are further issues that limit the usage of bacterins [15], [16]. These needs could be met by proteins that are conserved across species and serovars. Immunogenic proteins from pathogenic *Leptospira*, particularly the outer membrane surface proteins, may be effective vaccines.

Leptospiral OMPs have been found as a key way to discover vaccine candidates in studies. They are potentially relevant in pathogenesis due to their exposed location at the interface between leptospires and the mammalian host. So far, three types of leptospiral OMPs have been identified: 1) lipoproteins, which include LipL32, LipL41, LipL48, LipL36, and LipL21 as well as the temperature-regulated Qlp42 ;2) transmembrane protein OMPL1 (12); and 3) peripheral membrane proteins such LipL45 [17]–[20].

LipL32, also known as Hap-1 (haemolysis associated protein-1), is a potential vaccine candidate. It is the principal outer membrane protein that is exposed to the environment [21], [22]. During human leptospirosis, LipL32 is the most significant protein in the leptospiral protein profile and is an immunodominant antigen [23]– [25]. LipL41 is a surface-exposed lipoprotein that works in tandem with OMPL1 to give immunoprotection. OMPL1 is 320-amino-acid long transmembrane outermembrane protein initially identified by Haake and colleagues in 1993, is a porin expressed by all pathogenic *Leptospira* species investigated [26]–

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0299.v1

[30]. OMPL1 and LipL41 are exposed on the surface, according to surface immunoprecipitation studies [18], [20]. Both are expressed in the mammalian host after infection and are conserved among pathogenic *Leptospira* species. In a hamster model of leptospirosis, studies have shown that when OMPL1 and LipL41 are produced as membrane proteins, they provide considerable protection against homologous challenge [28].

The presence of epitopes that are recognized by immune system cells, particularly B cells, T cells (TCs), and T helpers (THs), is one of the most critical concerns in vaccination [31]. Epitopes are a type of small protein component that comes in two types: continuous or linear epitopes and discontinuous or conformational epitopes. Since epitopes can prevent allergic reactions, cross-reactions, pathological immune responses, and undesirable consequences, they are superior immunogens than complete proteins [32], [33]. As a result, discovering immunogenic epitopes can aid in the development of more effective and safe vaccinations [34].

In this study, using various immuno-informatics approaches, epitopes were predicted from LipL 32, 41, 46 and OMPL 1 protein sequences retrieved from all common pathogenic *Leptospira* species. In the first stage of vaccine designing, appropriate B and T cell epitopes were predicted followed by the formation of multi-epitope vaccine constructs using TLR 4 adjuvant [APPHALS] and different linkers. Later these vaccine constructs were assessed for their physio-chemical properties, secondary structure, tertiary structure and its validation. In the last stage, docking and simulation studies were carried out for the vaccine constructs to get a firm idea about the interaction of the vaccine constructs with different receptors. Simulation studies were performed to check the overall stability of these multi-epitope vaccine constructs. Figure 1 represents the overview of the workflow followed for this research study.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0299.v1

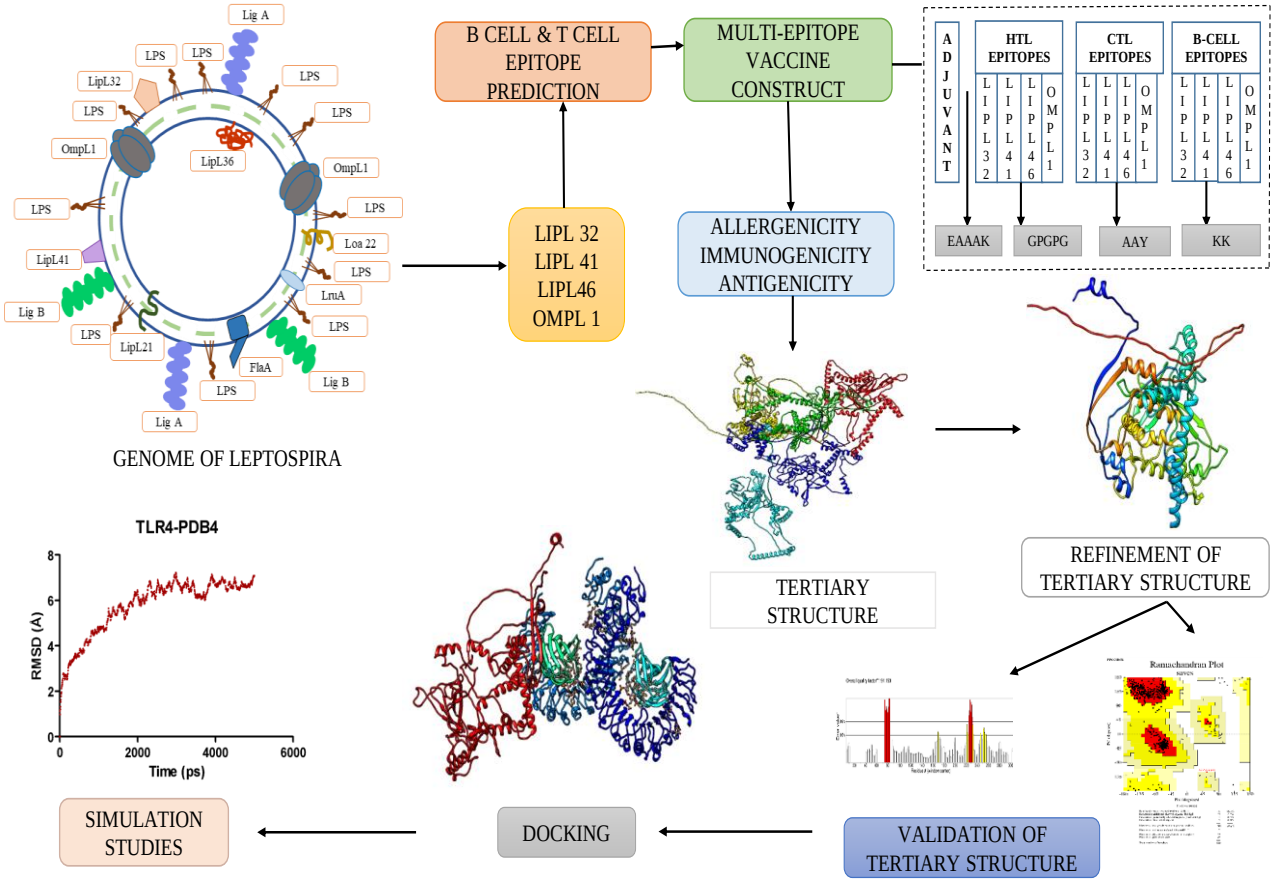


Figure 1. Over-view of the workflow followed for the designing of multi-epitope-based vaccine against Leptospirosis.

2. Materials and methods

2.1. Protein sequence retrieval

The amino-acid sequences of OMPL 1, LipL 32, 41, and 46 for 10 pathogenic different serovars and species of *Leptospira - L.interrogans, L.kischneri, L.noguchii, L.borgerpeterseni, L.alexanderi L.santarosai , L.weilli, L.kmetyi, L.mayottensis , L.alstonii,* were retrieved from the Uniprot database[1] in the standard FASTA format. The accession numbers for all the sequences retrieved are given in (Supplementary Table S1). These sequences are used for B Cell and T cell epitope prediction.

[1] http://www.uniprot.org/proteomes/

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0299.v1

2.2. Linear B cell epitope prediction

IEDB server[2] was used for predicting linear B cell epitopes. It predicts antigenic determinants on proteins based on the physicochemical qualities of amino acid residues as well as their frequency of occurrence in empirically known segmental epitopes. B cell epitopes were predicted using Kolaskar and Tongaonkar method [35].

2.3. Cytotoxic T Cell (CTL) Epitope prediction

The NetCTL 1.2 server was used to identify potential T-cell epitopes that are recognized by CD8+ T cells and triggers a long-lasting and exclusive cytotoxic immune response. NetCTL 1.2 server[3] predicts the epitope using artificial neural network algorithm with threshold value of 0.75, indicating 0.80 sensitivity and 0.97 specificity [36]. Three factors determining the prediction efficiency of NetCTL 1.2 are (i) TAP transport efficiency (ii) peptides binding to MHC class I and (iii) proteasomal C-terminus cleavage.

2.4. Helper T Cell (HTL) Epitope prediction

HTL is important for triggering and creating an efficient humoral or cytotoxic T-cell response. The NetMHCIIpan 4.0 server[4] was used to find the epitopes that may trigger the MHC-II restricted T-cell response. The threshold value was set as 1% and 5% for strong binding epitopes/peptides and weaker binding epitopes/peptides respectively.

2.5. Conservancy of predicted epitopes

IEDB Epitope Conservancy Analysis tool was used to determine the conservancy of the predicted B-Cell, CTL and HTL epitopes. For all the predicted epitopes, conservancy was

[2] http://tools.iedb.org/main/bcell/
[3] http://www.cbs.dtu.dk/services/NetCTL
[4] https://services.healthtech.dtu.dk/service.php?NetMHCIIpan-4.0

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0299.v1

evaluated among ten pathogenic species- *L.interrogans* , *L.kischneri*, *L.noguchii*, *L.borgerpeterseni*, *L.alexanderi*, *L.santarosai* , *L.weilli*, *L. kmetyi*, *L.mayottensis* and *L.alstonii*. The degree of conservancy is defined as the portion of protein sequences that contain the epitope at a specified identity level [37].

2.6. Population coverage analysis

Prediction of the world-wide population coverage along with the South America, South Asia and South Africa of the selected epitopes for MHC -I and MHC- II alleles was carried out using IEDB Population coverage tool[5] [38]. The MHC-I alleles assessed included HLA-A*01:01, HLA-A*03:01, HLA-A*30:02, HLA-A*33:01, HLA-A*68:02, HLA-A*30:01, HLA-A*31:01, HLA-A*68:01, HLA-A*11:01, HLA-A*02:03, HLA-A*23:01, HLA-A*24:02, HLA-A*26:01, HLA-A*02:01, HLA-A*02:06, HLA-A*32:01, HLA-B*15:01, HLA-B*40:01, HLA-B*07:02, HLA-B*51:01, HLA-B*08:01, HLA-B*44:03, HLA-B*44:02, HLA-B*53:01, HLA-B*57:01, HLA-B*58:01, HLA-B*35:0. For MHC-II alleles assessed included: HLA-DRB1*07:01, HLA-DRB1*15:01, HLA-DRB1*03:01.

