

# Epitope Mapping and Vaccine Candidate Prediction for Listeriosis and Tetanus

Naravula Jalaja\*, Mundlamuri Ramesh and Yarramane Ajay

Department of Biotechnology, Vignana's Foundation for Science, Technology and Research, Vadlamudi, Guntur 522 213, INDIA  
\*jalajanaravula@gmail.com

## Abstract

*In the pursuit to combat infectious diseases, the development of vaccines remains a cornerstone of public health strategy. This study presents a novel in silico approach to the design of multi-epitope vaccines against *Listeria monocytogenes* and *Clostridium tetani*, pathogens responsible for significant morbidity and mortality worldwide. Utilising advanced computational tools, we identified and characterised antigenic and non-allergenic epitopes from surface proteins of these bacteria, essential for eliciting a targeted immune response. Our analysis included the prediction of linear B-cell epitopes using the ABCPred server, assessment of antigenicity with VaxiJen and determination of epitope orientation within the cellular membrane via DeepTMHMM. Molecular dynamics simulations provided insights into the stability and interactions of the protein-peptide complexes, with RMSD and RMSF analyses confirming the structural integrity conducive to vaccine efficacy. The strategic linking of shortlisted epitopes, facilitated by the KK linker, led to the construction of vaccine candidates with broad protective capabilities.*

*Our findings demonstrate the potential of computational methods in streamlining vaccine development, offering a blueprint for rapid and efficient generation of vaccine candidates against complex pathogens. The implications of this research are far-reaching, providing a method that is both scientifically robust and technically sound, capable of addressing the urgent need for new vaccines in the face of emerging infectious diseases.*

**Keywords:** Vaccine Development, Multi-Epitope Vaccine, *Listeria monocytogenes*, *Clostridium tetani*, In silico Analysis, Antigenic Epitopes, Non-Allergenic Epitopes, Molecular Dynamics Simulation, RMSD and RMSF Analysis, Protein-Peptide Interactions, Computational Immunology, Transmembrane Proteins, ABCPred Server, VaxiJen Server, DeepTMHMM, Protein-Ligand Interaction Profiler (PLIP).

## Introduction

The global endeavour to combat infectious diseases has been revolutionised by the advent of vaccines, which have now expanded their reach to address a myriad of conditions including cancer, Ebola and allergies. The intersection of

biotechnology and vaccine development has been particularly transformative, enabling the detailed elucidation of pathogen molecular structures and the discovery of protective antigens. These advancements have shifted the paradigm from empirical vaccine research to a more rational, evidence-based approach, significantly reducing the error rates and costs associated with traditional vaccine development methods.<sup>4</sup>

A key aspect of this progress lies in the understanding of the human leukocyte antigen (HLA) system, a set of proteins found on the surface of cells that plays a critical role in the immune response. HLAs are responsible for presenting peptide fragments including pathogen-derived epitopes, to T cells, thereby triggering an immune response. The specificity of HLA-peptide binding is paramount in designing vaccines that can elicit a robust and targeted immune response.

In our context, the HLA-A\*02:01 allele emerges as a key player due to its widespread prevalence and its ability to present a diverse range of epitopes to T cells. Epitopes that bind to HLA-A\*02:01 have the potential to induce strong immune responses and are therefore of particular interest in vaccine development.<sup>4,17,19,20</sup>

Another promising development in this field is the concept of cross-immunity which posits that vaccines designed for one disease may confer protection against multiple pathogens. This is achieved through the activation of both innate and adaptive immunity, with innate immune cells like dendritic cells and macrophages playing a pivotal role in inducing a broad immune response. This is particularly relevant in the context of HLA molecules, as their polymorphic nature can influence the repertoire of epitopes presented and the breadth of the immune response elicited. Thus, identifying epitopes that bind promiscuously to various HLA alleles can be instrumental in developing multivalent vaccines.<sup>4,17,19,20</sup>

In the pursuit for multivalent vaccines, researchers are actively seeking bacterial epitopes that can elicit such cross-immunity. Identifying common virulence factors and immune structure domains across different bacterial species is crucial for this endeavour. Despite the lack of sequence homology, certain virulence factors such as the pore-forming toxins *Listeriolysin O* from *Listeria monocytogenes*, *pneumolysin* from *Streptococcus pneumoniae* and the ESX-1 secretion system components ESAT-6 and CFP-10 from *Mycobacterium tuberculosis*, share functional similarities that can be exploited for vaccine development.<sup>5,14</sup>

Furthermore, the glycolytic enzyme glyceraldehyde-3-phosphate dehydrogenase (GAPDH), found in these pathogens, has been identified as a key player in virulence and immunogenicity. GAPDH's ability to bind to various immune-related proteins and intracellular transport GTPases presents a unique opportunity to identify broad immunogenic epitopes suitable for the development of multivalent vaccines.<sup>1,3,8,9,11,13,15,18,21</sup>

This study aims to harness these insights to predict B-cell epitopes and to develop vaccine candidates against *Listeria monocytogenes* and *Clostridium tetani*. By leveraging computational techniques to map epitopes and design vaccines, we aspire to contribute to the global effort in creating more effective and comprehensive immunisation strategies against these formidable pathogens. We seek to identify epitopes that bind promiscuously to various HLA alleles, specifically HLA-A\*02:01, thereby maximising the potential for inducing broad and robust immune responses. Additionally, we aim to elucidate the molecular interactions between these epitopes and their corresponding proteins, shedding light on potential vaccine targets and mechanisms of action. Ultimately, our goal is to pave the way for the development of novel vaccines capable of conferring broad protection against infections caused by *Listeria monocytogenes* and *Clostridium tetani*.

## Material and Methods

### *In silico* analysis of HLA-A\*02:01:

**Protein structure:** HLA-A\*02:01 is one of the most common human leukocyte antigen (HLA) alleles in many populations worldwide. Its prevalence makes it an attractive target for vaccine development as it has the potential to provide broad coverage across diverse populations. We extracted the three dimensional PDB structures of proteins analogous to the HLA-A\*02:01 epitope from the PDB database. We downloaded and examined the three dimensional structure of the protein 4U6Y which is similar to the HLA-A\*02:01 epitope. The protein's spatial configurations were analyzed and their formats were converted for compatibility with docking software (AutoDock).

**Identification and preparation of peptide:** We have used IEDB tool to identify the potential peptides that can bind to HLA-A\*02:01 epitope by giving our protein of interest(4U6Y) and selecting HLA-A\*02:01 as the allele for prediction. Based on their IC<sub>50</sub> values, the peptides were selected from the IEDB as shown in table 1. Peptide sequences were converted into PDB format using the Open babel tool.

**Molecular Docking:** In the process of molecular docking and binding energy estimation, the ADCP Crank pep toolkit was employed as a robust de novo method for the docking of proteins and peptides. This toolkit utilises the potential energy generated by the given peptide to fold it, thereby facilitating the interaction between the peptide and the

protein receptor. ADCP, a peptide docking software, orchestrates the sampling and folding of peptides by manipulating the backbone conformation. To ascertain the potential energy of the receptor, AutoDock, in conjunction with ADCP, utilises affinity maps. By utilising AutoGridFR, binding pocket affinity maps of the protein, previously unknown for the peptides binding sites, are generated. Subsequently, a target file is generated and the peptides are docked according to their sequence onto the protein receptor.

