Hyrcanian Journal of Clinical Medicine
eISSN: 3150-4213

Original Article

OPEN ACCESS

Engineering a multi-epitope vaccine for *Leishmania infantum*: a combination of traditional immuno-informatic tools and python programming approaches

Mohammad Rahmati[1 2], Fatemeh Fathali[3], Ali Kia-Karimi[4], Abolfazl Nikkhoy[5], Amirhossein Larijani[5 6*]

[1] Department of Medical Biotechnology, School of Paramedicine, Guilan University of Medical Sciences, Rasht, Iran
[2] Medical Biotechnology Research Center, School of Paramedicine, Guilan University of Medical Sciences, Rasht, Iran
[3] Department of Cellular and Molecular Biology, Faculty of Basic Science, University of Mazandaran, Babolsar, Mazandaran, Iran
[4] Student Research Committee, School of Medicine, Mazandaran University of Medical Sciences, Sari, Iran
[5] Student Research Committee, School of Medicine, Guilan University of Medical Sciences, Rasht, Iran
[6] Regenerative Medicine Research Center, Guilan University of Medical Sciences, Rasht, Iran

Corresponding author:
Amirhossein Larijani
Student Research Committee, School of Medicine, Guilan University of Medical Sciences, Rasht, Iran
Regenerative Medicine Research Center, Guilan University of Medical Sciences, Rasht, Iran
E-mail: Amir1313801360@gmail.com

Abstract

Leishmaniasis, a significant health challenge in subtropical regions, requires effective and safe therapeutic options. Our primary objective was to design and evaluate a multi-epitope peptide-based vaccine, LIDVC, targeting Leishmania infantum. LIDVC was developed by integrating key epitopes from GP63, KMP11, and LACK antigens with the adjuvant Ov-ASP-1. Advanced bioinformatics tools assessed the vaccine's antigenicity, allergenicity, toxicity, and population coverage. Structural integrity was confirmed through AlphaFold and DeepCDpred analyses, and molecular docking studies with TLR4 were performed. Our comprehensive computational evaluation demonstrated robust interactions between LIDVC and TLR4, favorable dynamic properties, and strong protective responses akin to those from natural infections. Codon optimization and in silico cloning showed potential for efficient expression in Escherichia coli. The LIDVC vaccine candidate shows promise against L. infantum, warranting experimental validation and preclinical studies. Our computational evaluation highlights LIDVC's potential as a safe and effective vaccine against leishmaniasis, addressing the need for improved therapeutic options in subtropical regions. Further studies will focus on experimental validation and developing a safe and effective vaccine against leishmaniasis.

Article Info:
Received: 01 February 2027
Accepted: 10 March 2027

Copyright:
© 2027 by the authors

This work is licensed under a Creative Commons Attribution- NonCommercial 4.0 International (CC BY) license.
Noncommercial uses of the work are permitted, provided the original work is properly cited.

Keywords: Leishmaniasis, Multi-epitope vaccine, Python, In silico design, Immunoinformatics, Vaccine development, Computational biology

1. Introduction

Leishmaniasis, caused by the protozoan parasite Leishmania, represents a significant global health challenge, particularly in subtropical regions where the disease imposes a profound burden on affected populations (1). Transmitted through the bites of infected sandflies, leishmaniasis manifests in various forms, ranging from cutaneous to visceral diseases (see **Fig. 1**). The limitations and adverse effects associated with existing therapeutic options underscore the urgent need for effective and safe vaccines to prevent and manage this disease (2).

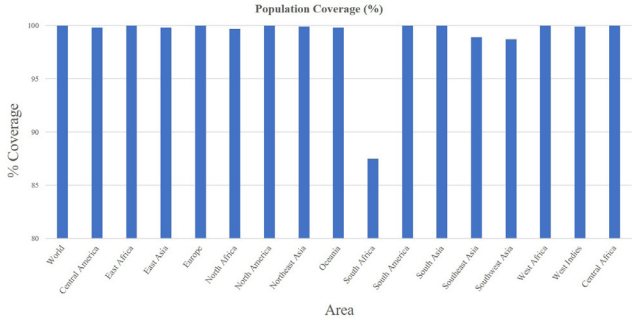


Figure 1. *Population coverage analysis of the predicted epitopes demonstrating broad global representation and variable regional coverage.*

Significant research efforts have identified various antigenic proteins expressed throughout the Leishmania life cycle using high-throughput methodologies and bioinformatics (3-5). Among these, glycoprotein 46 (gp46), cathepsin L-like and B-like proteases, histone H2A, glucose-regulated protein 78 (grp78), and stress-inducible protein 1 (STI-1) have shown promise as vaccine candidates (6). Preclinical studies have demonstrated the ability of these antigens to elicit protective immune responses against Leishmania infections.

The immune response to Leishmania infection comprises a complex interplay between pro-inflammatory (Th1) and anti-inflammatory (Th2) cytokines (7, 8). Protective immunity predominantly relies on Th1-mediated responses characterized by the release of interferon-gamma (IFN-γ), which activates macrophages to eliminate intracellular pathogens (5, 8). Conversely, Th2 responses may facilitate parasite survival and contribute to disease progression. Understanding this intricate balance is vital for the design of effective vaccines against leishmaniasis.

Recent advancements in bioinformatics and computational biology have revolutionized vaccine development by enabling the identification of immunodominant epitopes and antigenic proteins through extensive computational analyses (4). In this context, the programming language Python has emerged as a powerful tool, utilizing its extensive libraries for various tasks, including antigenicity, allergenicity, and immunogenicity assessments. By leveraging Python programming scripts and advanced tools, our research focused on the computational evaluation of a multi-epitope vaccine candidate, referred to as LIDVC, incorporating prominent epitopes from KMP11, GP63, and LACK alongside the adjuvant Ov-ASP-1.

Our in silico methodology facilitated rapid predictions of the vaccine candidate's potential, supporting the hypothesis that LIDVC could effectively elicit robust immune responses similar to those induced by natural infections. The present work demonstrates the efficacy of using in silico approaches, particularly Python, for accelerating vaccine development against leishmaniasis, laying the groundwork for subsequent experimental validation and preclinical studies aimed at establishing a safe and effective vaccine against L. infantum.

2. Methods

2.1 Data collection and epitopes prediction

We have used the validated epitopes predicted in the previous articles to design a vaccine construct, the FASTA-formatted amino acid sequences for L. infantum GP63 (Accession no. QJF54184), KMP11 (Accession no. AGV77135), and LACK (Accession no. UQI50440) were collected from the National Center for Biotechnology Information (NCBI) website and servers like iedb and CTLPRED or VaxiJen were used for epitope prediction (19).