2.7. Construction of multi-epitope vaccine sequence

All the predicted conserved B cell epitopes, CTL & HTL epitopes from OMPL 1, LipL 32, 41, and 46 using various immuno-informatics tools were joined together to form a multi-epitope vaccine sequence. All epitopes were linked together with different linkers and adjuvants to form a multi epitope vaccine sequence. Linkers were added to ensure the proper separation of different epitopes required for effective functioning of each epitope. Adjuvant selected to construct the multi-epitope vaccine sequence was TLR 4 with sequence APPHALS. The adjuvant was linked with the EAAAK linker to the N-terminal of vaccine construct. All

[5] http://tools.iedb.org/population/

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0299.v1

CTL, HTL and B cell epitopes were linked by AAY, GPGPG and KK linkers respectively. All vaccine constructs designed were subjected to further analysis to assess the antigenicity, allergenicity, toxicity and immunogenicity.

2.8. Allergenicity, Antigenicity, Toxicity and Immunogenicity prediction

In order to predict the allergenicity of the vaccine construct, AllerTOP server[6] was used. The method is based on the translation of protein sequences into uniform equal-length vectors using auto cross covariance [39]. The antigenicity of the vaccine constructs was determined using VaxiJen v2.0 server[7] with default parameters [40]. The toxicity of the multi-epitope vaccine constructs was predicted using ToxIBTL server [41]. The immunogenic nature of the vaccine was predicted using IEDB Resource Analysis Tool[8] [42].

2.9. Physio-chemical parameters analysis & Secondary structure prediction

The ProtParam tool of ExPASy database server[9] was used to determine the physio-chemical parameters of all designed vaccine constructs such as amino acid composition, theoretical isoelectric point, molecular weight, aliphatic index, instability index, extinction coefficient, grand average of hydropathicity, and total number of positive and negative residues.

The secondary structure of the vaccine constructs was predicted using SOPMA server[10] [43].

2.10. Generation of Tertiary Structure

[6] https://www.ddg-pharmfac.net/AllerTOP/
[7] http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html
[8] http://tools.iedb.org/immunogenicity/
[9] http://us.expasy.org/tools/protparam
[10] https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0299.v1

The tertiary structure of the vaccine sequence was predicted using Robetta ab initio method[11] [44]. Ab initio is a method to determine the tertiary structure of protein in the absence of experimentally solved structure of a similar/homologous protein. The structure obtained has 5 models based on their confidence score. The structure with the best confidence score is selected for refinement.

2.11. Tertiary structure refinement

Refinement of the tertiary structure is done to improve the quality of the tertiary structure. The Galaxy-Web Refine server[12], a publicly available server, is used to refine the tertiary structure of the vaccine constructs [45].

2.12. Tertiary structure validation

Since the tertiary structure is generated using ab initio method, it needs to be validated using Prosa-Web[13], PROCHECK and ERRAT[14]. Prosa-Web computes the quality score of the input structure [46]. PROCHECK provides the Ramachandran plot, a visualization technique that predicts energetically acceptable and prohibited dihedral angles psi (Ψ) and phi (φ) of an amino acid based on the side chains van der Waal radius [47]. ERRAT generates the overall quality score of the refined tertiary structure of vaccine constructs. It is a programme that verifies crystallographically determined protein structures [48].

2.13. Molecular docking of vaccine construct with different immune receptors

[11] https://robetta.bakerlab.org/
[12] https://galaxy.seoklab.org/
[13] https://prosa.services.came.sbg.ac.at/prosa.php/
[14] https://saves.mbi.ucla.edu/

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0299.v1

The antigenic molecule must engage with the immune receptor molecule in order to generate the optimal immune response. The interaction between the immunological receptor - TLR 4 [PDB ID: 4G8A], TLR 2 [PDB ID: 2Z7X], MHC II [PDB ID: 4U6Y], MHC I [PDB ID: 1I1Y] and the ligand (multi-epitope vaccine) was explored using molecular docking. PatchDock [15] server was used to evaluate the immunological responses. PatchDock algorithm is categorised into 3 stages: the first is molecular shape representation, then surface patch matching, filtering, and scoring. It gives the result in the form of 10 best solutions for final refinement on the basis of binding score. Atomic contact energy, Van der Waals interaction, partial electrostatics, and binding energy calculations are all included in this score. The patchdock results are further refined using FireDock. The hydrogen bonding interactions of the docked structures were also analyzed using chimera.

2.14. Prediction of binding affinity

The docked structures were checked for their binding affinity using PRODIGY web server [49]. This server is used to predict binding affinity based solely on a protein-protein complex's structural properties.

2.15. Molecular Dynamics simulation

Molecular Dynamic simulation is done to minimize and evaluate the stability of the docked complex. The Desmond tool of Schrodinger was used to perform simulation. After carrying out stepwise energy minimization and equilibration protocol of the solvated systems using TIP3P water model, a production simulation of 5ns at 300K temperature and 1 bar pressure was obtained for each of the docked complexes. The force field used for the simulation was OPLS4 [50]. In addition, a trajectory analysis was carried out to look at the Root Mean Square

[15] http://bioinfo3d.cs.tau.ac.il/PatchDock/

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0299.v1

Deviation (RMSD). Further, simulation was carried out at 50ns to achieve the final RMSD value for the best docked complex.

2.16. Prediction of immune response for the mice model

Considering the importance of pre-clinical trials, epitope screening for different mice alleles was performed using NetMHCIIpan 4.0 server using default parameters 1% and 5% for strong and weak binding epitopes, respectively.

3. Results

The study aims to identify the effective and conserved epitopes from LipL32, 41, 46 and OMPL 1 sequences of pathogenic *Leptospira* sp. and to construct multiepitope vaccine candidate against Leptospirosis using extensive immuno-informatics approach. The *in-silico* approach has been widely used for the development of epitope-based vaccine candidates against various pathogens as it has been proven rapid, cost-effective and more efficient as compared to conventional approaches [51]. In the present study, OMPL and LipL protein sequences from ten pathogenic *Leptospira* sp. have been identified, of which B-cell and T-cell epitopes have been predicted for designing the multiepitope vaccine constructs using appropriate linkers and adjuvant against Leptospirosis.

3.1. Retrieval of Leptospira surface protein sequences for epitope prediction

The first step of the designed workflow was to retrieve protein sequences in FASTA format. A total of forty-three sequences comprising two OMPL1, twenty-two LipL32, eighteen 41 and one LipL46 sequences from major ten pathogenic Leptospira sp. were obtained from the UniprotKB database.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0299.v1

Further B cell, HTL and CTL epitopes were predicted for making multi-epitope vaccine constructs. List of *Leptospira* species sequences along with their accession number is given in **Supplementary Table 1**. Further B cell, HTL and CTL epitopes were predicted for making multi-epitope vaccine constructs.

Table 1. B cell epitopes predicted using IEDB B-cell epitope prediction tool.

S.no	LIPL/OMP	predicted epitopes	length	total no.
1	LipL32	KKLLVRGLYRI	11	4
		GLPSLKSSFVLSE	13	
		KTLLPYGSVINYYGYVKPG	19	
		AYYLYVWIPAVIAE	14	
2	LipL41	AEAILYIGYQ	10	3
		DAVAAG	6	
		MLIPLDATLIKV	12	
3	LipL46	TTRLISALAVISFAVNCG	18	3
		NPRVAVLVLGKV	12	
		LIKILAEALQAEVLVLG	17	
4	OMPL1	TCTVGPSDPACFQN	14	4
		SSIVIPATVGIKLN	14	
		FWRVAAE	7	
		IGVAPRKA	8	

3.2. Linear B-cell epitope prediction

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0299.v1

The IEDB B-cell epitope prediction tool was used to determine B-cell epitopes from the specified lipoproteins. Recognizing and characterizing B-cell epitopes in target antigens is a crucial step in developing epitope-based vaccinations. The recognition of these epitopes by B lymphocytes elicit antibody production, which is very important for the adaptive immunity [52]. The Kolaskar and Tongaonkar approach was used to predict antigenic peptides by analyzing the physicochemical properties of amino acid residues and their abundance in experimentally determined antigenic epitopes [35]. The authors demonstrated that by applying this method to a large number of proteins, the system can predict antigenic determinants with approximately 75% accuracy, which is better than most other known methods (35). The analysis generated a total of 342 peptides. The analysis identified that a total of 14 B cell epitopes predicted were found to be conserved in all selected pathogenic species of Leptospira. Moreover, within predicted epitopes **KKLLV**RGLYRI, KTLLPY**GSVINYYGY**VKPG and **AEAILY**IGYQ, some short stretches of sequences were also predicted as CTL epitopes and hence can be considered as more effective epitopes. The total length of predicted epitopes was in the range of 6 to 19 amino-acids **(Table 1).**

3.3. T cell epitopes prediction

Prediction of conserved CTL epitopes is necessary for an effective vaccine design. The NetCTL 1.2 server predicts CTL epitopes by integrating predictions of proteasomal cleavage, TAP transport efficiency, and MHC class I binding. The alleles used for CTL and HTL epitope predictions were HLA *-DRB1_0101, DRB1_0102, DRB1_0103, DRB1_0301, DRB1_0305, DRB1_0401, DRB1_0402, DRB1_0403, DRB1_0404, DRB1_0405, DRB1_0408, DRB1_0701, DRB1_0801, DRB1_0803, DRB1_0901, DRB1_1001, DRB1_1101, DRB1_1104, DRB1_1201, DRB1_1301, DRB1_1302, DRB1_1303, DRB1_1401, DRB1_1402, DRB1_1454, DRB1_1501, DRB1_1503, DRB1_1601, DRB3_0101,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0299.v1

DRB3_0202, DRB4_0101, DRB4_0103, DRB5_0101, DRB5_0202. Total 244 CTL and 389 HTL epitopes were predicted using the online tools. A total of 14 epitopes of 9 amino-acids length were found to be conserved in all selected pathogenic strains of Leptospira as shown in **Table 2**.

Table 2.CTL epitopes predicted from NetCTL 1.2 server.

S.no	LIPL/OMP	Predicted epitopes	Length	Total no.
1	LipL32	FTTYKPGEV	9	4
		GSVINYYGY	9	
		LVDGNKKAY	9	
		NIDTKKLLV	9	
2	LipL41	LIGAEAILY	9	4
		YTECSTENK	9	
		STENKIDAV	9	
		AKANLATYY	9	
3	LipL46	TSDTNGSEW	9	2
		STAVDNLLA	9	
4	OMPL1	LTAAGAGAV	9	4
		VTKADIAGY	9	
		YTQKISGGV	9	
		ESDFGKYFF	9	

The NetMHCIIpan 4.0 server was used to determine the Helper T lymphocyte (HTL) epitopes. Multiple epitopes were predicted against the selected alleles for each protein sequence

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0299.v1

in which a total of 12 epitopes of 15 amino-acids length were found to be conserved in all selected pathogenic species of *Leptospira*. The predicted HTL epitopes are listed in **Table 3**.