The assessment of the docking process hinges on the determination of binding affinities, measured in terms of kilocalories per mole (kcal/mol). A lower negative energy value indicates a more stable system and a stronger binding relationship between the peptide and the protein receptor. This process serves as a pivotal step in understanding the molecular interactions between proteins and peptides, providing crucial insights into potential therapeutic targets and drug design strategies.

**Molecular Dynamics Simulation:** Molecular dynamics (MD) simulation, a technique employed to examine atomic interactions over time, was utilised to analyse protein-peptide interactions in this study. The GROMACS software package, incorporating CHARMM-36 force fields, was employed to manipulate atom mobility and study molecular interactions. The parameters of the force fields defined the forces acting on particles within the system.

To maintain the system's equilibrium, solvation was created and Cl<sup>-</sup> and Na<sup>+</sup> ions were incorporated to preserve elemental balance and form the system's topology. The simulations were conducted at 1 bar and 310K to optimise forces below 10 KJ/mol. The Berendsen thermostat and barostat algorithms were utilised to control pressure and temperature by modifying volumes. MD simulations were run for a duration of 10 nanoseconds (ns) with a time step of 2 femtoseconds (fs) using the leapfrog approach to facilitate analysis. These computational methods ensured simulation stability and energy conservation, thereby enabling a comprehensive understanding of molecular characterization and stable arrangements of protein-peptide interactions.

**Trajectory analysis:** In the trajectory analysis, root mean square deviation (RMSD) and root mean square fluctuation (RMSF) were employed to evaluate the stability and interactions observed during protein-peptide stimulation. RMSD quantifies the average deviation of a set of atoms, typically the backbone atoms of the protein from their reference positions over time. On the other hand, RMSF measures the average variation of individual atoms from their reference positions throughout the simulation.

During the analysis, the simulations' trajectories were examined using RMSD and RMSF graphs. These graphs provide insights into the structural changes and fluctuations occurring within the protein-peptide complex over the course of the simulation. RMSD reveals the overall stability

of the complex by tracking the deviation of backbone atoms from their initial positions while RMSF highlights the flexibility and dynamic behaviour of individual atoms within the complex.

Through trajectory analysis, we were able to assess the stability and conformational changes of the protein-peptide complex, gaining valuable insights into the molecular interactions driving their binding affinity. These analyses contribute to a deeper understanding of the structural dynamics underlying protein-peptide interactions and to aid in the rational design of therapeutic interventions.

**Prediction of Multi-Epitope Vaccine Candidates Against *Listeria monocytogenes* and *Clostridium tetani*:** To generate vaccine candidates against *Listeria monocytogenes* and *Clostridium tetani*, we initiated our study by identifying surface proteins of these pathogens from previous studies. The prediction of bacterial membrane proteins was conducted using DeepTMHMM, a tool that calculates posterior probabilities of proteins being located in bacterial membranes. Subsequently, the primary sequences of these proteins were utilised to predict linear B cell epitopes using the ABCPred web server, with a prediction threshold set at 0.8 to ensure high specificity.<sup>6,7,10,12</sup>

The antigenicity of the predicted epitopes was assessed using the VaxiJen v2.0 web server, employing a default threshold of 0.4 to distinguish antigenic epitopes. To evaluate the allergenicity of the identified epitopes, the AllerTOP v2.0 web server was utilised. Epitopes deemed antigenic and non-allergenic were selected for further analysis.<sup>2,22</sup>

In the development of vaccine candidates, a strategic approach to epitope linking was employed. The identified antigenic and non-allergic epitopes were joined using specific linkers and an adjuvant. In one strategy, the B cell epitopes were linked to each other using the KK linker. Subsequently, the linked epitope sequence was connected to the adjuvant through the EAAAK linker.

The selected adjuvant, beta-defensin, was chosen for its dual functionality as an antimicrobial and an immunomodulatory agent<sup>16</sup>. Beta-defensin was connected to the very first epitope in the sequence via the EAAAK linker. This strategic epitope linking approach facilitated the enhancement of the immune response, contributing to the development of potent vaccine candidates against *Listeria monocytogenes* and *Clostridium tetani*.

## Results and Discussion

**In-silico analysis of HLA-A\*02:01:** To investigate the affinity between the reference protein and the putative peptide inhibitor, we employed a comprehensive array of prediction tools including AutoDockFR, AutoDock4 and Auto Dock Crank Pep suites. The binding pocket of the protein was meticulously generated utilising an adapted Metropolis Monte Carlo method, laying the groundwork for

precise docking simulations. The principle guiding peptide folding posits that longer peptides necessitate a greater number of Monte Carlo steps for accurate folding, ensuring meticulous molecular modelling.

Docking scores, a crucial metric indicative of binding affinity, were derived through ADCP docking and meticulously tabulated in table 4 for comprehensive analysis. Subsequently, Pymol software was harnessed to visualise protein-peptide interactions, enabling the ranking of peptides based on their intermolecular interactions. Furthermore, a two-dimensional plot of the complex's secondary structure was meticulously constructed to pinpoint the residues exerting influence on the system's interactions. The process of peptide construction from amino acid sequences was meticulously executed, with docking results further validated against experimental data to ensure accuracy and reliability. The ADCP program, integral to the analysis of protein-peptide interactions, was seamlessly executed on the Linux operating system, underscoring the robustness and versatility of our computational approach.

Of particular note, peptide 2 emerged as a standout candidate, boasting the highest ranking score of all peptides at -9.64 kcal/mol. This exceptional score, highlighted in fig. 1, underscores its potential as a promising inhibitor and emphasises its critical role in facilitating coordinate interactions within the protein-peptide complex.

In figure 2, X-axis, labelled time (ns), indicates the simulation time frame while the Y-axis, labelled RMSD (nm) measures the deviation of the protein structure from a reference conformation. The fluctuating blue line graph reflects changes in the protein conformation over time, with higher RMSD values suggesting greater structural variation. Figure 2 represents the RMSD plot which illustrates the dynamic stability of the protein structure throughout the simulation. Initially, the protein experiences conformational adjustments, reflected by the rising RMSD values. As the simulation progresses, the RMSD stabilises, indicating that the protein has reached a consistent conformation. This stabilisation phase is crucial for reliable interpretations of the protein behaviour in a simulated physiological environment.

Figure 3 represents the graph of root mean square fluctuation (RMSF) plot, which is used to analyse the flexibility of protein structures at the atomic level. The X-axis represents individual atoms or residues within the protein, while the Y-axis shows the RMSF values in nanometers. Peaks in the graph indicate regions of high flexibility, suggesting these areas of the protein are more dynamic and can move more freely. Conversely, areas with lower RMSF values represent more rigid and stable parts of the protein. This is crucial for understanding protein dynamics, as regions with higher flexibility often play key roles in protein function, such as binding sites for other molecules or regions involved in enzymatic activity.