2.2 Population coverage analysis

Population coverage analysis was performed using the IEDB Population Coverage tool (http://tools.iedb.org/population/) to estimate the global and continental representation of the selected epitopes. This analysis represents the proportion of individuals potentially capable of recognizing the selected epitopes based on the distribution of their corresponding HLA class I and class II alleles across different geographical populations. The analysis was conducted for the World population as well as 16 regional populations, including Central Africa, Central America, East Africa, East Asia, Europe, North Africa, North America, Northeast Asia, Oceania, South Africa, South America, South Asia, Southeast Asia, Southwest Asia, West Africa, and West Indies. The population coverage and PC90 were calculated for

the combined MHC class I and class II epitope set.

2.3 Toxicity Assessment of Epitopes

The toxicity evaluation and measurement of the predicted epitopes, as well as the exclusion of known toxic ones, were carried out utilizing Server ToxIBTL, which uses transfer learning to refine a model trained on a protein dataset. This server's algorithm, located at https://server.wei-group.net/ToxIBTL/Server.html, is based on three fundamental methods: sequence encoding, optimization, and classification. Evolutionary profiles are created from raw sequences and processed using the CNN_BiGRU network to collect local and global information. The FEGS model gathers physicochemical information and graphical and statistical properties. The information bottleneck approach is used in the second phase to concatenate and optimize the evolutionary and physicochemical features. The optimized characteristics are then used to assess the toxicity status of each sequence (20).

2.4 Using Python programming for organizing ToxiBTL server output

After checking the toxicity level of epitopes using the ToxiBTL server, the received output needed to be checked, and non-toxic epitopes must be separated for further analysis in the next step. However, one of the challenges in the data received from this server was the existence of a large amount of information that confused us. We used Python programming to overcome this challenge and shorten the vaccine design time. As mentioned above, the server developed a Python application to optimize and enrich the output file. The Python script first reads the input file to reach the target word "non-toxic." It then lists all the identified records in another file with the title "non_toxic.txt." then the number of rows that contain the non-toxic phrases was selected, and these numbers matched to the duplicate rows in another input file named (epitope.txt). Utilizing this methodology has supported processes of data extraction, formatting, and augmentation to assure the quality of its output and speed up our operations. The code for this application is located in a repository on GitHub with the following access link: https://github.com/Amirlrj128/bioinf-applications.

2.5 Assessment of immunogenicity level

The VaxiJen v2.0 server was utilized to estimate the protein's antigenicity and remove epitopes that could not stimulate the immune system based on the 0.7 threshold. This server, available at http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html, employs an ACC transformation method to convert the provided protein sequence into standardized vectors. The predictions heavily rely on the chemical properties of proteins. Notably, VaxiJen v2.0 has reported accuracy rates ranging from 70% to 89% (19).

2.6 Utilizing Python Programming to classify the output of the VaxiJen server

The second software, designed in Python, is coded to classify the output data in summary mode on the VaxiJen server. This software was developed for more straightforward classification and access to select immunogen epitopes for further evaluation. The approach of the commands written in Python for this software is that it first receives the output file of the VaxiJen server in summary mode as input. Then, it divides them into two categories among rows with the phrase's "antigen" and "non-antigen" and finally provides us with two files with the mentioned separate names. Following the link in GitHub, refer to the repository of this software code: https://github.com/Amirlrj128/bioinf-applications.

2.7 Forecasting the allergenic properties of epitopes

The AllerTOP v2.0 server assessed the allergenicity of the predicted epitopes. Accessible at https://www.ddg-pharmfac.net/AllerTOP/, the AllerTOP server uses an auto cross-covariance (19) method to convert protein sequences into uniform vectors of equal length. It is noteworthy that AllerTOP v2.0 has reported accuracy rates of 65% (21).

2.8 Finalize the vaccine candidate sequence

In this step, the selected epitope sequences are combined with the adjuvant sequence downloaded from the Uniport website, available at https://www.uniprot.org/, and various linkers to finalize the vaccine candidate sequence. The sequence of epitopes will be inserted as B cells first, then MHCs, and finally CTLs and connected with linker sequences. Then, the secondary structure of the vaccine candidate protein was obtained using the probe server https://npsa-prabi.ibcp.fr/cgi-bin/secpred_sopma.pl; the stability of the vaccine protein was evaluated and observed with the program server to finally get a stable structure (22, 23).

2.9 Docking Analysis

The Swarm Dock Server (SDS), https://bio.tools/swarmdock, is used to design protein-protein interactions. It requires only PDB format files as input to create molecular docking with flexible algorithms that

can predict molecular binding sites. In fact, SDS can be used to guide drug and vaccine design studios for evaluation (24).

2.10 Flexibility of the Docked TLR4-Designed Vaccine Complex

CABS-flexhttps://biocomp.chem.uw.edu.pl/CABSflex2 is a Python package for fast simulations of protein structure flexibility. This server simulates protein dynamics using the CABS model and makes selected modeling results using all-atom representation models and analysis of different models (11).

2.11 Codon Optimization and in Silico Cloning

For this step, the server Java Codon Adaptation Tool (JCat) is used; JCat is available at the address https://www.jcat.de/. This server applies the sequenced target gene codon for most prokaryotes and eukaryotes that host the gene expression process to develop heterologous protein production. Unlike most existing tools, it doesn't need to define genes with high expression manually, and this feature makes it simple and convenient (25).

2.12 Prediction and validation of the Secondary and 3D Structures PSIPRED software

To predict the second structure of the vaccine candidate protein, there is an excellent tool available at https://bio.tools/psipred. This tool detects the folding of the protein and predicts the topology of the MEMSAT-2 transmembrane by accessing GenTHREADER. This server provides all the previous analysis methods on one website and has a very efficient user interface that significantly helps scientists discover better results (26).

2.13 ProSA-web service

Protein structure analysis (ProSA) is available at https://prosa.services.came.sbg.ac.at/prosa.php. This tool is used to find errors in the predicted models of the third structure of the vaccine candidate protein, which analyzes the protein structure and is very simple and practical. In particular, it displays the quality scores of a protein in the context of all known protein structures. Is given, and problematic structure parts are shown and highlighted in a 3D molecule viewer (27).

2.14 Galaxy Refine

To understand the biological function and regulation of the protein structure, a tool is needed to predict and improve the protein structure with pattern-based modeling, and for this purpose, the GalaxyRefine server was selected for structure refinement, one of the best bioinformatics tools available in this field. Refine server uses a CAPS9 (9th Critical Assessment of Techniques for Protein Structure Prediction)-based method (28). The Galaxy Refine server is available at https://galaxy.seoklab.org/cgi-bin/submit.cgi?type=REFINE.