From results of predicted B cell and HTL Epitopes, **YYLYVWIPAVIAE**MG, **SSFVLSE**DTIPGTNE, **INYYGY**VKPGQAPDG, KPTGVRM**MLIPLDAT**, and EGNY**IGVAPRKA**IPA were found to be common. Whereas within T cell epitopes, **INYYGY**VKPGQAPDG and KGGYDIL**TAAGAGAV** sequences were similar which suggest them as promising protein vaccine candidates.

Table 3. HTL epitopes predicted from NetMHCIIpan 3.1 server.

S.no	LIPL/OMP	Predicted epitopes	Length	Total no.
1	LipL32	YYLYVWIPAVIAEMG	15	4
		MSAIMPDQIAKAAKA	15	
		SSFVLSEDTIPGTNE	15	
		INYYGYVKPGQAPDG	15	
2	LipL41	RMMLIPLDATLIKVE	15	4
		RIKVFVKDEDEEVKE	15	
		KPTGVRMMLIPLDAT	15	
		VKKAVVSSPAKIFNS	15	
3	LipL46	DRGFSFLATKSLEVK	15	2
		GFSFLATKSLEVKSA	15	
4	OMPL1	KGGYDILTAAGAGAV	15	2
		EGNYIGVAPRKAIPA	15	

3.4. Conservancy of predicted epitopes

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0299.v1

Conservancy of all the predicted epitopes was found in the range of 26.67 to 100% for ten selected pathogenic *Leptospira* species. After the similarity search, epitopes GLPSLKSSFVLSE, AEAILYIGYQ, LIGAEAILY, NPRVAVLVLGKV, TSDTNGSEW, STAVDNLLA, AEAILYIGYQ and LIGAEAILY show 100% conservancy amongst selected pathogenic Leptospira species for the study. The conservancy results and exact length for each predicted epitope is summarized in **Supplementary Table 2**.

3.5. Population coverage of selected epitopes

All the selected epitopes were then analyzed to determine the percentage of the population coverage for the following regions – South Asia, South America, South Africa and World for selected MHC-I and MHC-II alleles. The coverage for these alleles was found to be 93.03% for South Africa, 88.3% for South America, 94.73% for South Asia and 98.55% for World, thus indicating high population coverage for all selected epitopes. Results for population coverage for the world are presented in graphical form in Figure 2.

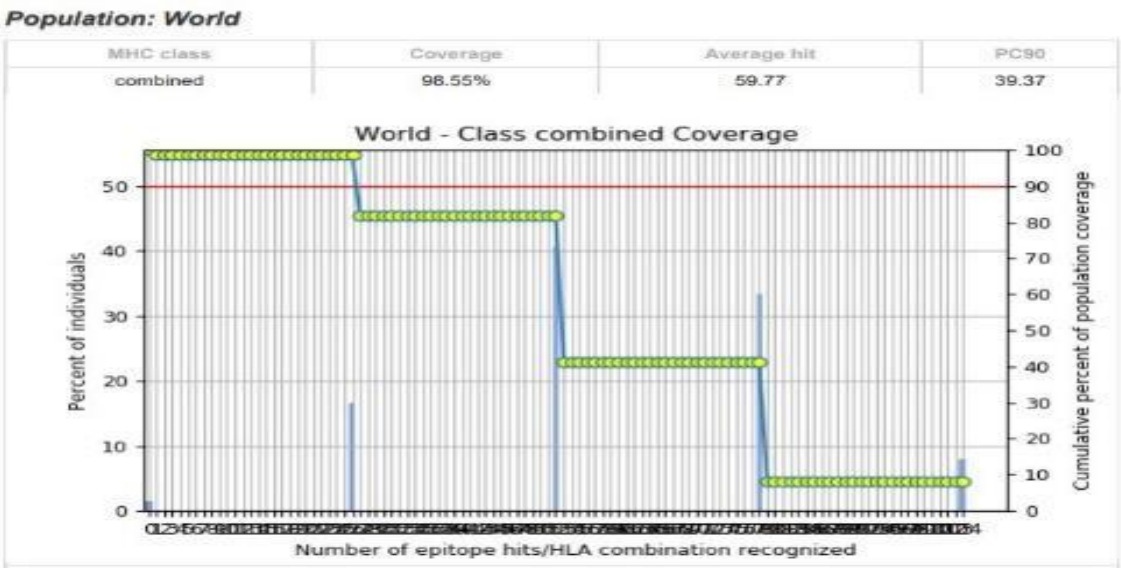


Figure 2. Population coverage for the MHC-I and MHC-II alleles for all selected epitopes.

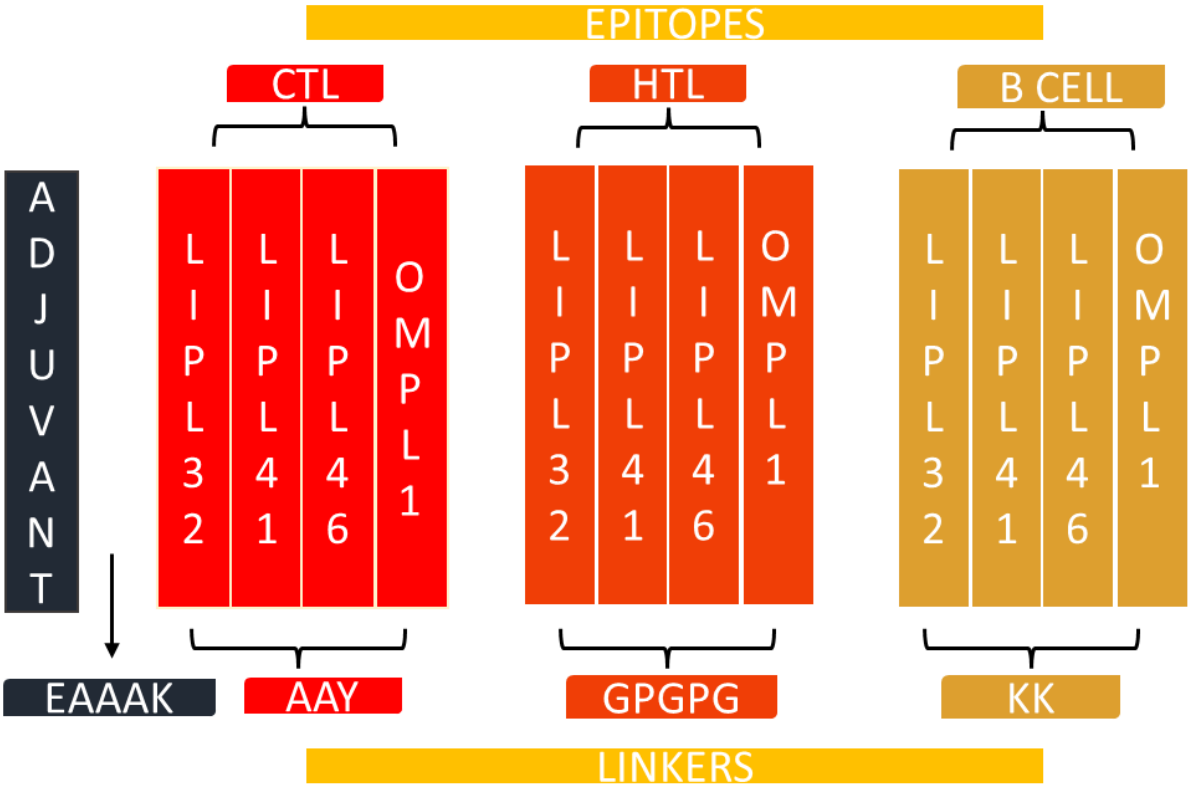
3.6.Construction of Multi-epitope sub-unit vaccine

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0299.v1

Multi-epitope or sub-unit vaccines were constructed by joining all the best predicted B cell, HTL and CTL epitopes with the suitable linkers. The N terminal portion of the vaccine construct starts with a TLR 4 sequence– APPHALS, added as an adjuvant to increase the immunogenicity of the vaccine construct and joined with the help of the linker EAAAK. B cell epitopes were fused together by KK linker, CTL epitopes were joined by AAY linkers and the HTL epitopes were linked by GPGPG linkers. A total of 6 constructs were made by shuffling the B cell, CTL and HTL epitopes. The total length of the designed vaccine constructs was found out to be 618 to 620 amino-acids. All these 6 constructs were further assessed as per the workflow of the study. The arrangement of one such construct and basic design of the multi-epitope vaccine construct is shown in Figure 3.

A)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0299.v1



B)

Figure 3. A) sequence of the vaccine construct (red – adjuvant; blue – linker for adjuvant; pink – B cell epitopes; green – CTL epitopes; purple – HTL epitopes); **B)** schematic presentation of multi-epitope vaccine construct.

3.7. Prediction of allergenicity, antigenicity, toxicity & immunogenicity of the vaccine construct

]All designed vaccine constructs were subjected to the prediction of allergenicity, antigenicity, toxicity and immunogenicity. Construct 6 exhibited the highest antigenicity score with the value of 0.6021, whereas construct 3 exhibited highest immunogenicity of -3.043All the constructs were found to be non-allergenic, probable antigenic, non-toxic and immunogenic in nature. The result for allergenicity, antigenicity, toxicity and immunogenicity of all 6 constructs are summarized in **Table 4**.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0299.v1

Table 4. Results for allergenicity, antigenicity, toxicity and immunogenicity prediction for vaccine constructs.

Construct no	Allergenicity	Antigenicity	Toxicity	Immunogenicity
1	Non-allergenic	0.5825	Non-toxic	-3.014
2	Non-allergenic	0.5846	Non-toxic	-2.961
3	Non-allergenic	0.5988	Non-toxic	-3.043
4	Non-allergenic	0.6021	Non-toxic	-2.952
5	Non-allergenic	0.6014	Non-toxic	-2.560
6	Non-allergenic	0.6069	Non-toxic	-2.546

3.8. Physio-chemical analysis of vaccine constructs

The various physio-chemical properties such as number of amino acids, molecular weight, isoelectric point, aliphatic index, stability index, grand average of hydropathy of the vaccine constructs were assessed using the Protparam server. The length of the constructs formed was found to be 618 to 620 amino-acids with average molecular weight of 64637.015 Daltons. The instability index was computed to be in the range of 18.22 to 18.81, implying that all the constructs formed are stable in nature. The aliphatic index was calculated in the range of 87.61 to 88.22 for all the constructs. A protein that has a high aliphatic index is thermostable across a wide temperature range. All the constructs made were thus found to be thermostable in nature. The Grand average of hydropathicity (GRAVY) was calculated to be positive, implying that all constructs formedare hydrophobic. The results for physio-chemical properties of all 6 multi-epitope vaccine constructs are summarized in **Table 5**.