The results of the MD simulations revealed a steady RMSD graph, indicating minimal divergence from the reference protein structure and consistent stability of the protein-peptide complex (Figure 2). Furthermore, the RMSF plots displayed fluctuations in residue positions over time with

some residues exhibiting higher deviations than others (Figure 3). These findings suggest that the complex maintains structural stability while allowing for localized flexibility in certain regions

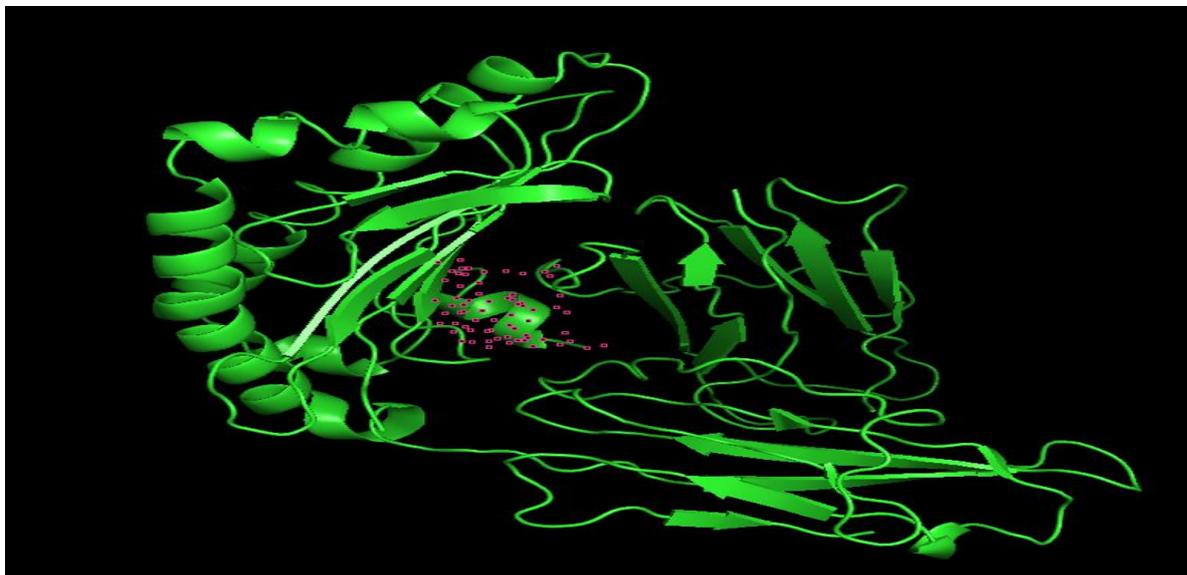


Figure 1: Represents by molecular docking studies shows interactions between protein - peptide

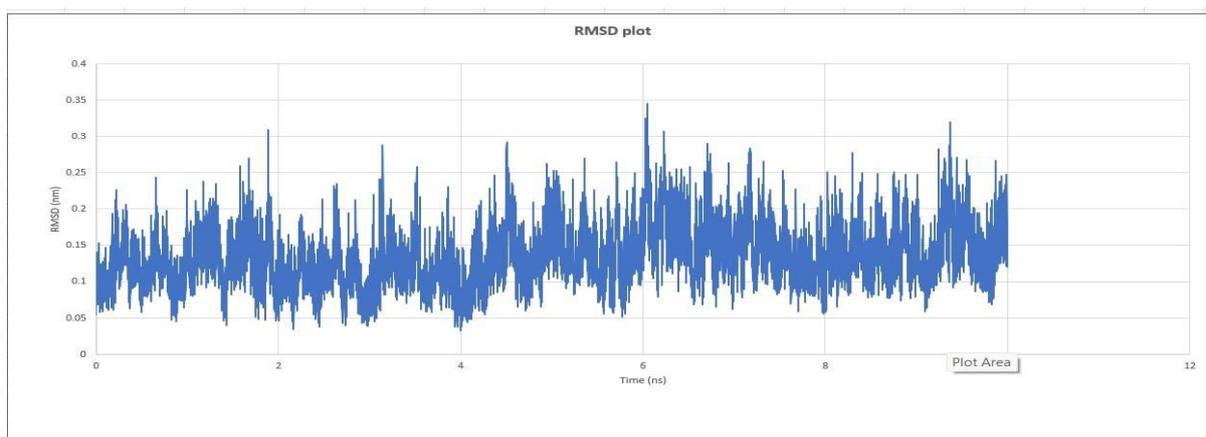


Figure 2: RMSD plot for protein-peptide

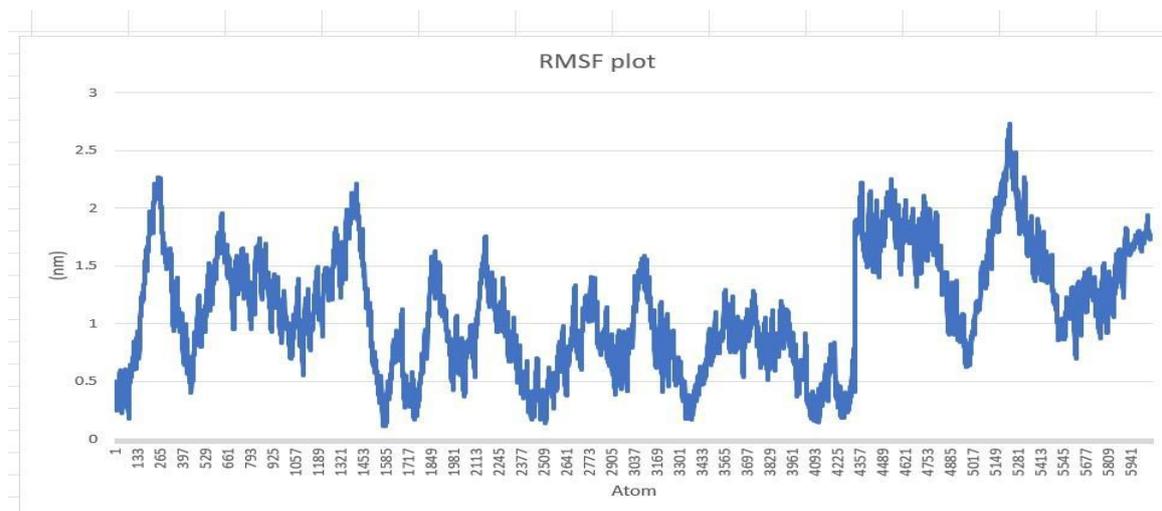


Figure 3: RMSF plot for each Residue

Visual inspection of the complex trajectories using Pymol software corroborated the RMSD and RMSF analyses, confirming the overall stability and dynamic behavior of the protein-peptide interactions. Additionally, analysis using the Protein-Ligand Interaction Profiler (PLIP) tool revealed a significant number of hydrophobic interactions within the peptide, contributing to the stability of the complex. Overall, the MD simulations provided valuable insights into the structural dynamics and stability of protein-peptide interactions, highlighting the robust nature of the studied complex.

#### Prediction of Multi-Epitope Vaccine Candidates against

*Listeria monocytogenes* and *Clostridium tetani*: This study identified 4 surface proteins (LMOF2365\_0578, LMOF2365\_0581, LMOF2365\_0639, LMOF2365\_2117) in *Listeria monocytogenes* and 1 surface layer protein (CTC00462) in *Clostridium tetani* (Table 5). The proteins were shown to be present in the outer region of the membrane (Figures 4 to 8). Our study predicted linear B cell epitopes of 16 amino acids length using ABCPred webserver. It predicted 33 epitopes in LMOF2365\_0578, 15 epitopes in LMOF2365\_0581, 29 epitopes in LMOF2365\_0639, 30 epitopes in LMOF2365\_2117 and 49 epitopes in CTC00462 (Supplementary Tables 1 to 5). Among them, some epitopes are allergenic, non-allergenic, antigenic and non-antigenic.