3. Results

3.1 Candidate vaccine's construct design

As reported previously by our team (9), in the present in silico study, several prediction and screening protocols were employed to design the final construction, which was named LIDVC (Leishmania's in silico-designed vaccine candidate). The best-scored epitopes from GP63, KMP11, and LACK antigens were evaluated, refined, and arranged in the construct (**Fig. 4C**). Additionally, an adjuvant, Ov-ASP-1 (Onchocerca volvulus Activation-associated secreted protein-1; UniProt accession number O16854) amino acid sequences were inserted. Linear and conformational B-cell epitopes within the entire length of the designed vaccine structure were predicted, and high-ranked CTL epitopes were selected. MHC class I and II binding epitopes were also predicted, and a thorough antigenicity evaluation of the target proteins was conducted.

3.2 Population coverage analysis

Population coverage analysis indicated that the selected epitopes were predicted to cover 56.25% of the global population, suggesting a moderate level of overall population representation (**Fig 2**.). However, substantial differences were observed among the examined geographic regions. The highest predicted coverage was recorded for Europe (62.60%), followed by North America (60.41%) and North Africa (50.57%). Moderate coverage levels were observed in the West Indies (46.58%), Southeast Asia (44.87%), South Asia (43.66%), Southwest Asia (42.42%), East Asia (41.67%), East Africa (40.85%), Northeast Asia (40.54%), and West Africa (38.35%). In contrast, relatively lower coverage was predicted for South Africa (34.58%), South America (30.76%), and Central Africa (30.67%). The lowest values were observed in Oceania (19.74%) and Central America (4.99%). Across all evaluated regions, the mean predicted population coverage was 39.58%, with an average of 0.58 epitope hits per HLA combination and a PC90 value of 0.17. These findings indicate that the selected epitopes have the po-

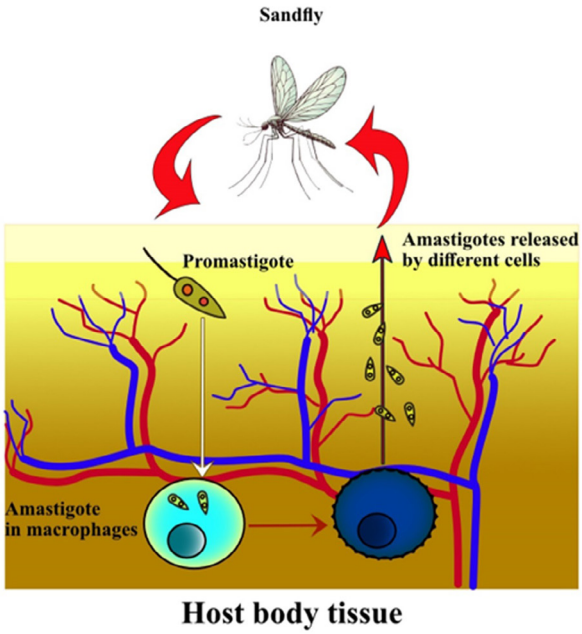


Figure 2. *Life cycle and transmission of Leishmania infantum in the sandfly and host body.*

tential to provide immune recognition across a broad range of geographically and genetically diverse populations, although predicted coverage varies considerably in association with regional differences in HLA allele frequencies. Overall, the population coverage analysis supports the potential broad applicability of the designed multi-epitope vaccine, while also highlighting differences in its predicted effectiveness among distinct populations.

3.3 Toxicity assessment of epitopes

To mitigate the risk of incorporating toxic epitopes into the final vaccine construct, the ToxIBTL server was utilized for a comprehensive toxicity assessment. After optimizing and refining the results using a Python program, the analysis confirmed that all selected epitopes exhibited no toxicity, as detailed in **Table 1**. This step is crucial for ensuring the safety and efficacy of the vac-

Table 1. L. infantum KMP11, LACK, and GP63 selected epitopes toxicity evaluation by ToxIBTL server.

Antigens		Epitopes	Score	Toxicity
KMP11	Linear B cell	DEEFNRKMQALNAKFFADKPDESTLSPEMKE HYEKFERMIKEHTEKFNKKMHEHSEHFKQK	0.00018532519	No
		AELLEQQKAAQY	4.6070228e-27	No
	CTL	FFADKPDES	6.855844e-24	No
		KFFADKPDE	6.6241714e-09	No
		RLDEEFNRK	7.5820573e-28	No
	MHC	FNRKMQALN	0.39676902	No
		FSAPFKRLD	0.26262674	No
		HTEKFNKKM	0.027506351	No
LACK	Linear B cell	NPDRHSVDSDYGLPSHRLEG	3.9384704e-05	No
		CMHEFLRDGHE	0.0011695899	No
		GGKCERTLKGH	4.9916435e-15	No
	CTL	SHRLEGHTG	1.1731253e-31	No
		KFLKHTKDV	3.48751e-14	No
		FVSCVSLAH	1.7928064e-15	No
	MHC	SLACPQQAG	0.18910587	No
		LACPQQAGS	0.16934268	No
		ACPQQAGSY	3.309424e-05	No
GP63	Linear B cell	KVRQVQDKWNATGMVDEICGDFKVPPAHITEGFS	3.584472e-07	No
		IEDQGGAGSAGSHIKMRNAQ	9.6080445e-21	No
		FFEGARILESISNVRHKDF	3.7275776e-21	No
	CTL	VRQVQDKWN	3.7414862e-26	No
		PQALQLHTE	2.722246e-30	No
		RILESISNV	6.7163546e-36	No
	MHC	PSEEGVLAW	0.18029849	No
		CGDFKVPPA	0.15298171	No
		CDTLEYLEI	0.02768834	No

cine candidate.

3.4 Assessment of epitopes immunogenicity and allergenicity

The immunogenicity of the selected antigens was systematically evaluated using the VaxiJen v2.0 server, which analyzed their amino acid sequences (see **Table 2**). To enhance the classification process, a dedicated Python program was developed to categorize the output into "antigen" and "non-antigen" groups. This method not only simplifies the identification of immunogenic epitopes but also facilitates their subsequent analysis, aligning with best practices observed in similar studies. Furthermore, the allergenicity of the top-ranked epitopes was rigorously assessed using the AllerTOP v2.0 server, with findings confirming that none of the evaluated epitopes demonstrated allergenic potential (refer to **Table 2**). This comprehensive evaluation is crucial for ensuring the safety and efficacy of the vaccine candidates in development.

3.5 Physico-chemical parameters in silico analysis

The LIDVC construct comprises 1028 amino acids, with 247 positively charged (Arg, Lys) residues and 116 negatively charged (Asp, Glu) residues. It has a molecular weight (10) of approximately 118 kDa and an isoelectric point (pI) value of 9.98. The extinction coefficient of the chimeric protein was determined to be 124565 M−1cm−1 at 280 nm, and its estimated half-life was approximately 30 hours in mammalian reticulocytes in vitro, over 20 hours in yeast in vivo, and more than 10 hours in Escherichia coli in vivo. The instability index was calculated to be 29.05. Additionally, the aliphatic index and GRAVY (Grand average of hydropathicity) of the final construct were 61.96 and -1.018, respectively, as detailed in **Table 3**. These findings are consistent with existing literature and provide valuable insights into the biochemical characteristics of the LIDVC construct.