Table 5. Physio-chemical properties for the vaccine constructs.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0299.v1

Construct no	No of amino acids	Molecular weight (daltons)	Isoelectric point	Aliphatic index	Instability index	Gravy
1	618	64662.24	9.57	88.22	18.63 STABLE	0.012
2	618	64662.24	9.57	88.22	18.63 STABLE	0.012
3	620	64722.29	9.57	87.61	18.22 STABLE	0.001
4	619	64559.12	9.58	87.75	18.81 STABLE	0.003
5	620	64608.10	9.54	87.94	18.36 STABLE	0.020
6	620	64608.10	9.54	87.94	18.80 STABLE	0.020

3.9. Secondary structure prediction

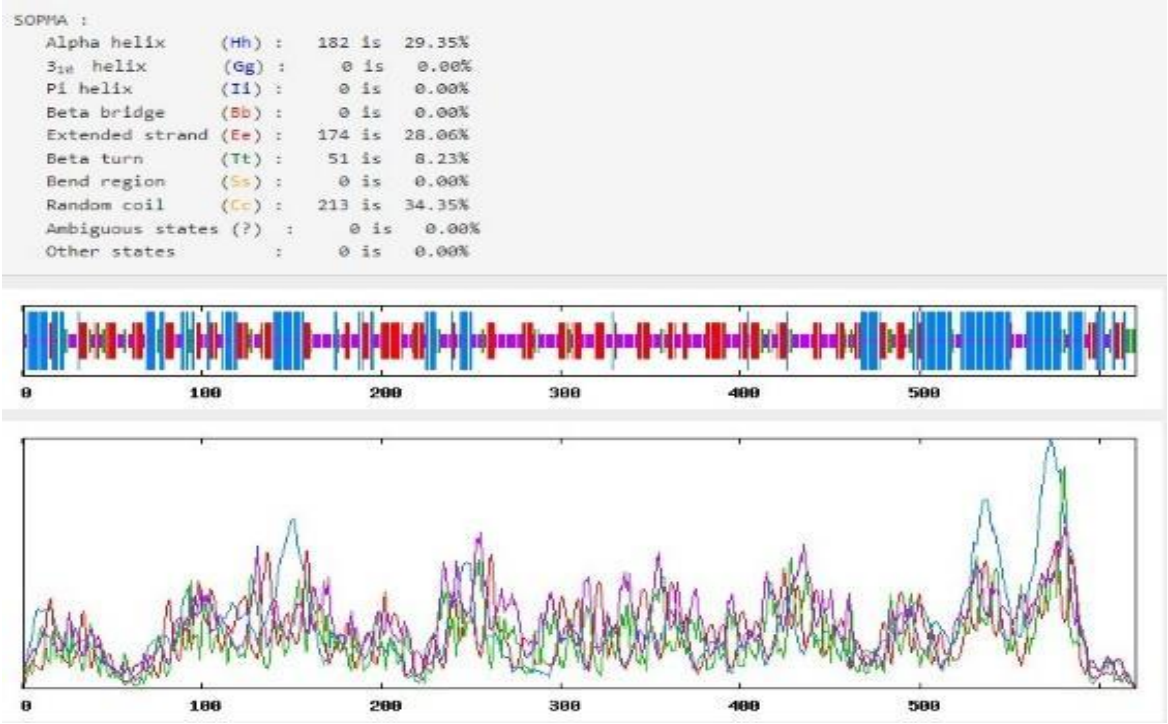
The secondary structure of the vaccine constructs was predicted using the SOPMA server. This server predicts the percentage of alpha helix, extended strand, beta turn and random coil of the submitted vaccine construct sequence. Figure 4 depicts the result obtained for secondary structure prediction of 4th construct. The calculated percentage of alpha helix (28.76 to 29.52), extended strand (27.74 to 28.64), beta turn (7.10 to 8.25) and random coils (28.76 to 35.65) in all 6 constructs is presented in **Table 6**. Since all constructs have more percentage of alpha helix than beta sheets, it can be concluded that all are stable in nature as stability of the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0299.v1

secondary structure of the protein is associated with the alpha helix formation. Constructs 4 and 5 show equal and highest percentage of alpha helix.

Table 6. Secondary structure indicating percentage of alpha helix, extended strand, beta turn and random coil for all 6 constructs.

Construct no.	Alpha helix	Extended strand	Beta turn	Random coil
1	29.45	27.83	8.25	34.07
2	29.13	28.64	7.44	28.76
3	29.03	28.06	7.58	35.32
4	28.76	27.95	7.75	35.54
5	28.76	27.95	7.75	35.54
6	29.52	27.74	7.10	35.65



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0299.v1

Figure 4. Secondary structure prediction using SOPMA server for 4th construct, stating the percentage of alpha helix, extended strand, beta turn and random coils. (Blue - alpha helix; red – extended strand; green – beta turn; yellow- random coil)

3.10. Tertiary structure modeling

For tertiary structure, Robetta server (ab-initio method) was used. Sequence of the vaccine construct was given as input with ab-initio method (AB) used as optional parameters and 5 models of 3d structure were generated in output along with the confidence score. The confidence score lies in the range of -5 to 2, where high value denotes a model with higher confidence and correct topology [37]. Figure 5 represents tertiary structure generated for 4th vaccine construct. Later the top ranked model given by the robetta server is refined using Galaxy- Refine server.

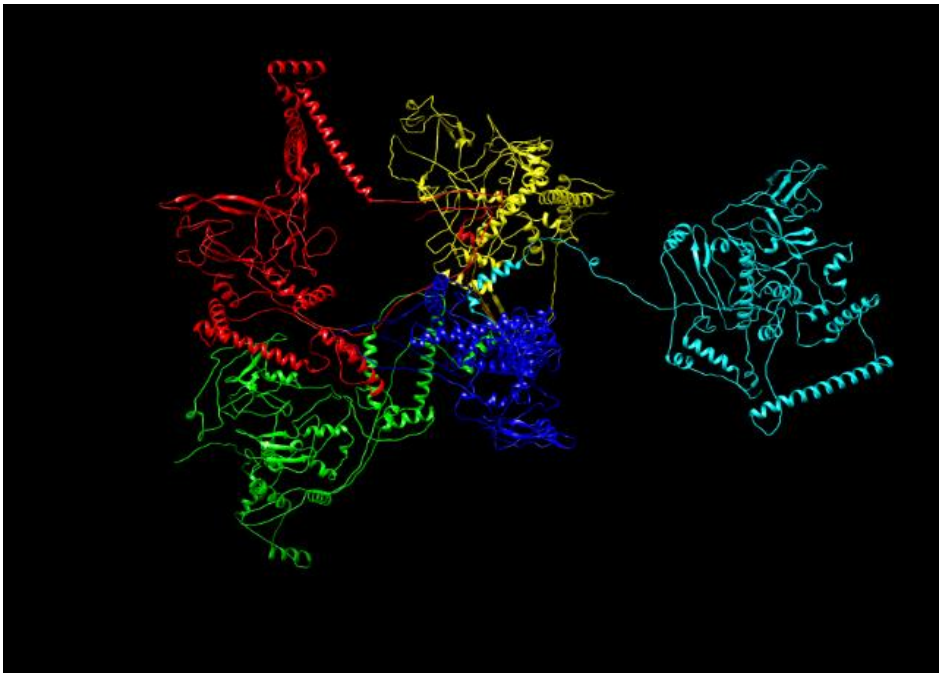


Figure 5. Tertiary structure modelling performed using Robetta server (ab-initio method) for 4th vaccine construct.

3.11. Tertiary structure refinement

Refinement of the tertiary structure is a crucial step as it improves the overall quality of the tertiary structure generated. Through conformational sampling, protein structure refinement

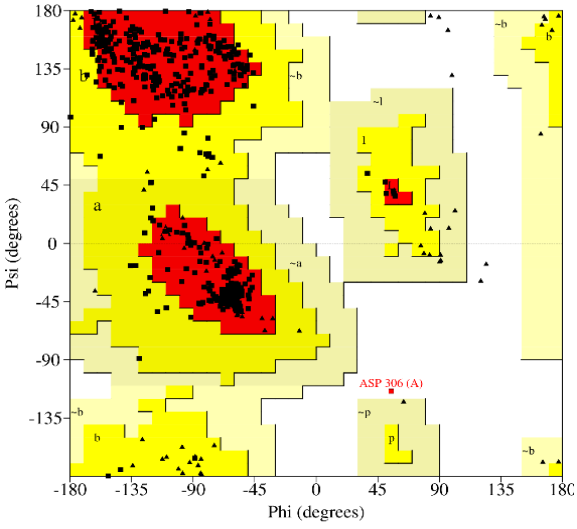
Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0299.v1

tries to get moderately accurate template-based protein models closer to the native state. For refining the tertiary structure of the vaccine constructs, Galaxy-Refine server was used. The server gives five refined models stating their GDT-HA, RMSD, MolProbity, Clash score and Poor rotamers and percentage of residues in Rama favored regions of Ramachandran Plot.