For a vaccine candidate, epitopes should be antigenic in order to induce immune response and non-allergenic to avoid any allergic reaction in the host. So, epitopes that are antigenic, non-allergenic and lie on the outer region of the membrane have been shortlisted. Finally, we obtained 16 epitopes from protein LMOF2365\_0578, 7 epitopes from LMOF2365\_0581, 13 epitopes from LMOF2365\_0639, 5 epitopes from LMOF2365\_2117, 22 epitopes from CTC00462. Tables 6 to 10 show amino acid sequences of shortlisted antigenic and non-allergenic epitopes with their start position in the protein, BCEPred score and VaxiJen v2.0 score. All of these epitopes reside in the outer region of the membrane (Figures 4 to 8).

The epitopes together have been linked to generate a vaccine candidate. To generate vaccine candidate for *Listeria monocytogenes*, the shortlisted epitopes from proteins LMOF2365\_0578, LMOF2365\_0581, LMOF2365\_0639 and LMOF2365\_2117 have been joined (Table 6-9). To generate vaccine candidates for *Clostridium tetani*, shortlisted epitopes from proteins CTC00462 have also been joined (Table 10). For dual vaccines against both *Listeria monocytogenes* and *Clostridium tetani*, epitopes identified for *Listeria monocytogenes* proteins and *Clostridium tetani* each with KK linker have also been joined (Table 11).

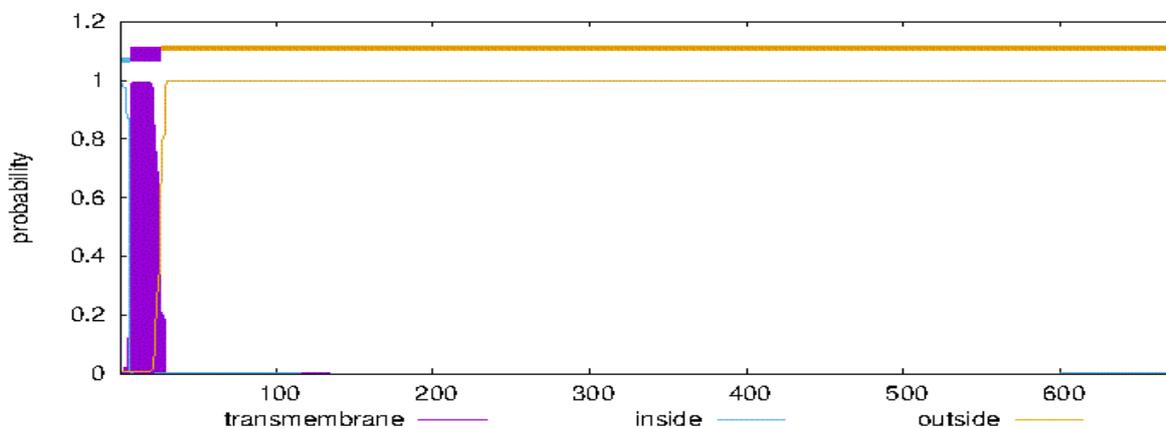


Figure 4: DeepTMHMM posterior probabilities for LMOF2365\_0578.

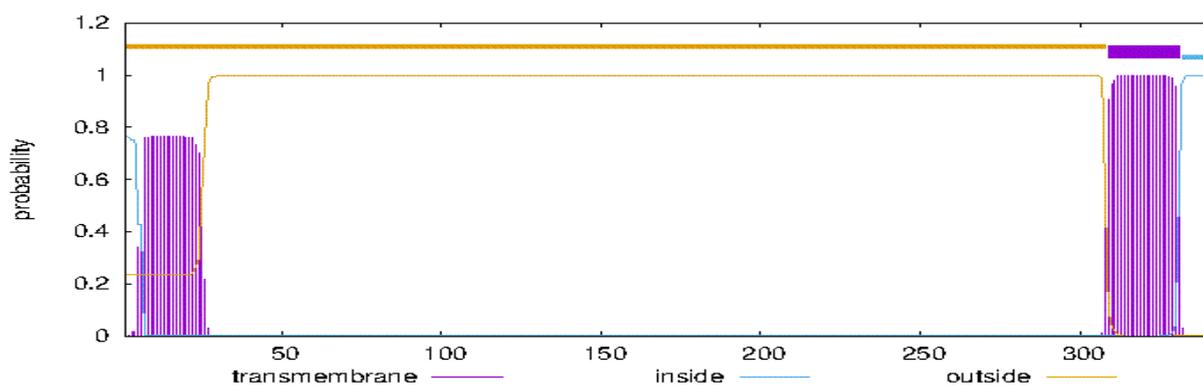
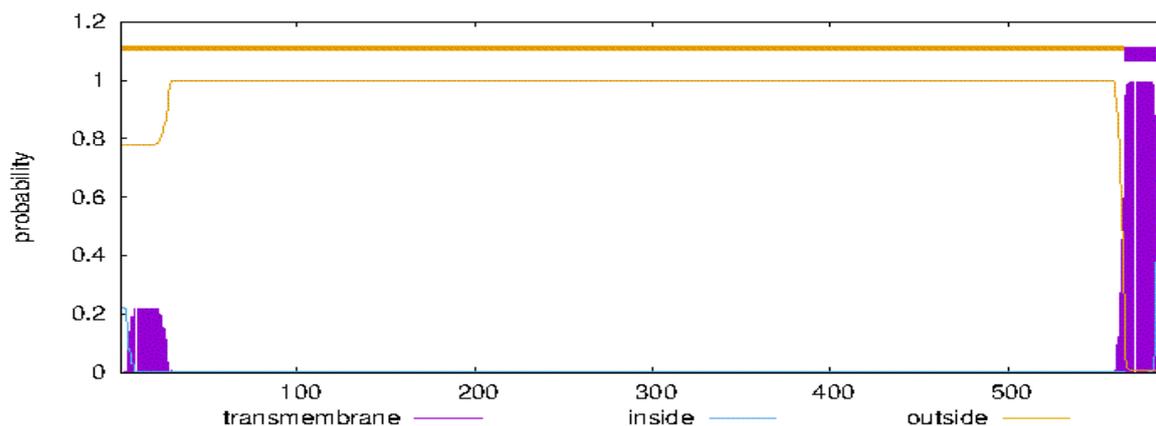
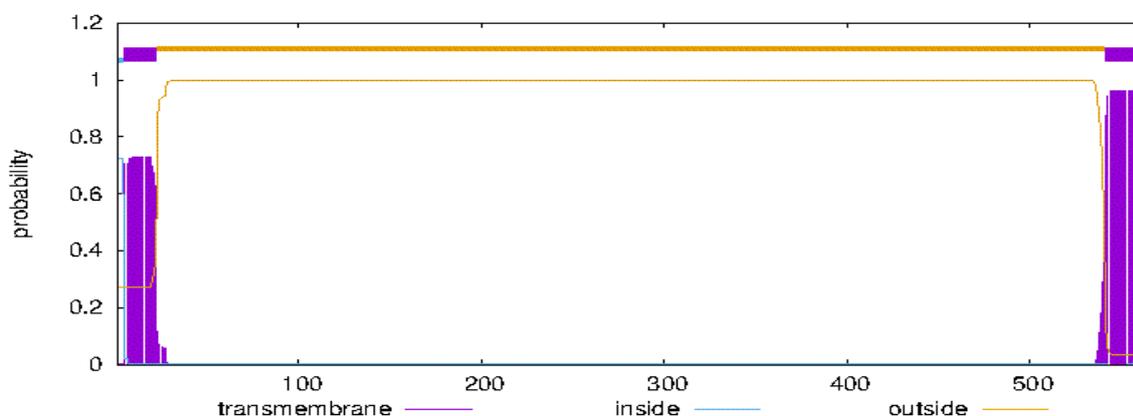