3.6 Prediction and validation of the Secondary and 3D Structures

The PRABI server available at Self-Optimized Prediction Method with Alignment (SOPMA) (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html) web server tool was used to assess and forecast the secondary structure of the LIDVC, predicted result illustrated in the **Fig. 3A**. The resultant structure predominantly consists of 45.43% alpha-helix, accompanied by 17.32% beta strands and 29.86% coils. Subsequently, the most accurate 3D model predicted by Phyre2 was chosen for validation analysis, depicted in **Fig. 3**. Additionally, the 3D model

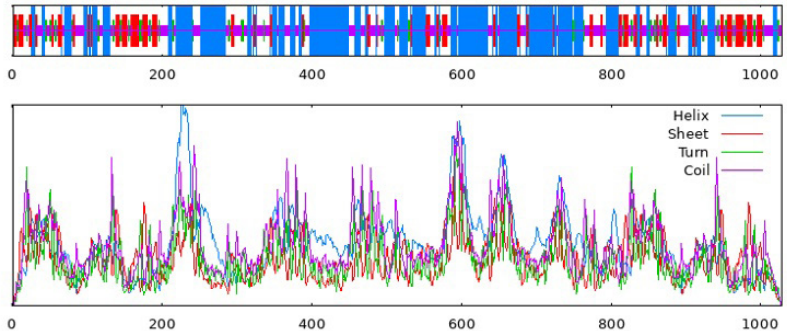
Table 2. L. infantum KMP11, LACK, and GP63 selected epitopes' immunogenicity and allergenicity were evaluated by the ToxIBTL and AllerTOP servers, respectively.

Antigens		Epitopes	Score	Immunogenicity	Allergenicity
KMP11	Linear B cell	DEEFNRKMQALNAKFFADKPDESTLSPEMKE HYEKFERMIKEHTEKFNKKMHEHSEHFKQK	0.8626	Yes	No
	CTL	FFADKPDES	1.4285	Yes	No
		KFFADKPDE	1.5829	Yes	No
	MHC	QALNAKFFA	1.6665	Yes	No
		FFADKPDES	1.4285	Yes	No
		QALNAKFFA	1.6665	Yes	No
LACK	Linear B cell	GGKCERTLKGH	2.0231	Yes	No
		CMHEFLRDGHE	1.0095	Yes	No
		NPDRHSVDSDYGLPSHRLEG	0.8747	Yes	No
	CTL	KFLKHTKDV	1.4556	Yes	No
		SHRLEGHTG	1.3804	Yes	No
		FVSCVSLAH	0.7503	Yes	No
	MHC	SLACPQQAG	0.7061	Yes	No
GP63	Linear B cell	IEDQGGAGSAGSHIKMRNAQ	1.3162	Yes	No
		IASRYDQ	0.7758	Yes	No
	CTL	PQALQLHTE	0.9175	Yes	No
	MHC	YVASVPSEE	1.2281	Yes	No
		YSALTMAIF	1.1573	Yes	No
		LHTERLKVR	1.1653	Yes	No

Table 3. Physico-chemical parameters analysis of the LIDVC. *Grand average of hydropathy.

	No. of residues	Molecular weight	pI	Positive residues (Arg+Lys)	Negative residues (Asp+Glu)	Extinction coefficient (at 280 nm in H_2O)	Half-life (h)	Instability index	Aliphatic index	GRAVY*
LIDVC	1028	118 KDa	9.98	247	116	124565 $M^{-1}cm^{-1}$	~30 >20 >10	29.05	91.96	-1.018

A



Sequence length: 1028

SOPMA:

Alpha helix (Hh): 467 is 45.43%

3_{10} helix (Gg): 0 is 0.00%

Pi helix (Ii): 0 is 0.00%

Beta bridge (Bb): 0 is 0.00%

Extended strand (Ee): 178 is 17.32%

Beta turn (Tt): 76 is 7.39%

Bend region (Ss): 0 is 0.00%

Random coil (Cc): 307 is 29.86%

Ambiguous states (?): 0 is 0.00%

Other states: 0 is 0.00%

B

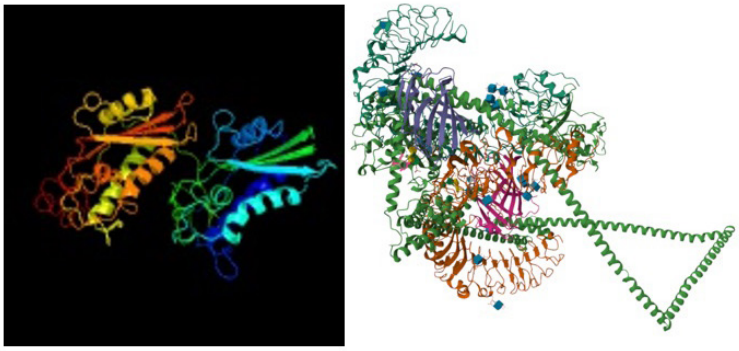


Figure 3. *The secondary and 3D structures of LIDVC prediction and validation. (A) Composition and position of α-helices, β-sheets, turns, and coils of 2D structure of final construct illustrated graphically. (B) The final 3D model of LIDVC was predicted by Phyre2/Robetta (Left) and further evaluated by the I-TASSER server (Right).*

DOI: 10.66224/hjcm.1.6.4

Table 4. Structure Information (Galaxy Refine result)

Model	GDT-HA	RMSD	MolProbity	Clash score	Poor rotamers	Rama favored
Initial	1.0000	0.000	1.304	2.3	0.0	95.7
MODEL 1	0.9881	0.286	1.796	11.0	0.8	96.4
MODEL 2	0.9903	0.274	1.791	11.1	0.6	96.5
MODEL 3	0.9878	0.296	1.808	10.8	0.8	96.2
MODEL 4	0.9827	0.311	1.849	11.5	0.6	96.0
MODEL 5	0.9837	0.308	1.796	10.7	0.8	96.3

of the engineered LIDVC was subjected to further evaluation using the I-TASSER server, as shown in **Fig. 3B**.

To refine and validate the tertiary modeling of LIDVC, we utilized the ProSA-web service and GalaxyRefine server to analyze the quality and potential errors in the 3D structure of the final selected model (**Table 4**). The final model's Z-score was -8.91 (**Fig. 3B**). The ProSA-web results indicated that the selected model required no further refinement, as the Z-score fell within the acceptable range.

Furthermore, the GalaxyRefine server output revealed that five models met the acceptable parameters. As detailed in **Table 4**, model 1 exhibited the most favorable characteristics, with a GDT-HA of 0.9881, RMSD of 0.286, MolProbity of 1.796, Clash score of 11.0, Poor rotamers of 0.8, and Rama favored of 96.4.