3.12. Validation of refined tertiary structure

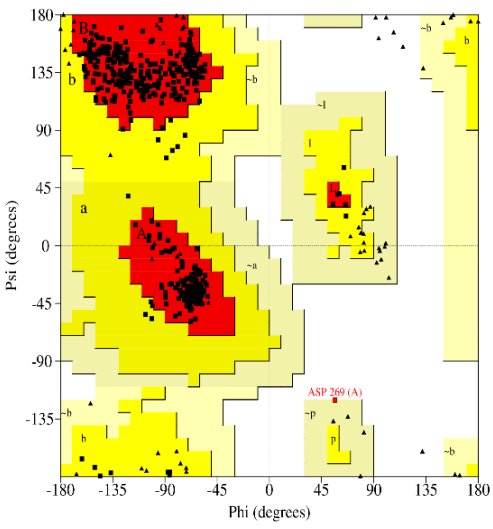
Procheck and ERRAT were used to validate the refined tertiary structure. Procheck gives the Ramachandran plot, which explains the statistical distribution of the combinations of the backbone dihedral angles ϕ and ψ. The Ramachandran plot can be used to assess the precision of the predicted protein structure. It describes the percentage of residues falling in the favored regions, allowed regions and disallowed regions **(**Figure 6A**)**. For all the constructs the percentage of residues falling in allowed regions were 95.3 to 97.5 %, additionally allowed regions were 2.2 to 4.5 % and dis-allowed regions were up to 0.2%. It is expected that no residue should be found in the disallowed or outlier region, and no more than 2% of residues should be in the allowed region. All constructs met the requirements, hence were found to be stable and could be used for further analysis. Vaccine construct 4 showed 95.3% residues in most favored regions of Ramachandran plot. ERRAT predicts the overall model quality and attempts to determine any potential errors in the structure **(**Figure 6B**)**. Regions of the ERRAT graph that are coloured red and yellow stand for the problematic parts of the structure, while white stands for the normal parts. The plot analysis quickly identifies residues with error values greater than 95% and 99%. The ERRAT value was found to be in the range of 90 to 95. The Prosa-Web (z-score) value indicates the model quality in terms of z score and value ranged from -6.34 to -4.14. Value of z score in the range of -6.0 represents good quality of the model. The Procheck, Prosa-Web and ERRAT result obtained for the validation of the refined structures of the vaccine constructs is summarized in **Table 7 and 8;** Figure 6A, B and C.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0299.v1



Plot statistics

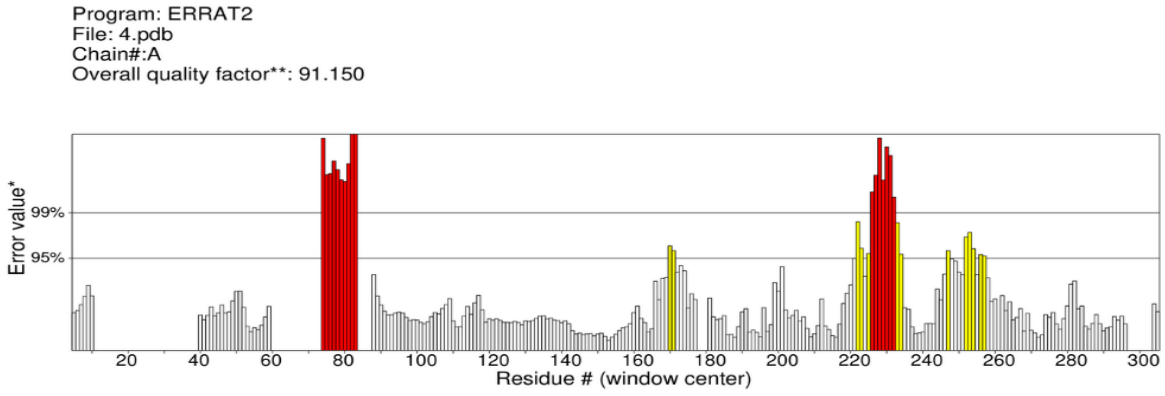
Residues in most favoured regions [A,B,L]	459	92.5%
Residues in additional allowed regions [a,b,l,p]	36	7.3%
Residues in generously allowed regions [~a,~b,~l,~p]	0	0.0%
Residues in disallowed regions	1	0.2%
Number of non-glycine and non-proline residues	496	100.0%
Number of end-residues (excl. Gly and Pro)	1	
Number of glycine residues (shown as triangles)	75	
Number of proline residues	48	
Total number of residues	620	



Plot statistics

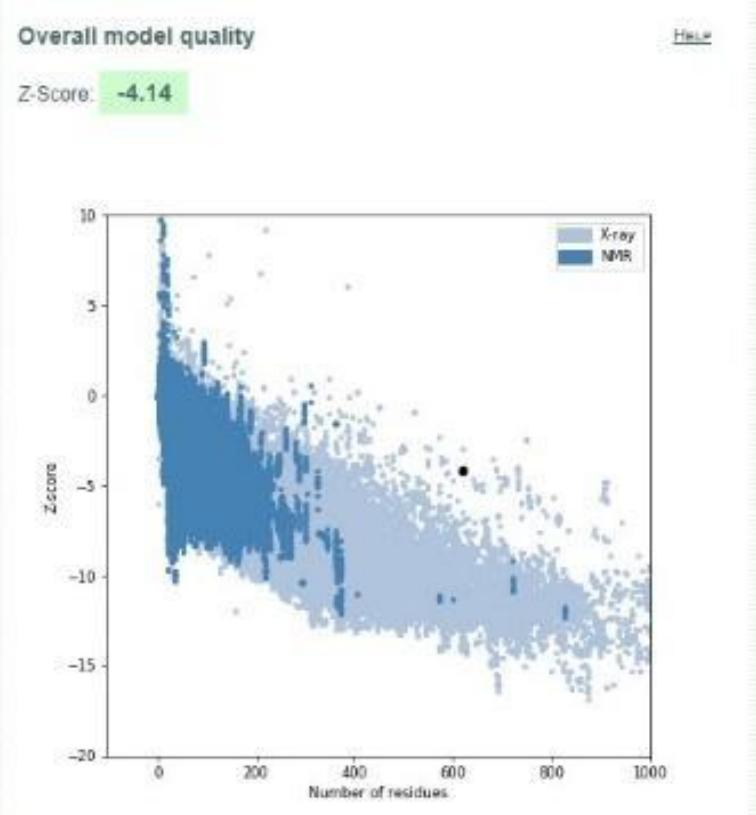
Residues in most favoured regions [A,B,L]	471	95.3%
Residues in additional allowed regions [a,b,l,p]	22	4.5%
Residues in generously allowed regions [~a,~b,~l,~p]	1	0.2%
Residues in disallowed regions	0	0.0%
Number of non-glycine and non-proline residues	494	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	75	
Number of proline residues	48	
Total number of residues	619	

A.1) Ramachandran plot for unrefined tertiary structure of the 4th construct
A.2) Ramachandran plot for refined tertiary structure of the 4th construct



B)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0299.v1



C)

Figure 6.A.) Ramachandran plot for 4th vaccine construct **B)** ERRAT graph showing the error value for the 4th vaccine construct **C)** Prosa-web (z-score) showing the overall quality of the 4th construct.

Table 7. Procheck validation results of the refined tertiary structures for all vaccine constructs

S.NO	residues in favoured regions		residues in additionally allowed regions		residues in dis-allowed regions	
	Before refinement	**After refinement**	**Before refinement**	**After refinement**	**Before refinement**	**After refinement**
1	92.8	96%	7.8%	4.0%	0.0%	0.0%
2	93.2%	97%	6.6%	3.0%	0.0%	0.0 %
3	92.1%	97.6%	7.7%	2.2%	0.2%%	0.2%
4	92.5%	95.3%	7.3%	4.5%	0.2%	0.0%
5	94.1%	96.2%	5.5%	3.6%	0.2%	0.2%
6	95.4%	96.8%	4.6%	3.2%	0.0%	0.0%

Table 8. ERRAT validation results of the refined tertiary structures for all vaccine constructs.

S.NO	ERRAT
1	90.047

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0299.v1

2	91.927
3	91.162
4	91.150
5	90.646
6	95.165

3.13. Molecular docking of subunit vaccine construct with different immune receptors

After the validation, docking studies were performed in which the vaccine constructs were docked with different receptors (mentioned in the methods section). Docking was performed using Patchdock server to check the interaction between the receptor and the constructs. The docked structures were further refined using Firedock server and the global energy for the 10 top ranked refined docked structures was obtained for each construct. Further, the docked complex was visualized using Chimera software [53] in Figure 7. The construct 2 ranked top amongst docked structures with MHC I receptor based on global energy (-32.91), Attractive Van Der Waals interactions (-23.49), Atomic contact energy (-1.53) and Hydrogen bonding (-3.51) (**Table 9)**.

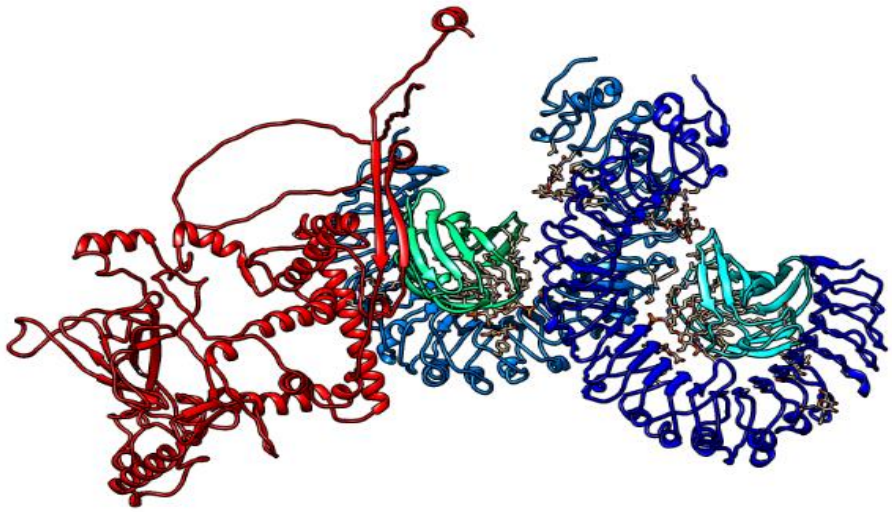
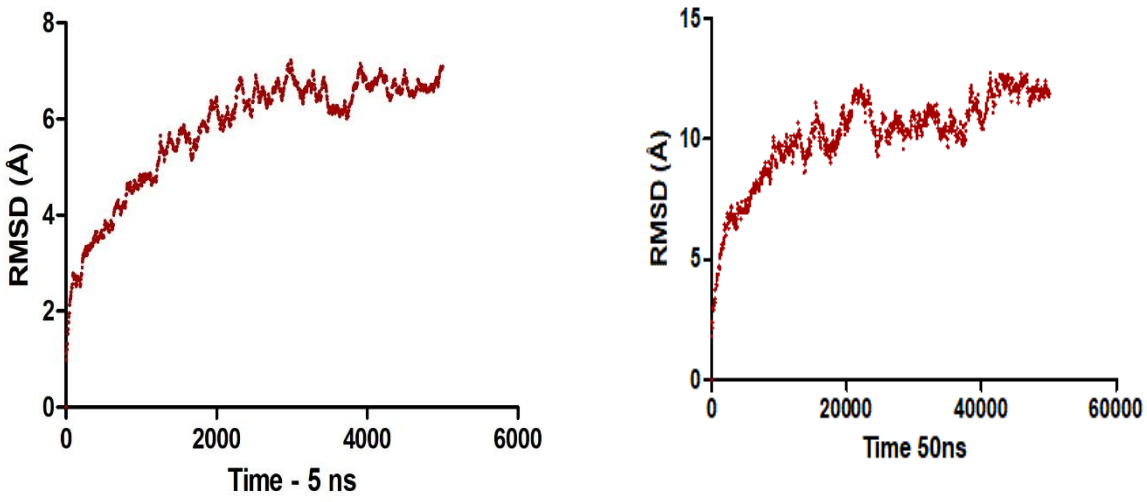


Figure 7. Docked complex (vaccine construct 4 with receptor TLR 4) visualized in chimera

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0299.v1



8.A) **8.B)**

Figure 8.A) RMSD graph obtained at 5 ns for construct 4 with receptor TLR 4; **B)** RMSD graph obtained at 50 ns for construct 4 with receptor TLR 4

Table 9. Results for docking of the vaccine constructs and the receptors.