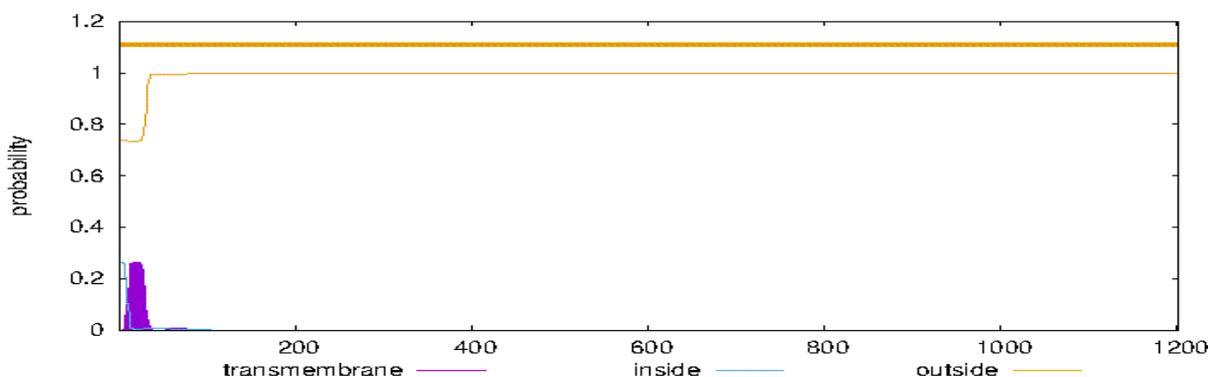
Figure 5: DeepTMHMM posterior probabilities for LMOF2365\_0581



**Figure 6: TMHMM posterior probabilities for LMOF2365\_0639**



**Figure 7: DeepTMHMM posterior probabilities for LMOF2365\_2117**



**Figure 8: DeepTMHMM posterior probabilities for CTC00462**

**Table 1**

**List of predicted epitopes against HLA allele using IEDB**

Peptide	Sequence
Peptide 1	MGQIVTMFE
Peptide 2	GQIVTMFEA
Peptide 3	QIVTMFEAL
Peptide 4	IVTMFEALP
Peptide 5	VTMFEALPH
Peptide 6	TMFEALPHI
Peptide 7	MFEALPHII
Peptide 8	FEALPHIID

**Table 2**  
Hydrogen bonds within the complex

Index	Residue	AA
1	32 A	GLN
2	35A	ARG
3	36A	PHE
4	48A	ARG
5	51B	HIS
6	237A	GLY

**Table 3**  
Hydrophobic Interactions within the complex

Index	Residue	AA
1	27A	TYR
2	32A	GLN
3	33A	PHE
4	35A	ARG
5	48A	ARG

**Table 4**

Docking scores of peptides obtained by ADCP docking

Peptides	Ranking scores (kcal/mol)
Peptide 1	-6.098884
Peptide 2	-9.640905
Peptide 3	-6.789289
Peptide 4	-9.178511
Peptide 5	-2.411339
Peptide 6	-7.535579
Peptide 7	-6.478482
Peptide 8	-6.462157

**Table 5**

Identified surface proteins in *Listeria monocytogenes* and *Clostridium tetani*

Protein name	Length (aa)	Location	NCBI protein accession
LMOF2365_0578	677	Cell wall	WP_003727281.1
LMOF2365_0581	343	Plasma membrane	WP_010958763.1
LMOF2365_0639	589	Cell wall	WP_010958775.1
LMOF2365_2117	562	Cell wall	WP_003731372.1
CTC00462	1202	Cell wall	WP_035124690.1

**Table 6**

Shortlisted B cell epitopes from LMOF2365\_0578

Epitope sequence	Start position	BcePred Score	VaxiJen score	VaxiJen prediction	AllerTOP prediction
DSLMEIEPN DALVND	646	0.94	1.11	Antigen	Nonallergen
IEKKQSAVTD PKYDST	625	0.89	0.58	Antigen	Nonallergen
PELQDDKSSTLKNVNT	36	0.89	1.08	Antigen	Nonallergen
GNTEFTT SVLIKYKPP	437	0.88	0.71	Antigen	Nonallergen
TSKITYSAEVM AKRPK	313	0.87	0.71	Antigen	Nonallergen
PSLEEISVERNNISDV	211	0.87	0.4	Antigen	Nonallergen
ELEVIDRRTVRQGWTI	570	0.86	0.87	Antigen	Nonallergen
KPVKITATNP KATIDP	296	0.86	1.12	Antigen	Nonallergen
EGIQYLPNLYNVQLQF	128	0.86	0.67	Antigen	Nonallergen
ANSTIADLFPDEGMAK	64	0.85	0.41	Antigen	Nonallergen
KYQSKSPSSSPIYLNT	605	0.84	0.6	Antigen	Nonallergen
VGDKNATQFRADVNA	464	0.84	1.11	Antigen	Nonallergen
TVANQLGR TENNFQT	80	0.82	0.89	Antigen	Nonallergen
GSSGKAELEVIDRRTV	564	0.82	1.18	Antigen	Nonallergen
EVM AKRPKYSSNRVSG	321	0.82	1.05	Antigen	Nonallergen
KNLKASPLTIPANSTI	53	0.81	0.52	Antigen	Nonallergen

**Table 7**  
Shortlisted B cell epitopes from LMO2365\_0581

Epitope sequence	Start position	BcePred score	VaxiJen score	VaxiJen prediction	AllerTOP prediction
GLHFQEKKEEDKEKDS	145	0.89	1.71	Antigen	Nonallergen
KEEITVNCVANTAVTN	69	0.86	0.71	Antigen	Nonallergen
NVEITNQSKEEITVNC	61	0.85	1.21	Antigen	Nonallergen
KEEDKEKDSSENDVQI	152	0.85	2.01	Antigen	Nonallergen
YKKGSEKVLHESKRTD	228	0.84	0.93	Antigen	Nonallergen
KVHLVAKNKEDKWEWT	269	0.83	0.71	Antigen	Nonallergen
LELNQIKPATRNYRNV	190	0.81	0.51	Antigen	Nonallergen