The Ramachandran plot analysis (RAMPAGE) demonstrated that 92.8% of the residues were located in the favored region, with 5.7% in the allowed region and only 1.5% in the outlier region (**Fig. 4A**), indicating the high quality of the predicted structure. Additionally, the overall quality factor of the selected model, as determined by the ERRAT server, was 93.4% (**Fig. 4B**). This comprehensive validation process confirmed the robustness and accuracy of the selected 3D model of LIDVC.

3.7 Candidate vaccine validating

The candidate vaccine, which was designed carefully and named LIDVC, as shown in **Fig. 4C**, totally contains 220 amino acids of Ov-ASP-1 as an adjuvant at both the N- and C-terminal, and composition of selected and refined epitopes from KMP (15 epitopes), LACK (5 epitopes), and GP63 (26 epitopes) and two types of linkers (KK & EAAAK) between them which contains 808 amino acids (some of epitopes were listed in **Table 5**).

Table 5. LIDVC protein epitopes antigenicity. Some high-ranked epitopes are listed below.

Antigens	Epitopes		Score	Antigenicity
KMP11	Linear B cell	DEEFNRKMQALNAKFFADKPDESTLSPEMKEH YEKFERMIKEHTEKFNKKMHEHSEHFKQK	0.8626	Yes
	MHC	FFADKPDES	1.4285	Yes
		FADKPDEST	1.0419	Yes
		ADKPDESTL	0.9647	Yes
	CTL	FFADKPDES	-	Yes
		KFFADKPDE	-	Yes
LACK	Linear B cell	NPDRHSVDSDYGLPSHRLEG	-	Yes
		GGKCERTLKGH	-	Yes
	MHC	SLACPQQAG	-	Yes
	CTL	SHRLEGHTG	-	Yes
		KFLKHTKDV	-	Yes
GP63	Linear B cell	IASRYDQ	-	Yes
		IEDQGGAGSAGSHIKMRNAQ	-	Yes
	MHC	YSALTMAIF	1.1573	Yes
		LHTERLKVR	1.1653	Yes
		HTERLKVRQ	1.1211	Yes

DOI: 10.66224/hjcm.1.6.4

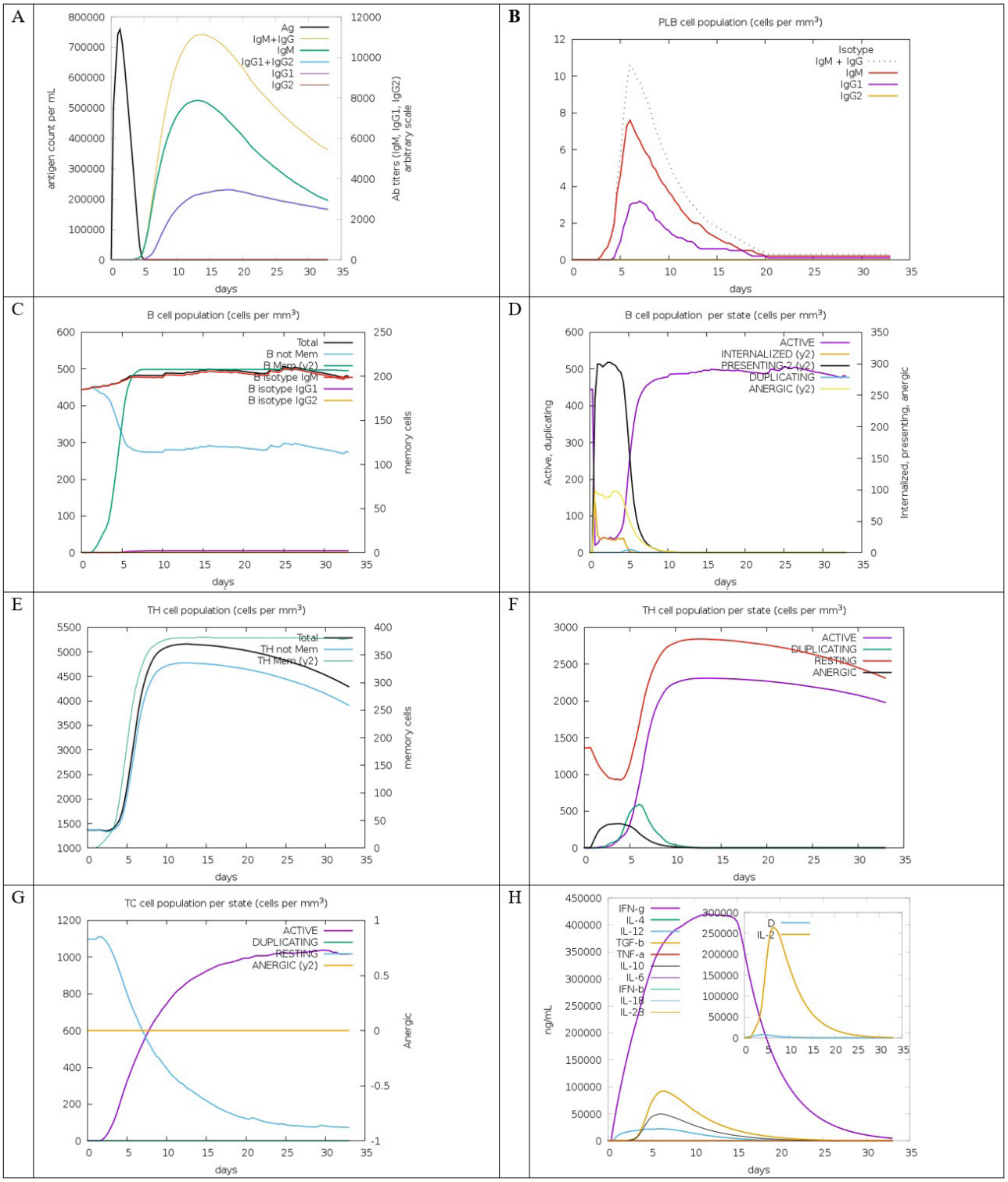


Figure 4. *In silico immune simulation with the subunit vaccine provided by the C-IMMSIM simulation server (available at http://kraken.iac.rm.cnr.it/C-IMMSIM). (29) Different isotypes and combinations of antibody production (colored peaks) by PLB cells after one dose of antigen (Black peak) administration.* ***(C, D)*** *Populations of various subsets of B cells (cells per mm3) induced by the vaccine injection.* ***(E)*** *Memory T-Helper cells (not Mem and y2) and Total T-Helper cell proliferation curve.* ***(F)*** *T-Helper cell population contains active, resting (not encountered the antigen), and anergic (T-cells tolerant to antigens) cells after LIDVC injection.* ***(G)*** *Proliferation of active, resting, and anergic cytotoxic T-cells induced by LIDVC.* ***(H)*** *The various cytokines and interleukin concentrations are shown in colored peaks with prominent up-regulation of pro-inflammatory IFN-γ (shown in violet). The inside graph shows a measure of diversity* ***(D)*** *of IL-2.*