Construct no.	Receptor	Global energy	Attractive vdw	Repulsive vdw	Ace	Hb
1	TLR 2	4.79	-0.01	0.00	0.10	0.00
2	TLR2	-0.90	-20.43	11.56	13.75	-1.36
3	TLR2	3.71	-10.79	1.19	5.92	-1.12
4	TLR2	-10.25	-19.25	11.54	-0.38	-0.95
5	TLR2	0.36	-14.18	20.73	1.94	-1.06
6	TLR2	-5.43	-11.39	9.87	3.37	0.31
1	TLR4	-14.56	-19.71	9.95	9.74	-2.82

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0299.v1

2	TLR4	-11.94	-24.36	16.95	8.32	-3.84
3	TLR4	-0.54	-31.49	19.31	13.23	-3.74
4	TLR4	-12.07	-31.73	18.43	11.03	-2.58
5	TLR4	1.05	-1.90	0.0	2.04	0.0
6	TLR4	-30.29	-29.70	11.47	8.67	-1.47
1	MHC1	-3.80	-2.02	0.00	1.10	-0.44
2	MHC1	-32.91	-23.49	10.14	-1.53	-3.51
3	MHC1	-19.08	-32.24	14.44	7.72	-2.67
4	MHC1	-5.28	-15.21	1.53	1.82	-1.93
5	MHC1	3.17	-4.79	1.31	3.72	-0.72
6	MHC1	-1.0	-14.78	10.10	1.56	-1.32
1	MHC2	-13.99	-21.67	12.43	7.91	0.00
2	MHC2	-0.18	-5.38	3.87	3.43	-1.94
3	MHC2	2.09	-2.95	0.41	1.12	-0.21
4	MHC2	-8.89	-7.19	2.24	-1.01	-0.26
5	MHC2	-3.04	-3.76	1.18	-3.60	0.00
6	MHC 2	5.43	-2.10	0.00	0.76	0.00

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0299.v1

(vdw – vanderwalls ; ACE – Atomic Contact Energy ; Hb – Hydrogen bonding)

3.14. Binding affinity analysis

The Prodigy (PROtein binDIng enerGY prediction) server was used to predict binding affinity which is based solely on a protein-protein complex's structural properties. In summary, construct 1 showed best binding affinity with the receptors TLR 4, TLR 2, MHC-I and MHC-II. The predicted binding free energy ΔG in kcal mol^{-1} is included in the results; binding affinity was predicted for all docked complexes and is presented in **supplementary Table 3**.

3.15. Molecular dynamic simulations

On the basis of docking and binding affinity results obtained, construct 4 was further subjected to simulation studies. Vaccine construct no. 4 with receptor TLR 4 was simulated at 5 nanoseconds using the TIP3P water model; NPT ensemble; cubic geometry and Na+ Cl- ions for neutralizing the environment. RMSD obtained at 5ns was 7 Å. Later the same docked complex was simulated at 50 ns to check the stability at the similar conditions and RMSD was found to be 12 Å. The results for simulation studies are presented in Figure.8.

3.16. Prediction of immune response for the mice model

Along with the epitope screening for human alleles, it is also necessary to screen the epitopes for mice alleles to predict their immune response for multi-epitope vaccine construct designed. A total of 243 strong binding epitopes were predicted for all the selected mice alleles shown in **supplementary table 4**. On the basis of the screening it can be predicted that the designed multi-epitope vaccine construct is capable of generating effective immune response even in mice as a number of epitopes for the selected mice alleles were found to be overlapping with the vaccine construct.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0299.v1

4. Discussion

The global incidence of leptospirosis is growing year by year. Vaccination is one of the most effective ways to promote public health in a cost-effective, timely, and affordable manner, as well as the best strategy to control infectious disease in society. The traditional vaccine development methodology is a time-consuming and ineffective method that costs a lot of money. Vaccine development using a revolutionary immuno-informatics approach, on the other hand, is more stable, safe, specific, efficacious, and time and money efficient [54]. Even with the availability of massive immuno-informatics tools, preventing the zoonotic disease Leptospirosis remains difficult. Targeting bacterial species soft-core or core genes could aid in the identification of improved vaccine candidates to combat infectious illness [55]. Hence, in the present study, outermembrane proteins were targeted for the vaccine construct using immuno-informatics approach.

Subunit vaccines, as opposed to whole organism vaccines, are the focus of current research because subunit vaccines contain specific immunogenic components of the pathogens rather than the entire pathogenic entity. OMPs, lipoproteins, and transmembrane proteins are all identified vaccine targets as they are present on the surface of leptospires and are considered as key virulence factors that interact with host tissues and responsible for immune response [56].

In this study, we have tried to prepare multi-epitope vaccine constructs using outermembrane proteins and lipoproteins obtained from ten different common pathogenic strains of *Leptopsira* by screening B cell and T cell (HTL and CTL) epitopes. . This is the first study undertaken with multiple pathogenic species of *Leptospira* for designing vaccine constructs using four specific proteins. A total of 43 sequences comprising LipL 32, 41, 46 and OMPL 1 sequences from pathogenic species of *Leptospira- viz. L. interrogans , L. kischneri, L. noguchii, L.*

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0299.v1

borgerpeterseni, L. alexanderi L. santarosai , L.weilli, L. kmetyi, L.mayottensis , L.alstonii were selected for the study.

In diverse preceding studies, a single epitope vaccine has been proposed towards leptospirosis. However, our research differs from theirs in terms of the prediction and choice of multiple epitopes and consequently the potential of the vaccine construct to cause a more impregnable immune reaction in-silico. When it comes to creating a prolonged, extensive immune reaction, the involvement of each T cells and B cells is vital to set off and concatenate each the cellular and humoral immunity respectively. Another factor of selecting kinds of epitopes is that often, antigens can get away memory B cells, wherein case, their detection stays covert. In such cases, the contingency plan could be to have the T cells understand the ones overlooked antigens or epitopes and neutralize them so that no remains of the pathogen are found.

To elicit a more distinct immunological response from T cells, we predicted both CD8+ (MHC-I) and CD4+ (MHC-II) epitopes to have a greater variety of possible epitopes. We exclusively predicted linear epitopes for B cell epitope prediction since they are more stable than conformational epitopes. A total of 14 B cell and 26 T cell (12 HTL and 14 CTL) epitopes were selected. We evaluated conservancy as a major factor in addition to antigenicity, allergenicity, and toxicity when choosing the epitopes from the anticipated results because conserved sequences can enhance innate immunity by triggering improved pattern recognition.

The predicted immunogenic epitopes against proteins, LipL32, LipL41, LipL46 and OmpL1, were found to be conserved amongst selected ten pathogenic species of *Leptospira*. Out of ten selected, seven of these species: *L. interrogans*, *L. borgpetersenii*, *L. santarosai*, *L. noguchii*, *L. weilli*, *L. kirschneri* and *L. alexanderi* are the main agents of leptospirosis [57] and the remaining three have also been found to be the causative agents of leptospirosis. Three major species (*Leptospira borgpetersenii*, *L. interrogans* and *L. kirschneri*) are predominant in

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 December 2022 doi:10.20944/preprints202212.0299.v1

reports from Africa and isolates from a diverse range of serogroups have been reported in human and animal infections. [58]. Approximately half of the pathogenic serovars belong to *L interrogans* or *L. borgpetersenii* [59]. The interaction analysis of these epitopes with HLA alleles indicated that the multiple epitope vaccine constructs against leptospirosis will cover 98.55% of the world population. The population coverage for the epitopes we chose was 93.03% in South Africa, 88.3% in South America, and 94.73% in South Asia. A 98.55% coverage of the world's population was also attained.

When used alone, epitope-based peptide vaccinations generate a relatively mild immune response (60). Appropriate adjuvants could be used to increase the immunoreactivity. Linkers in the vaccine construct play a vital role in antigen processing, effective separation by the host immune system and prevent junctional epitope formation (61). In this study, B cell, CTL and HTL epitopes were connected through linkers KK, AAY and GPGPG respectively. The multi epitope constructs purely cannot amplify superior immune response. Hence, it was coupled with an adjuvant APPHALS with EAAAK linker at the N-terminal of the construct. An adjuvant helps in protection against infections, enhance the stability, durability of the vaccine formulation, and also extends the vaccine ability to induce strong immune response without an allergic reaction towards humans. Linkers will cause the efficient separation, stimulation of HTL response and prevention of vaccine interactions with its own sections (62).

The obtained six multi-epitope vaccine constructs were 618 to 620 amino acids length with 64.63 kDa average molecular weight. All the constructs showed positive antigenicity score (vaxijen score), which indicates the vaccine constructs were highly antigenic in nature. All constructs were found to be non-allergenic, nontoxic and immunogenic in nature. For all the constructs, physio-chemical properties were found to be satisfactory and stable. The reliability of the vaccine model was high as suggested by the Ramachandran plot wherein

 doi:10.20944/preprints202212.0299.v1

majority of the residues fell in the most favorable region in the range of 95.3 % to 97.6 % with the overall model quality ranging from 90 to 91.9 indicated by ERRAT .

Further, molecular docking studies revealed stable interactions between multi-epitope vaccine constructs and different receptors using PatchDock server. Negative binding energy for all docked complexes indicates positive binding interactions between vaccine and receptor. Construct no. 4 with TLR 4 receptor had less energy (-12.07) which was needed for proficient binding. The binding of multi-epitope vaccine with Toll-Like Receptors 4 (TLR4) implies that it might play an important role in the innate immune response. Later to substantiate the docking results, molecular dynamic simulation studies were performed. During the simulation analysis using Schrodinger, 300K temperature and 1 bar pressure was maintained in the system. Water molecules and salts addition was carried out to neutralize the system. To evaluate the stability of vaccine constructs interaction with immune receptors, RMSD was observed for 5 and 50 ns. The RMSD plot versus time displayed mild fluctuation after 2.5 ns for Construct no. 4 with TLR 4 receptor. The results are comparable with the docking of vaccine constructs with a receptor reported in previous study done on novel multi-epitope vaccine against SARS-COV-2 using immunoinformatics approach in 2021. This was the best docked model as inferred from global energy score (-12.07) indicating the proficient binding with protein vaccine and presented high stability of vaccine construct and TLR4 interaction with RMSD value of 12 Å. Overall construct no.4 appears to elicit both innate and adaptive immune response against most common pathogenic species of *Leptospira*. The epitopes selected for the vaccine construct were found to be overlapping with the epitopes predicted for the selected mice alleles. Thus, it can be concluded that the designed vaccine construct is capable of eliciting effective immunization in mice. Many reports have applied a similar approach to design multi epitope vaccine candidates against cancer, SARS-CoV-2, dengue, malaria and other life-threatening infections (34,60,61,63–65). Multi-epitope vaccines are extensively designed in cutting edge

 doi:10.20944/preprints202212.0299.v1

immunoinformatics studies due to high immunogenicity, stability, specificity, least production cost and in vivo efficacy.