**Table 8**  
Shortlisted B cell epitopes from LMO2365\_0639

Epitope sequence	Start position	BcePred Score	VaxiJen score	VaxiJen prediction	AllerTOP prediction
GGSYDHYFDIDHSLTI	336	0.94	0.69	Antigen	Nonallergen
LQDISAKTSDGSKVTS	449	0.93	1.6	Antigen	Nonallergen
PPAPGPDPTDPTPNP	510	0.92	0.91	Antigen	Nonallergen
PTPDPTPNPNPNINP	517	0.91	0.82	Antigen	Nonallergen
DKSISYTKDSTKTDQQ	432	0.91	1.06	Antigen	Nonallergen
TVTIHAKPVITADKSI	420	0.91	0.66	Antigen	Nonallergen
YFDIDHSLTITNDSAI	342	0.86	0.64	Antigen	Nonallergen
EVNAALPNTGDASQAT	552	0.84	0.94	Antigen	Nonallergen
TVLINSSIKSSALNFD	222	0.84	0.55	Antigen	Nonallergen
HAETDDGTPVTSDFNT	375	0.83	0.87	Antigen	Nonallergen
AKIPSLTNLSIAGDNL	111	0.83	0.46	Antigen	Nonallergen
TSSASTYFTLNETKID	268	0.82	0.76	Antigen	Nonallergen
YGEQTTVTEEQFLKDV	359	0.81	0.48	Antigen	Nonallergen

**Table 9**  
Shortlisted B cell epitopes from LMO2365\_2117

Epitope sequence	Start position	BcePred Score	VaxiJen score	VaxiJen prediction	AllerTOP prediction
TKEIRVTIPFNPQKTI	481	0.9	0.6	Antigen	Nonallergen
TFTITYGDTNPVKLTF	393	0.9	1.14	Antigen	Nonallergen
TGTAEGLPPGKYTATE	363	0.9	0.52	Antigen	Nonallergen
QKTINIFSDNKIMVP	493	0.88	0.49	Antigen	Nonallergen
DVKSTDGTTLKKVTTN	436	0.81	0.95	Antigen	Nonallergen

**Table 10**  
Shortlisted B cell epitopes from CTC00462

Epitope sequence	Start position	BcePred Score	VaxiJen score	VaxiJen prediction	AllerTOP prediction
DKTITANKVNVKGDVI	935	0.93	0.87	Antigen	Nonallergen
DGKVDEDTAEDVRNYE	361	0.9	1.46	Antigen	Nonallergen
KRIAKGADRFDTNLKV	229	0.89	0.64	Antigen	Nonallergen
QKVIDTKVEDKANYTE	1167	0.89	1.18	Antigen	Nonallergen
SEAIEKDLKGEKESTG	326	0.88	1.75	Antigen	Nonallergen
DTKITTSKDGKAARLE	822	0.87	2.07	Antigen	Nonallergen
TIEINDKLKDAYGNKI	642	0.87	1.25	Antigen	Nonallergen
STVKMDAEPMTLKKDT	891	0.84	0.74	Antigen	Nonallergen
GFEADYSEVKNSIKAK	502	0.84	1.12	Antigen	Nonallergen
SEKNNDKLFKVTVTE	1132	0.84	1.48	Antigen	Nonallergen
VEKGNAAGDKDWAUNI	1018	0.84	0.77	Antigen	Nonallergen

LGAKNIYIVGGKGVVT	105	0.82	0.7	Antigen	Nonallergen
DFAGRTIEKEDLSKDK	979	0.81	0.68	Antigen	Nonallergen
MKTSTTNKVENYKFKD	796	0.81	1.18	Antigen	Nonallergen
IRVIYSKDVSEKVAKE	687	0.81	0.89	Antigen	Nonallergen
PSGTNKLYTPDGKDGE	535	0.81	0.9	Antigen	Nonallergen
LNLNQIKVVFDGKVDE	351	0.81	1.08	Antigen	Nonallergen
VERIGGNSRYETNAEI	133	0.81	0.93	Antigen	Nonallergen
THIYTLDEGTERLQK	1077	0.81	0.48	Antigen	Nonallergen
ILLTDASDKPSADLTA	85	0.8	0.65	Antigen	Nonallergen
DGKVRDLPSDTKITTS	813	0.8	1.09	Antigen	Nonallergen
KVTVTEADTKVDQASK	1142	0.8	1.33	Antigen	Nonallergen

Table 11

VaxiJen and AllerTop Prediction: Predicted vaccine candidates generated from the B cell epitopes identified from the proteins LMOF2365\_0578, LMOF2365\_0581, LMOF2365\_0639, LMOF2365\_2117 of *Listeria monocytogenes* and CTC00462 of *Clostridium tetani*

<p><b>Vaccine candidate derived from the proteins LMOF2365_0578, LMOF2365_0581, LMOF2365_0639, LMOF2365_2117 of <i>Listeria monocytogenes</i></b></p> <p><b>VaxiJen prediction: Antigen (score: 0.89), AllerTOP prediction: non-allergen (with adjuvant)</b></p> <p><b>VaxiJen prediction: Antigen (score: 0.90), AllerTOP prediction: non-allergen (without adjuvant)</b></p> <p>GIINTLQKYYCRVRGGRCVLSCLPKEEQIGKCSTRGRKCCRRKKEAAAKDLSLSMEIEPN DAL VNDK<del>KK</del>IEKKQSAVTDPKYDST<del>KK</del>PELQDDKSSTLKNVNT<del>KK</del>GNTEFTT<del>SV</del>LIKYKPP<del>KK</del>TSKI TYSAEVMAKRP<del>KK</del>PSLEEISVERNNDISV<del>KK</del>ELEVIDRRTVRQGWTI<del>KK</del>KPVKITATNPKATI DP<del>KK</del>KEGIQYLPNLVYNVQLQF<del>KK</del>ANSTIADLFPDEGMAK<del>KK</del>KYQSKSPSSSPIYLN<del>TK</del>VGDKV NATQFRADVNA<del>KK</del>TVANQLGRTENNNFQT<del>KK</del>GSSGKAELEVIDRRTV<del>KK</del>EVMAKRPKYSSN RVSG<del>KK</del>KNLKASPLTIPANSTI<del>KK</del>GLHFQEKKEEDKEKDS<del>KK</del>KKEITVNCVANTAVTN<del>KK</del>NVE ITNQSKEEITVNC<del>KK</del>KEEDKEKDSSENDVQI<del>KK</del>YKKGSEKVLHESKRTD<del>KK</del>KVHLVAKNKED KWEWT<del>KK</del>LELNQIKPATRNYRNV<del>KK</del>GGSYDHYFDIDHSLTI<del>KK</del>LQDISAKTSDGSKVTS<del>KK</del>PP APGPDPTPDPTPNP<del>KK</del>PTPDPTPNPNPNINP<del>KK</del>DKSISYTKDSTKTDQ<del>KK</del>TVTIHAKPVITAD KSI<del>KK</del>YFDIDHSLTITNDSA<del>KK</del>EVNAALPNTGDASQAT<del>KK</del>TVLINSSIKSSALNFD<del>KK</del>HAETDD GTPVTSDFNT<del>KK</del>AKIPSLTNLSIAGDNL<del>KK</del>TSSASTYFTLNETKID<del>KK</del>YGEQTTVTEEQFLKDV <del>KK</del>TKEIRVTIPFNPQTI<del>KK</del>TFTITYGDTNPVKLTF<del>KK</del>TGTAEGLPPGKYTATE<del>KK</del>QKTINITF SDNKIMVP<del>KK</del>DVKSTDGTTLKKVTTN</p>
<p><b>Vaccine candidate derived from the protein CTC00462 of <i>Clostridium tetani</i></b></p> <p><b>VaxiJen prediction: antigenic (score: 1.01), AllerTOP prediction: non-allergenic (with adjuvant)</b></p> <p><b>VaxiJen prediction: antigenic (score: 1.05), AllerTOP prediction: non-allergenic (without adjuvant)</b></p> <p>GIINTLQKYYCRVRGGRCVLSCLPKEEQIGKCSTRGRKCCRRKKEAAAKDKTITANKVNVK GDVI<del>KK</del>DGKVD<del>ED</del>TAEDVRN<del>YE</del>KKKRIAKGADRFD<del>TNL</del>KV<del>KK</del>QKVIDTKVEDKANYTE<del>KK</del>S EAIEKDLKGEKESTG<del>KK</del>DTKITTSKDGKAARLE<del>KK</del>TIEINDKLKDAYGNKI<del>KK</del>STVKMDAEP MTLKKDT<del>KK</del>GFEADYSEVKNSIKAK<del>KK</del>SEKNNDKLFKVTVTE<del>KK</del>VEKGNAAAGDKDWAVNI <del>KK</del>LGAKNYIVGGKGVVT<del>KK</del>DFAGRTIEKEDLSKDK<del>KK</del>MKTSTTNKVENYKFKD<del>KK</del>IRVIYS KDVSEKVAKE<del>KK</del>PSGTNKLYTPDGKDGE<del>KK</del>LNLNQIKVVFDGKVDE<del>KK</del>VERIGGNSRYETNA EI<del>KK</del>THIYTLTDEGTERLQK<del>KK</del>ILLTDASDKPSADLTA<del>KK</del>DGKVRDLPSDTKITTS<del>KK</del>KVTVTE ADTKVDQASK</p>
<p><b>Vaccine candidate derived from the proteins LMOF2365_0578, LMOF2365_0581, LMOF2365_0639, LMOF2365_2117 of <i>Listeria monocytogenes</i> and CTC00462 of <i>Clostridium tetani</i></b></p> <p><b>VaxiJen prediction: antigenic (score: 0.95), AllerTOP prediction: non-allergenic (with adjuvant)</b></p> <p><b>VaxiJen prediction: antigenic (score: 0.96), AllerTOP prediction: non-allergenic (without adjuvant)</b></p> <p>GIINTLQKYYCRVRGGRCVLSCLPKEEQIGKCSTRGRKCCRRKKEAAAKDLSLSMEIEPN DAL VNDK<del>KK</del>IEKKQSAVTDPKYDST<del>KK</del>PELQDDKSSTLKNVNT<del>KK</del>GNTEFTT<del>SV</del>LIKYKPP<del>KK</del>TSKI</p>