DOI: 10.66224/hjcm.1.6.4

3.8 In silico immune simulation of LIDVC

The immune simulation was accomplished by utilizing the C-IMMSIM Online server available at http://kraken.iac.rm.cnr.it/C-IMMSIM, which can be in the silico-based simulation of both humoral and cell-mediated immune response. After the application of a single dose, the antigen count reached around 7.6×105/mL and dipped to count 0 on the 5th day. Significant high amount of IgG+IgM and slightly lower IgG1+IgG2 population was indicated after 12 days of the LIDVC single dose as part of the primary immune response (**Fig. 5A**). A high amount of IgM and IgG1 isotypes produced by PLB cells population after six days (**Fig. 5B**). Amounts of activated memory B-cells persisted till 33 days as the entire simulation period. Furthermore, activation of cell-mediated immunity was predicted from the increased population of active TH, CTL, and memory T-Helper cells. Significant proliferation of memory T-Helper cell from day ten till day 35 has

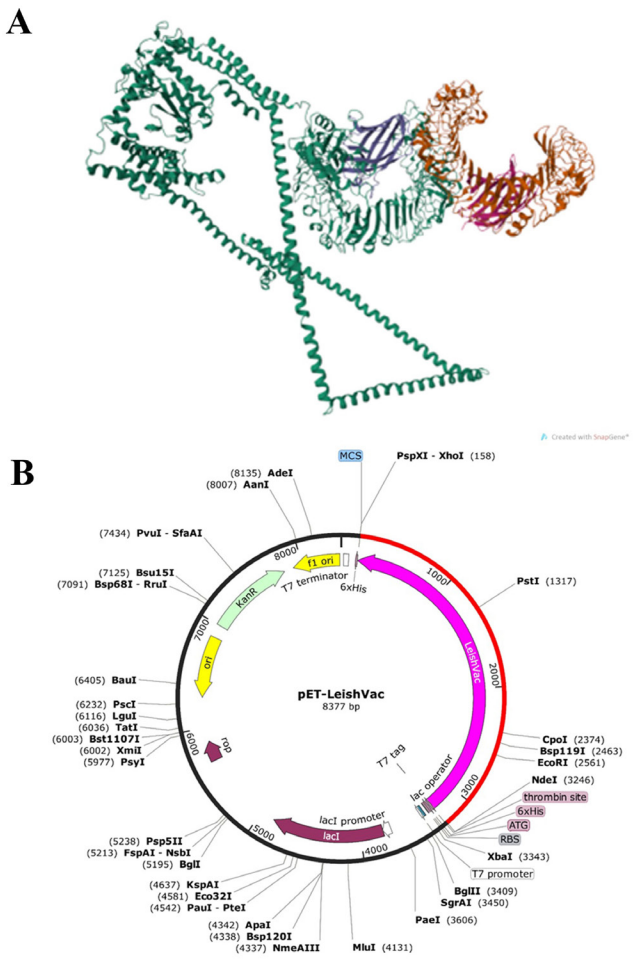


Figure 5. ***(A)*** *The representation of the docked LIDVC-TLR4 complexes was analyzed by the SwarmDock server.* ***(B)*** *In silico cloning of the LIDVC sequences into pET28a (+) vector as pET-LeishVac construct using SnapGene 5.3.1.*

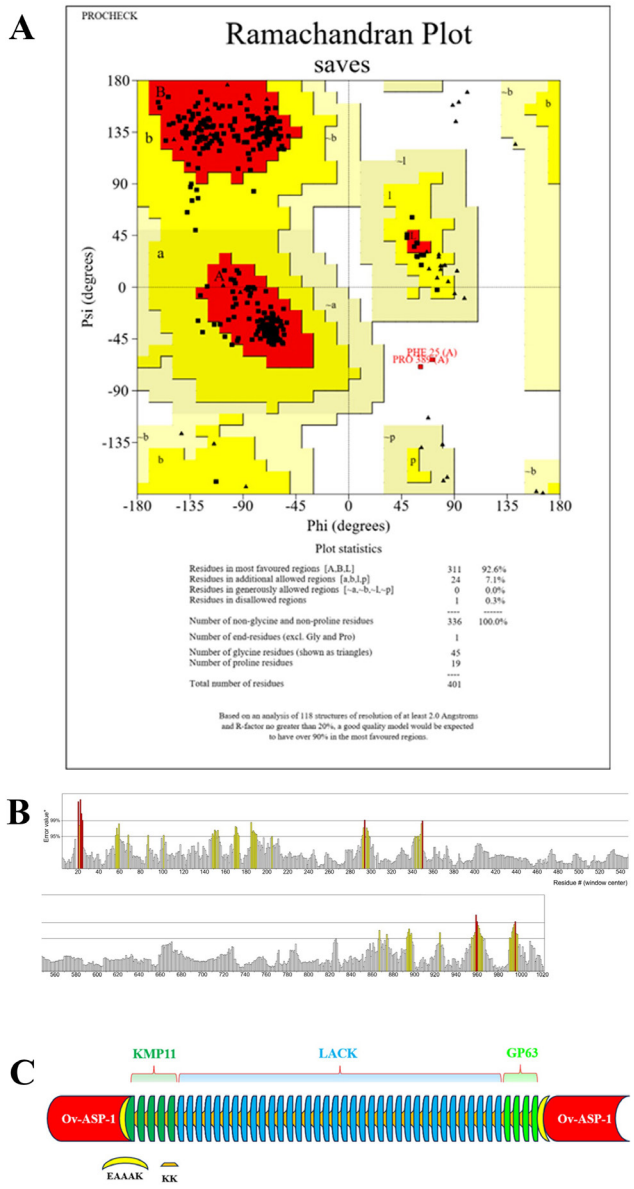


Figure 6. ***(A)*** *The 3D vaccine models were validated through refinement using Ramachandran plot analysis with the PROCHECK web tool.* ***(B)*** *The ERRAT diagram illustrates the quality of the modeled structure of LIDVC. It includes two lines on the error axis, representing the confidence level for rejecting regions surpassing the error value. *Expressed as the percentage of the protein for which the calculated error value falls below the 99% rejection limit indicates the overall quality. High-resolution structures typically yield values around 95% or higher, while lower resolutions (2.5 to 3A) average around 91% or higher. #residue refers to a single unit that makes up a polymer, such as an amino acid in a polypeptide or protein. The overall quality factor of the LIDVC construct is 93.412.* ***(C)*** *Schematic illustration of LIDVC, which contains Ov-ASP-1 as adjuvants in both N- and C-terminal of construct, 15 epitopes from KMP11, five from LACK, and 26 from GP64 antigens, and KK or EAAAK as linkers between them.*

DOI: 10.66224/hjcm.1.6.4

been shown in Fig. 5E. Active T-Helper cell population rose to 2,200 cell/mm3 after ten days (**Fig. 4F**). The active cytotoxic T-lymphocytes population reached ~ 1000 cells/mm3 after 15 days (**Fig. 4G**). The number of resting cytotoxic T-cells population dropped slightly and remained at ~100 cells/mm3 with just a single dose of LIDVC. Finally, in contrast to the very low level of anti-inflammatory cytokine IL-10, the level of pro-inflammatory cytokine IFN-γ was vigorously elevated (**Fig. 4H**).