In the present study, potential epitopes against Leptospirosis were identified by using immunoinformatics approach. With promiscuous CTL, HTL, and B-cell epitopes and appropriate adjuvants, protein vaccines were designed. The computational analysis for physicochemical and antigenicity of vaccine constructs were performed. The molecular interactions and their stability between vaccine constructs and immune receptors were evaluated by molecular docking and dynamic simulation. The findings of the study suggest that the designed multi-epitope vaccine construct can induce an optimal protective immune response against most commonly prevailing pathogenic species of *Leptospira* infection to provide broad range protection taking a lead towards development of universal vaccine. Further, in vivo studies addressing safety and efficacy of the vaccine candidate would provide proof of concept for developing a vaccine candidate.

Supplementary Materials: **Table 1.** Lipoprotein sequences of different *Leptospira* species retrieved from Uniprot and their accession number. **Table 2.** Conservancy results for the predicted epitopes of B cell, CTL and HTL. **Table 3.** Binding affinity analysis for all docked complexes. **Table 4**. Epitope prediction against selected mice alleles.

Author Contributions: Conceptualization, C.G.J. (Chaitanya G. Joshi) and A.P. (Amrutlal Patel); Methodology, A.P. (Amrutlal Patel), S.S. (Sataymitra Shekh), K.K. (Kopal Kapoor); Software, K.K. (Kopal Kapoor); Validation, A.P. (Amrutlal Patel), K.K. (Kopal Kapoor); Formal Analysis K.K. (Kopal Kapoor); Writing-original draft preparation, K.K. (Kopal Kapoor); Writing-review and editing, K.K. (Kopal Kapoor); Supervision, A.P. (Amrutlal Patel). All authors have read and agreed to the published version of the manuscript.

Funding: The research work is supported by the Department of Science and Technology, Government of Gujarat. (Grant number: GBRC/GoG/DST/JD-2/BS14/2017).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in the study are included in this article and supplementary material. Further inquiries can be directed to the corresponding authors.

Acknowledgments: We would like to express our sincere gratitude to the Department of Science and Technology, Govt. of India for the support and granting facilities for research work.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] Z. May, P. Soo, and N. A. Khan, "Acta Tropica Leptospirosis : Increasing importance in developing countries," vol. 201, no. September 2019, 2020, doi: 10.1016/j.actatropica.2019.105183.

[2] P. N. Levett and P. N. Levett, "Leptospirosis Leptospirosis," vol. 14, no. 2, 2001, doi: 10.1128/CMR.14.2.296.

[3] F. Costa *et al.*, "Global Morbidity and Mortality of Leptospirosis : A Systematic

Review," pp. 0–1, 2015, doi: 10.1371/journal.pntd.0003898.

[4] C. Donati and R. Rappuoli, "Reverse vaccinology in the 21st century : improvements over the original design," pp. 1–18, 2013, doi: 10.1111/nyas.12046.

[5] M. B. Heinemann, S. A. Vasconcellos, E. C. Romero, and L. T. O. Ana, "Evaluation of two novel leptospiral proteins for their interaction to human host components," no. 5511, 2016.

[6] A. L. T. O. Nascimento *et al.*, "Comparative Genomics of Two Leptospira interrogans Serovars Reveals Novel Insights into Physiology and Pathogenesis †," vol. 186, no. 7, pp. 2164–2172, 2004, doi: 10.1128/JB.186.7.2164.

[7] A. M. Khan *et al.*, "A systematic bioinformatics approach for selection of epitope-based vaccine targets," *Cell. Immunol.*, vol. 244, no. 2, pp. 141–147, 2006, doi: 10.1016/j.cellimm.2007.02.005.

[8] and J. M. V. Ajay R Bharti, Jarlath E Nally, Jessica N Ricaldi, Michael A Matthias, Monica M Diaz, Michael A Lovett, Paul N Levett, Robert H Gilman, Michael R Willig, Eduardo Gotuzzo, "Leptospirosis: a zoonotic disease of global importance," *Lancet Infect Dis*, vol. 3, no. December, pp. 176–180, 2003.

[9] S. A. Plotkin, "Vaccines: Past, present and future," *Nat. Med.*, vol. 11, no. 4S, p. S5, 2005, doi: 10.1038/nm1209.

[10] J. B. Kathiriya, S. H. Sindhi, and S. N. Ghodasara, "Leptospirosis in India: a Veterinary Perspective," *Ijset.Net*, no. April, 2016, [Online]. Available: www.ijset.net

[11] Y. Peters, A., Vokaty, A., Portch, R., & Gebre, "Leptospirosis in the Caribbean: a literature review," *Pan Am. J. public Heal.*, 2017, [Online]. Available: https://doi.org/10.26633/RPSP.2017.166

[12] A. A. Grassmann, J. D. Souza, and A. J. A. McBride, "A universal vaccine against leptospirosis: Are we going in the right direction?," *Front. Immunol.*, vol. 8, no. MAR,

 doi:10.20944/preprints202212.0299.v1

pp. 1–8, 2017, doi: 10.3389/fimmu.2017.00256.

[13] C. Sonrier, C. Branger, V. Michel, N. Ruvoën-Clouet, J. P. Ganière, and G. André-Fontaine, "Evidence of cross-protection within Leptospira interrogans in an experimental model," *Vaccine*, vol. 19, no. 1, pp. 86–94, 2000, doi: 10.1016/S0264-410X(00)00129-8.

[14] O. A. Dellagostin *et al.*, "Recombinant vaccines against Leptospirosis," vol. 8600, 2011, doi: 10.4161/hv.7.11.17944.

[15] O. A. Dellagostin *et al.*, "Reverse vaccinology: An approach for identifying leptospiral vaccine candidates," *Int. J. Mol. Sci.*, vol. 18, no. 1, 2017, doi: 10.3390/ijms18010158.

[16] de la P. M. A. Adler B, "Leptospira and leptospirosis.," *Vet Microbiol.*, 2010.

[17] C. Y. Feng *et al.*, "Immune strategies using single-component LipL32 and multi-component recombinant LipL32-41-OmpL1 vaccines against leptospira Immune strategies using single-component LipL32 and multi-component recombinant LipL32-41-OmpL1 vaccines against leptospira," *Braz J Med Biol Res 42(9)*, vol. 42, no. September, pp. 796–803, 2009.

[18] D. A. Haake, E. M. Walker, D. R. Blanco, C. A. Bolin, J. N. Miller, and M. A. Lovett, "Changes in the surface of Leptospira interrogans serovar grippotyphosa during in vitro cultivation," *Infect. Immun.*, vol. 59, no. 3, pp. 1131–1140, 1991, doi: 10.1128/iai.59.3.1131-1140.1991.

[19] D. A. Haake and J. Matsunaga, "Characterization of the leptospiral outer membrane and description of three novel leptospiral membrane proteins," *Infect. Immun.*, vol. 70, no. 9, pp. 4936–4945, 2002, doi: 10.1128/IAI.70.9.4936-4945.2002.

[20] E. S. Shang and T. A. Summers, "Molecular Cloning and Sequence Analysis of the Gene Encoding LipL41 , a Surface-Exposed Lipoprotein of Pathogenic Leptospira Species," vol. 64, no. 6, pp. 2322–2330, 1996.

 doi:10.20944/preprints202212.0299.v1

[21] B. Cullen, Paul & Xu, Xiaoyi & Matsunaga, James & Sanchez, Yolanda & Ko, Albert & Haake, David & Adler, "Surfaceome of Leptospira spp.," *Infect. Immun.*, 2005.

[22] F. K. Seixas, C. H. Fernandes, D. D. Hartwig, F. R. Conceição, J. A. G. Aleixo, and O. A. Dellagostin, "Evaluation of different ways of presenting LipL32 to the immune system with the aim of developing a recombinant vaccine against leptospirosis," *Can. J. Microbiol.*, vol. 53, no. 4, pp. 472–479, 2007, doi: 10.1139/W06-138.

[23] A. I. Flannery, B., Costa, D., Carvalho, F. P., Guerreiro, H., Matsunaga, J., Da Silva, E. D., Ferreira, A. G., Riley, L. W., Reis, M. G., Haake, D. A., & Ko, "Evaluation of recombinant Leptospira antigen-based enzyme-linked immunosorbent assays for the serodiagnosis of leptospirosis.," *J. Clin. Microbiol.*, 2001, doi: https://doi.org/10.1128/JCM.39.9.3303-3310.2001.

[24] D. A. Guerreiro, H., Croda, J., Flannery, B., Mazel, M., Matsunaga, J., Galvão Reis, M., Levett, P. N., Ko, A. I., & Haake, "Leptospiral proteins recognized during the humoral immune response to leptospirosis in humans.," *Infect. Immun.*, 2001, doi: https://doi.org/10.1128/IAI.69.8.4958-4968.2001.

[25] D. A. Haake *et al.*, "The Leptospiral Major Outer Membrane Protein LipL32 Is a Lipoprotein Expressed during Mammalian Infection The Leptospiral Major Outer Membrane Protein LipL32 Is a Lipoprotein Expressed during Mammalian Infection," 2000, doi: 10.1128/IAI.68.4.2276-2285.2000.Updated.

[26] M. A. Haake, D. A., Champion, C. I., Martinich, C., Shang, E. S., Blanco, D. R., Miller, J. N., & Lovett, "Molecular cloning and sequence analysis of the gene encoding OmpL1, a transmembrane outer membrane protein of pathogenic Leptospira spp.," *J. Bacteriol.*, 1993, doi: https://doi.org/10.1128/jb.175.13.4225-4234.1993.