TYSAEVMAKRPK**KK**PSLEEISVERNNISDV**KE**ELEVIDRRTVRQGW**TI****KK**KPVKITATNPKATI  
 DP**KK**KEGIQYLPNLYNVQLQF**KK**ANSTIADLFPDEGMAK**KK**KYQSKSPSSSPIYLNT**KK**VGDKV  
 NATQFRADVNA**KK**TVANQLGRTE>NNNFQ**TK**GSSGKAELEVIDRRTV**KE**EVMAKRPKYSSN  
 RVSG**KK**KNLKASPLTIPANSTI**KK**GLHFQEKKEEDKEKDS**KK**KKEITVNCVANTAVTN**KK**NVE  
 ITNQSKEEITVNC**KK**KEEDKEKDSSENDVQI**KK**YKKGSEKVLHESKRTD**KK**KVHLVAKNKED  
 KWEWT**KK**LELNQIKPATRNYRNV**KK**GGSYDHYFDIDHSLT**IK**KLQDISAKTSDGSKVTS**KK**PP  
 APGPDPTPDTPNP**KK**PTPDTPNPNPNINP**KK**DKSISYTKDSTKTDQ**Q****KK**TVTIHAKPVITAD  
 KSI**KK**YFDIDHSLTITNDSA**IK**KEVNAALPNTGDASQAT**KK**TVLINSSIKSSALNFD**KK**HAETDD  
 GTPVTSDFNT**KK**AKIPSLTNLSIAGDNL**KK**TSSASTYFTLNETKID**KK**YGEQTTVTEEQFLKDV  
**KK**TKEIRVTIPFP**Q**TI**KK**TFTITYGDTNPV**KL**TF**KK**TGTAEGLPPGKYTATE**KK**QKTINITF  
 SDNKIMVP**KK**DVKSTDGTT**LL**KKVTTN**KK**DKTITANKVNVKGDV**IK**KDGKVEDTAEDVRNY  
**E****KK**KRIAKGADRFDNLKV**KK**QKVIDTKVEDKANYTE**KK**SEAIEKDLKGEKESTG**KK**DKIT  
 TSKDGKAARLE**KK**TIEINDKLKDAYGN**IK**KKSTVKMDAEPMT**LL**KKDT**KK**GFEADYSEVKNSI  
 KAK**KK**SEKNNDKLFKVT**VE****KK**VEKGNAAGDKDWA**VN****IK**KLGAKNYIVGGKGV**TK**DKD  
 FAGRTIEKEDLSKDK**KK**MKTSTTNKVENYKFKD**KK**IRVIYSKDVSEK**V**AKE**KK**PSGTNKLYT  
 PDGKDGE**KK**LNLNQIKVVFDGK**VE****KK**VERIGGNSRYETNAE**IK**KKTHIYTLTDEGTER**LQ****KK**  
 ILLTDASDKPSADLTA**KK**DGKVRDLPSDTKITTS**KK**KVTVTEADTKVDQASK

Supplementary Table 1  
 Predicted B cell epitopes from LMOF2365\_0578 (ABCPred threshold = 0.8)

Epitope sequence	Start position	ABCPre d score	VaxiJen score	VaxiJen prediction	AllerTOP prediction
DSLSMEIEPNDALVND	646	0.94	1.11	Antigen	Nonallergen
EGTLRDDFIYKVRNT	484	0.91	0.81	Antigen	Allergen
TYSWDEDIPFNSSNNL	341	0.9	0.37	Non Antigen	Nonallergen
IEKKQSAVTPKYDST	625	0.89	0.58	Antigen	Nonallergen
ADAEYTYLVGDKVNAT	456	0.89	0.49	Antigen	Allergen
TYELGTPLTEQQFLND	383	0.89	-0.19	Non Antigen	Nonallergen
PELQDDKSSTLKNVNT	36	0.89	1.08	Antigen	Nonallergen
GNTEFTTSLIKYKPP	437	0.88	0.71	Antigen	Nonallergen
VVTLSSPKSGYYKQDA	504	0.87	0.31	Non Antigen	Nonallergen
TSKITYSAEVMKRPK	313	0.87	0.71	Antigen	Nonallergen
PSLEEISVERNNISDV	211	0.87	0.4	Antigen	Nonallergen
TITGAMTPFTNSGGDI	584	0.86	-0.08	Non Antigen	Allergen
ELEVIDRRTVRQGW <b>TI</b>	570	0.86	0.87	Antigen	Nonallergen
KPVKITATNPKATIDP	296	0.86	1.12	Antigen	Nonallergen
NGYPQLYRLNINNGNI	158	0.86	0.54	Antigen	Allergen
EGIQYLPNLYNVQLQF	128	0.86	0.67	Antigen	Nonallergen
ANSTIADLFPDEGMAK	64	0.85	0.41	Antigen	Nonallergen
LLELQIPDEFMRMDIDA	531	0.85	1.22	Antigen	Allergen
HITDLSSLTNNKMPNL	246	0.85	0.16	Non Antigen	Allergen
NDSYESKITWTLEDAP	660	0.84	0.9	Antigen	Allergen
KYQSKSPSSSPIYLNT	605	0.84	0.6	Antigen	Nonallergen
VGDKVNATQFRADVNA	464	0.84	1.11	Antigen	Nonallergen
YVRSQSFYVETPVVTS	265	0.84	0.56	Antigen	Allergen
DQPTTITNFDKVLKD	404	0.83	0.06	Non Antigen	Nonallergen
YVEPPQVVSYSNGLTY	369	0.83	0.53	Antigen	Allergen
SVFGITKPVKITATNP	290	0.83	0.36	Non Antigen	Nonallergen
TVANQLGRTE>NNNFQ <b>T</b>	80	0.82	0.89	Antigen	Nonallergen
GSSGKAELEVIDRRTV	564	0.82	1.18	Antigen	Nonallergen
EVMKRPKYSSNRVSG	321	0.82	1.05	Antigen	Nonallergen
KNLKASPLTIPANSTI	53	0.81	0.52	Antigen	Nonallergen
RVSGVSGVTYSWDEDI	333	0.81	0.89	Antigen	Allergen
VETGKELPKSPELQDD	26	0.81	0.33	Non Antigen	Nonallergen
PDEGMAKTVANQLGRT	73	0.8	0.73	Antigen	Allergen