3.9 Docking Analysis

To further evaluate the structural dynamics of the docked TLR4-vaccine complex, the CABS-flex server (https://biocomp.chem.uw.edu.pl/CABSflex2) was employed. This Python package specializes in rapid simulations of protein structure flexibility, utilizing the CABS model to simulate protein dynamics (**Fig. 6A**). Through all-atom representation models, the CABS-flex server enables the detailed analysis of different modeling results (11). This approach not only validates the stability of the docked complex but also provides insights into its dynamic behavior, ensuring the reliability of the predicted interactions in the context of vaccine development.

3.10 Codon Optimization and in silico Cloning

The final construct-generated protein underwent reverse translation and codon optimization using EMBOSS Backtranseq and the Java Codon Adaptation Tool (JCAT) to ensure high-level protein expression in E. coli K12. The optimized nucleotide sequence achieved a Codon Adaptation Index (CAI) of 0.92, indicating good potential for expression in the chosen host, as a CAI of >0.8 is considered favorable. Additionally, the CG content, which falls within the optimum range of 30-70%, is crucial for efficient gene expression, with our sequence having a mean CG content of 47.04%. Finally, to facilitate gene cloning in the pET28a (+) vector using the SnapGene 5.3.1 tool, NcoI, and XhoI restriction sites were introduced at the N and C-terminals of the sequence, respectively (**Fig. 6B**).

4. Discussion

Leishmaniasis remains a significant health concern in subtropical regions, posing a considerable annual burden on affected populations (12). While various chemotherapeutic treatments for leishmaniasis are associated with adverse effects, the recent progress in vaccine development offers a promising strategy for enhancing public health and controlling this infectious disease (10, 13, 14).

The in silico design of the multi-epitope vaccine candidate LIDVC against L. infantum has undergone rigorous analyses, resulting in the identification of a viable immunogen. By incorporating high-scoring epitopes from GP63, KMP11, and LACK antigens alongside the adjuvant Ov-ASP-1, the construct demonstrates notable immunogenic and antigenic properties. The use of Python programming has been integral to this study, facilitating a variety of computational analyses through its robust libraries and tools. Noteworthy methods included VaxiJen v2.0 (15) for antigenicity evaluation, AllerTop v2.0 for allergenicity assessment, and ANTIGENpro for immunogenicity predictions and IEDB for Population coverage, which collectively provided significant insights into the vaccine candidate's immunological characteristics and its potential to induce a powerful immune response.

Docking studies with TLR4 and flexibility analyses via CABS-Flex further corroborated the vaccine candidate's potential efficacy and dynamic behavior. Moreover, the codon optimization and physicochemical parameter analyses have affirmed the feasibility of scaling up production and highlighted the stability of the LIDVC construct. The use of CAIcal (16) for codon optimization has enhanced our understanding of the construct's expression potential and immunogenicity. Together, these analyses have established a solid foundation for the development of a viable vaccine against.

Moreover, the validation of the 3D structure using state-of-the-art methods such as DeepCDpred (17) and AlphaFold (18) It has provided valuable insights into the structural integrity of the LIDVC construct, reinforcing its potential as an effective immunogen. The comprehensive in silico approach has set the stage for subsequent experimental validation and preclinical studies, aligning with the current trend in vaccine development, where in silico analyses play a crucial role in guiding experimental and clinical studies to accelerate the translation of vaccine candidates from the laboratory to clinical use.

In conclusion, the in silico analysis of the LIDVC vaccine candidate against L. infantum has demonstrated its potential as a promising immunogen. Integrating advanced computational tools and methodologies has provided valuable insights into the candidate's antigenic, structural, and immunogenic properties, laying the groundwork for future experimental validation and preclinical studies. These findings underscore the sig-

nificance of further investigation to assess the vaccine candidate's efficacy and safety in vivo, with the ultimate goal of contributing to developing a successful vaccine against L. infantum.

5. Conclusion

In conclusion, our comprehensive in silico analyses and design of the multi-epitope vaccine candidate, LIDVC, against L. infantum have provided compelling evidence of its potential as an effective and safe vaccine. The antigenicity, allergenicity, and immunogenicity assessments, as well as the docking studies and flexibility analysis, collectively support the vaccine candidate's promising efficacy.

Codon optimization, physicochemical parameters analysis, and validation of the 3D structure have further reinforced the construct's suitability for continued development as a vaccine. These findings underscore the significance of further investigation to assess the vaccine candidate's efficacy and safety in vivo, with the ultimate goal of contributing to developing a successful vaccine against L. infantum.

Our research has laid a strong foundation for subsequent experimental validation and preclinical studies, and we are excited to continue our efforts to advance the development of a safe and effective vaccine against L. infantum.

6. Declarations

Ethical approval and consent to participate

Not applicable.

Availability of data and materials

The datasets of Python programming for this study can be found in a repository on GitHub with the following access link: https://github.com/Amirlrj128/bioinf-applications.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Financial support

There is no funding for this work. This research was Self-funded.

Author's contributions

All authors were involved in conceptualizing and designing the methods, analyzing and interpreting the data, and writing the manuscript. All authors have read and approved the final version of the manuscript.

Acknowledgements

There are no acknowledgements.

Artificial intelligence statement

The authors used an artificial intelligence (AI)-assisted language tool solely to improve the readability and language of the manuscript. All scientific content, study design, data collection, data analysis, interpretation of the findings, and final editorial decisions were performed entirely by the authors. The authors have reviewed and approved the final manuscript and accept full responsibility for its accuracy and integrity.

7. References

1. Alvar J, Vélez ID, Bern C, Herrero M, Desjeux P, Cano J, et al. Leishmaniasis worldwide and global estimates of its incidence. PLoS One. 2012;7(5):e35671. doi: 10.1371/journal.pone.0035671.

2. Pradhan S, Schwartz RA, Patil A, Grabbe S, Goldust M. Treatment options for leishmaniasis. Clinical and Experimental Dermatology. 2022;47(3):516–21. doi: 10.1111/ced.14919.