[27] D. A. Shang, E. S., Exner, M. M., Summers, T. A., Martinich, C., Champion, C. I., Hancock, R. E., & Haake, "The rare outer membrane protein, OmpL1, of pathogenic

Leptospira species is a heat-modifiable porin.," *Infect. Immun.*, 1995, doi: https://doi.org/10.1128/iai.63.8.3174-3181.1995.

[28] D. A. Haake *et al.*, "Leptospiral outer membrane proteins OmpL1 and LipL41 exhibit synergistic immunoprotection," *Infect. Immun.*, vol. 67, no. 12, pp. 6572–6582, 1999, doi: 10.1128/iai.67.12.6572-6582.1999.

[29] R. L. Haake, D. A., Suchard, M. A., Kelley, M. M., Dundoo, M., Alt, D. P., & Zuerner, "Molecular evolution and mosaicism of leptospiral outer membrane proteins involves horizontal DNA transfer.," *J. Bacteriol.*, 2004, doi: https://doi.org/10.1128/JB.186.9.2818-2828.2004.

[30] H. Dong *et al.*, "Characterization of the ompL1 gene of pathogenic Leptospira species in China and cross-immunogenicity of the OmpL1 protein," *BMC Microbiol.*, vol. 8, pp. 1–12, 2008, doi: 10.1186/1471-2180-8-223.

[31] Y. Vakili, B., Eslami, M., Hatam, G. R., Zare, B., Erfani, N., Nezafat, N., & Ghasemi, "Immunoinformatics-aided design of a potential multi-epitope peptide vaccine against Leishmania infantum.," *Int. J. Biol. Macromol.*, 2018, doi: https://doi.org/10.1016/j.ijbiomac.2018.08.125.

[32] L. V. Fomsgaard, A., Karlsson, I., Gram, G., Schou, C., Tang, S., Bang, P., Kromann, I., Andersen, P., & Andreasen, "Development and preclinical safety evaluation of a new therapeutic HIV-1 vaccine based on 18 T-cell minimal epitope peptides applying a novel cationic adjuvant CAF01.," *Vaccine*, 2011, doi: https://doi.org/10.1016/j.vaccine.2011.07.025.

[33] R. A. Shey *et al.*, "In-silico design of a multi-epitope vaccine candidate against onchocerciasis and related filarial diseases," *Sci. Rep.*, vol. 9, no. 1, pp. 1–18, 2019, doi: 10.1038/s41598-019-40833-x.

[34] M. Rabienia *et al.*, "Exploring membrane proteins of Leishmania major to design a new

multi-epitope vaccine using immunoinformatics approach," *Eur. J. Pharm. Sci.*, vol. 152, p. 105423, 2020, doi: 10.1016/j.ejps.2020.105423.

[35] A. S. Kolaskar and P. C. Tongaonkar, "A semi-empirical method for prediction of antigenic determinants on protein antigens," *FEBS Lett.*, vol. 276, no. 1–2, pp. 172–174, 1990, doi: 10.1016/0014-5793(90)80535-Q.

[36] M. V. Larsen, C. Lundegaard, K. Lamberth, S. Buus, O. Lund, and M. Nielsen, "Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction," *BMC Bioinformatics*, vol. 8, pp. 1–12, 2007, doi: 10.1186/1471-2105-8-424.

[37] K. S. Lata, S. Kumar, V. Va, and P. Sharma, "Exploring Leptospiral proteomes to identify potential candidates for vaccine design against Leptospirosis using an immunoinformatics approach," no. October 2017, pp. 1–15, 2018, doi: 10.1038/s41598-018-25281-3.

[38] H. H. Bui, J. Sidney, K. Dinh, S. Southwood, M. J. Newman, and A. Sette, "Predicting population coverage of T-cell epitope-based diagnostics and vaccines," *BMC Bioinformatics*, vol. 7, pp. 1–5, 2006, doi: 10.1186/1471-2105-7-153.

[39] I. Dimitrov, I. Bangov, D. R. Flower, and I. Doytchinova, "AllerTOP v.2 - A server for in silico prediction of allergens," *J. Mol. Model.*, vol. 20, no. 6, 2014, doi: 10.1007/s00894-014-2278-5.

[40] I. A. Doytchinova and D. R. Flower, "VaxiJen: A server for prediction of protective antigens, tumour antigens and subunit vaccines," *BMC Bioinformatics*, vol. 8, pp. 1–7, 2007, doi: 10.1186/1471-2105-8-4.

[41] L. W. Lesong Wei, Xiucai Ye, Tetsuya Sakurai, Zengchao Mu, "ToxIBTL: prediction of peptide toxicity based on information bottleneck and transfer learning.," *Bioinformatics*, vol. 38, no. 6, 2022, [Online]. Available: https://doi.org/10.1093/bioinformatics/btac006

 doi:10.20944/preprints202212.0299.v1

[42] P. Vijayachari *et al.*, "Immunogenicity of a novel enhanced consensus DNA vaccine encoding the leptospiral protein LipL45," *Hum. Vaccines Immunother.*, vol. 11, no. 8, pp. 1945–1953, 2015, doi: 10.1080/21645515.2015.1047117.

[43] G. D. C. Geourjon, "SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments," *Bioinformatics*, vol. 11, 1995, doi: https://doi.org/10.1093/bioinformatics/11.6.681.

[44] T. I. Huszar, M. A. Jobling, and J. H. Wetton, "A phylogenetic framework facilitates Y-STR variant discovery and classification via massively parallel sequencing," *Forensic Sci. Int. Genet.*, vol. 35, no. April, pp. 97–106, 2018, doi: 10.1016/j.fsigen.2018.03.012.

[45] J. Ko, H. Park, L. Heo, and C. Seok, "GalaxyWEB server for protein structure prediction and refinement," *Nucleic Acids Res.*, vol. 40, no. W1, pp. 294–297, 2012, doi: 10.1093/nar/gks493.

[46] M. Wiederstein and M. J. Sippl, "ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins," *Nucleic Acids Res.*, vol. 35, no. SUPPL.2, pp. 407–410, 2007, doi: 10.1093/nar/gkm290.

[47] R. A. Laskowski, M. W. MacArthur, D. S. Moss, and J. M. Thornton, "PROCHECK: a program to check the stereochemical quality of protein structures," *J. Appl. Crystallogr.*, vol. 26, no. 2, pp. 283–291, 1993, doi: 10.1107/s0021889892009944.

[48] C. Colovos and T. Yeates, "Verification of protein structures: Patterns of nonbonded atomic interactions," pp. 1511–1519, 1993.

[49] L. C. Xue, J. P. Rodrigues, P. L. Kastritis, A. M. Bonvin, and A. Vangone, "PRODIGY: A web server for predicting the binding affinity of protein-protein complexes," *Bioinformatics*, vol. 32, no. 23, pp. 3676–3678, 2016, doi: 10.1093/bioinformatics/btw514.

[50] A. A. Omoniyi *et al.*, "In silico design and analyses of a multi-epitope vaccine against

Crimean-Congo hemorrhagic fever virus through reverse vaccinology and immunoinformatics approaches," *Sci. Rep.*, vol. 12, no. 1, pp. 1–17, 2022, doi: 10.1038/s41598-022-12651-1.

[51] K. Jalal *et al.*, "Reverse vaccinology approach for multi-epitope centered vaccine design against delta variant of the SARS-CoV-2," *Environ. Sci. Pollut. Res.*, no. 0123456789, 2022, doi: 10.1007/s11356-022-19979-1.

[52] K. M. Adam, "Immunoinformatics approach for multi-epitope vaccine design against structural proteins and ORF1a polyprotein of severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2)," *Trop. Dis. Travel Med. Vaccines*, vol. 7, no. 1, pp. 1–13, 2021, doi: 10.1186/s40794-021-00147-1.

[53] T. E. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, "UCSF Chimera--a visualization system for exploratory research and analysis.," *J. Comput. Chem.*, 2004, doi: https://doi.org/10.1002/jcc.20084.

[54] R. R. Donati C, "Reverse vaccinology in the 21st century: improvements over the original design.," *Ann N Y Acad Sci*, 2013, doi: doi: 10.1111/nyas.12046.

[55] A. Aslam, M., Shehroz, M., Hizbullah, Shah, M., Khan, M. A., Afridi, S. G., & Khan, "Potential druggable proteins and chimeric vaccine construct prioritization against Brucella melitensis from species core genome data.," *Genomics, Proteomics Bioinforma.*, 2020, [Online]. Available: https://doi.org/10.1016/j.ygeno.2019.10.009

[56] A. D. Odir, A. A. Grassmann, D. D. Hartwig, S. R. Félix, É. F. Da Silva, and A. J. A. McBride, "Recombinant vaccines against leptospirosis," *Hum. Vaccin.*, vol. 7, no. 11, pp. 1215–1224, 2011, doi: 10.4161/hv.7.11.17944.

[57] N. Ahmed *et al.*, "Multilocus sequence typing method for identification and genotypic classification of pathogenic Leptospira species," vol. 10, pp. 1–10, doi: 10.1186/1476-0711-5-Received.

 doi:10.20944/preprints202212.0299.v1

[58] K. J. Allan, H. M. Biggs, J. E. B. Halliday, and R. R. Kazwala, "Epidemiology of Leptospirosis in Africa : A Systematic Review of a Neglected Zoonosis and a Paradigm for ' One Health ' in Africa," pp. 1–25, 2015, doi: 10.1371/journal.pntd.0003899.

[59] A. F. B. Victoriano *et al.*, "Leptospirosis in the Asia Pacific region," vol. 9, pp. 1–9, doi: 10.1186/1471-2334-9-147.

CAPTIONS

Fig. 1- Overview of the workflow followed for the designing of multi-epitope-based vaccine against Leptospirosis.

Fig.2 - Population coverage for the MHC-I and MHC-II alleles for all selected epitopes.

Fig.3 A) – sequence of the vaccine construct; B) schematic presentation of multiple-epitope vaccine construct.

Fig. 4 - Secondary structure prediction using SOPMA server for 4th construct, stating the percentage of alpha helix, extended strand, beta turn and random coils. (Blue - alpha helix; red – extended strand; green – beta turn; yellow- random coil)

Fig. 5 – Tertiary structure modeling performed using Robetta server (ab-initio method) for 4th vaccine construct.

Fig. 6.A.) - Ramachandran plot for 4th vaccine construct B) ERRAT graph showing the error value for the 4th vaccine construct C) Prosa-web (z-score) showing the overall quality of the 4th construct.

Fig. 7 - Docked complex (vaccine construct 4 with receptor TLR 4) visualized in chimera.

Fig. 8.A) RMSD graph obtained at 5 ns for construct 4 with receptor TLR 4; 8.B) | RMSD graph obtained at 50 ns for construct 4 with receptor TLR 4.