3. Krishnamoorthy P, Raj AS, Kumar H. Machine learning-driven blood transcriptome-based discovery of SARS-CoV-2 specific severity biomarkers. J Med Virol. 2023;95(2):e28488. doi: 10.1002/jmv.28488.

4. Margaroni M, Agallou M, Tsanaktsidou E, Kammona O, Kiparissides C, Karagouni E. Immunoinformatics Approach to Design a Multi-Epitope Nanovaccine against Leishmania Parasite: Elicitation of Cellular Immune Responses. Vaccines (Basel). 2023;11(2). doi: 10.3390/vaccines11020304.

5. Yadagiri G, Singh A, Arora K, Mudavath SL. Immunotherapy and immunochemotherapy in combating visceral leishmaniasis. Front Med (Lausanne). 2023;10:1096458. doi: 10.3389/fmed.2023.1096458.

6. Nagill R, Kaur S. Vaccine candidates for leishmaniasis: a review. Int Immunopharmacol. 2011;11(10):1464–88. doi: 10.1016/j.intimp.2011.05.008.

7. Al-Qahtani AA, Alhamlan FS, Al-Qahtani AA. Pro-Inflammatory and Anti-Inflammatory Interleukins in Infectious Diseases: A Comprehensive Review. Tropical Medicine and Infectious Disease. 2024;9(1):13. doi: 10.3390/tropi-

calmed9010013.

8. Costa-da-Silva AC, Nascimento DO, Ferreira JRM, Guimarães-Pinto K, Freire-de-Lima L, Morrot A, et al. Immune Responses in Leishmaniasis: An Overview. Trop Med Infect Dis. 2022;7(4). doi: 10.3390/tropicalmed7040054.

9. Larijani A, Kia-Karimi A, Rahmati M, Derakhshan R. Computational Prediction of Leishmania Infantum Epitopes: A Bioinformatic-based Step to Leishmaniasis Vaccine Design. Journal of Visualized Medicine. 2024;1(6):–. Doi: 10.30699/jovm.006.e0109

10. Ponte-Sucre A, Gamarro F, Dujardin JC, Barrett MP, López-Vélez R, García-Hernández R, et al. Drug resistance and treatment failure in leishmaniasis: A 21st century challenge. PLoS Negl Trop Dis. 2017;11(12):e0006052. doi: 10.1371/journal.pntd.0006052.

11. Kurcinski M, Oleniecki T, Ciemny MP, Kuriata A, Kolinski A, Kmiecik S. CABS-flex standalone: a simulation environment for fast modeling of protein flexibility. Bioinformatics. 2019;35(4):694–5. doi: 10.1093/bioinformatics/bty685.

12. Knight CA, Harris DR, Alshammari SO, Gugssa A, Young T, Lee CM. Leishmaniasis: Recent epidemiological studies in the Middle East. Frontiers in Microbiology. 2023;13. doi: 10.3389/fmicb.2022.1052478.

13. Ghorbani M, Farhoudi R. Leishmaniasis in humans: drug or vaccine therapy? Drug Des Devel Ther. 2018;12:25–40. doi: 10.2147/DDDT.S146521.

14. Srivastava S, Shankar P, Mishra J, Singh S. Possibilities and challenges for developing a successful vaccine for leishmaniasis. Parasites & Vectors. 2016;9(1):277. doi: 10.1186/s13071-016-1553-y.

15. Ong E, Cooke MF, Huffman A, Xiang Z, Wong MU, Wang H, et al. Vaxign2: the second generation of the first Web-based vaccine design program using reverse vaccinology and machine learning. Nucleic Acids Research. 2021;49(W1):W671–W8. doi: 10.1093/nar/gkab279.

16. Puigbò P, Bravo IG, Garcia-Vallve S. CAIcal: A combined set of tools to assess codon usage adaptation. Biology Direct. 2008;3(1):38. doi: 10.1186/1745-6150-3-38.

17. Ji S, Oruç T, Mead L, Rehman MF, Thomas CM, Butterworth S, et al. DeepCDpred: Inter-residue distance and contact prediction for improved prediction of protein structure. PLOS ONE. 2019;14(1):e0205214. doi: 10.1371/journal.pone.0205214.

18. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021;596(7873):583–9. doi: 10.1038/s41586-021-03819-2.

19. Doytchinova IA, Flower DR. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinformatics. 2007;8:4. doi: 10.1186/1471-2105-8-4.

20. Wei L, Ye X, Sakurai T, Mu Z, Wei L. ToxIBTL: prediction of peptide toxicity based on information bottleneck and transfer learning. Bioinformatics. 2022;38(6):1514–24. doi: 10.1093/bioinformatics/btac006.

21. Savsani K, Jabbour G, Dakshanamurthy S. A New Epitope Selection Method: Application to Design a Multi-Valent Epitope Vaccine Targeting HRAS Oncogene in Squamous Cell Carcinoma. Vaccines (Basel). 2021;10(1). doi: 10.3390/vaccines10010063

22. Angamuthu K, Piramanayagam S. Evaluation of in silico protein secondary structure prediction methods by employing statistical techniques. Biomedical and Biotechnology Research Journal (BBRJ). 2017;1:29. Doi: 10.4103/bbrj.bbrj_28_17

23. Gasteiger E, C H, A G, Duvaud S, M.R W, Appel R, et al. In The Proteomics Protocols Handbook2005.

24. Torchala M, Bates P. Predicting the Structure of Protein–Protein Complexes Using the SwarmDock Web Server. Methods in molecular biology (Clifton, NJ). 2014;1137:181–97. doi: 10.1007/978-1-4939-0366-5_13.

25. Grote A, Hiller K, Scheer M, Münch R, Nörtemann B, Hempel DC, et al. JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. Nucleic Acids Research. 2005;33(suppl_2):W526–W31. doi: 10.1093/nar/gki376.

26. Buchan DWA, Minneci F, Nugent TCO, Bryson K, Jones DT. Scalable web services for the PSIPRED Protein Analysis Workbench. Nucleic Acids Research. 2013;41(W1):W349–W57. doi: 10.1093/nar/gkt381.

27. Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Research. 2007;35(suppl_2):W407–W10. doi: 10.1093/nar/gkm290.

28. Ko J, Park H, Heo L, Seok C. GalaxyWEB server for protein structure prediction and refinement. Nucleic Acids Research. 2012;40(W1):W294–W7. doi: 10.1093/nar/gks493.

29. Sabatier M, Birsen R, Lauture L, et al. C/EBPα Confers Dependence to Fatty Acid Anabolic Pathways and Vulnerability to Lipid Oxidative Stress-Induced Ferroptosis in FLT3-Mutant Leukemia. Cancer Discov. 2023;13(7):1720-1747. doi:10.1158/2159-8290.CD-22-0411