**Supplementary Table 2**  
**Predicted B cell epitopes from LMOF2365\_0581 (ABCPred threshold = 0.8)**

Epitope sequence	Start position	ABCPred score	VaxiJen score	VaxiJen prediction	AllerTOP prediction
TAAIQMPAENYDGIL	128	0.9	0.2	Non Antigen	Nonallergen
HESKRTDLSMAPNSNF	237	0.89	0.77	Antigen	Allergen
GLHFQEKKEEDKEKDS	145	0.89	1.71	Antigen	Nonallergen
MGYVDYSIPNTKPKDT	86	0.86	0.69	Antigen	Allergen
KKEITVNCVANTA VTN	69	0.86	0.71	Antigen	Nonallergen
NVEITNQSKEEITVNC	61	0.85	1.21	Antigen	Nonallergen
ILLIATYFVGKRAAK	321	0.85	0.52	Antigen	Nonallergen
KEEDKEKDSSSENDVQI	152	0.85	2.01	Antigen	Nonallergen
YKKGSEKVLHESKRTD	228	0.84	0.93	Antigen	Nonallergen
EFTISSDQAKKANKVA	286	0.83	1.38	Antigen	Allergen
KVHLVAKNKEDKWEWT	269	0.83	0.71	Antigen	Nonallergen
DGIILGGLHFQEKKEE	139	0.83	0.39	Non Antigen	Nonallergen
LGLEKDYTWMYIVGGV	302	0.82	0.05	Non Antigen	Nonallergen
TRNYRNVVEMNLQNTK	199	0.81	0.78	Antigen	Allergen
LELNQIKPATRNYRNV	190	0.81	0.51	Antigen	Nonallergen

**Supplementary Table 3**  
**Predicted B cell epitopes from LMOF2365\_0639 (ABCPred threshold = 0.8)**

Epitope sequence	Start position	ABCPred score	VaxiJen score	VaxiJen prediction	AllerTOP prediction
AGSYQTPPNFNYSVS	320	0.96	0.31	Non Antigen	Nonallergen
GGSYDHYFDIDHSLTI	336	0.94	0.69	Antigen	Nonallergen
TEAQMDTITNVTISNS	65	0.93	0.59	Antigen	Allergen
LQDISAKTSDGSKVTS	449	0.93	1.6	Antigen	Nonallergen
PPAPGPDPTDPTPNP	510	0.92	0.91	Antigen	Nonallergen
TVSGITKSYFDTITKM	295	0.92	-0.24	Non Antigen	Allergen
PTPDPNPNPNINP	517	0.91	0.82	Antigen	Nonallergen
DKSISYTKDSTKTDQQ	432	0.91	1.06	Antigen	Nonallergen
TVTIHAKPVITADKSI	420	0.91	0.66	Antigen	Nonallergen
YFDTITKMEYNALYNN	303	0.9	-0.04	Non Antigen	Nonallergen
VVEGNEPPTPPAPGPD	501	0.88	0.57	Antigen	Allergen
SITNIMPLKSIPNLAT	174	0.88	0.21	Non Antigen	Nonallergen
TLMTERGVNFDGYLFP	248	0.87	0.29	Non Antigen	Allergen
TVTLNAENAAGLKATP	401	0.86	1.18	Antigen	Allergen
YFDIDHSLTITNDSAI	342	0.86	0.64	Antigen	Nonallergen
TVLIGIIIAGVAILFF	568	0.85	0.5	Antigen	Nonallergen
EVNAALPNTGDASQAT	552	0.84	0.94	Antigen	Nonallergen
TITNDSAISYGEQTTV	350	0.84	0.58	Antigen	Allergen
TVLINSSIKSSALNFD	222	0.84	0.55	Antigen	Nonallergen
CVLLMMPFTISFSANV	11	0.84	1.66	Antigen	Nonallergen
HAETDDGTPVTSDFNT	375	0.83	0.87	Antigen	Nonallergen
AKIPSLTNLSIAGDNL	111	0.83	0.46	Antigen	Nonallergen
NPNINPNPDNGQSANS	527	0.82	1.71	Antigen	Allergen
TSSASTYFTLNETKID	268	0.82	0.76	Antigen	Nonallergen
YGEQTTVTEEQFLKDV	359	0.81	0.48	Antigen	Nonallergen
GQSSTSDITEAQMDTI	57	0.8	0.94	Antigen	Allergen
GQSANSENASNPNSSE	537	0.8	1.3	Antigen	Allergen
SKPGVYTVTLNAENAA	395	0.8	0.62	Antigen	Allergen
YFTLNETKIDGSRLTI	274	0.8	1.05	Antigen	Allergen

**Supplementary Table 4**  
**Predicted B cell epitopes from LMOF2365\_2117 (ABCpred threshold = 0.8)**

Epitope sequence	Start position	ABCpred Score	VaxiJen score	VaxiJen prediction	AllerTOP prediction
ATEVTAPLGYQKNPTP	376	0.94	-0.21	Non Antigen	Allergen
EGQTITRLTFNPISTA	173	0.94	0.7	Antigen	Allergen
SEKVQASPTSSNGWQL	22	0.93	1.26	Antigen	Allergen
ITEATAPPGYEKSTKE	468	0.92	0.34	Non Antigen	Nonallergen
ISTASGSLTSGNFLDD	185	0.91	0.31	Non Antigen	Allergen
PVYQDIRTIPGSNLTW	112	0.91	0.24	Non Antigen	Nonallergen
TKEIRVTIPFNPQKTI	481	0.9	0.6	Antigen	Nonallergen
TFTITYGDTNPVKLTF	393	0.9	1.14	Antigen	Nonallergen
TGTAEGLPPGKYTATE	363	0.9	0.52	Antigen	Nonallergen
EAWGTTNPTGNIEVWQ	72	0.89	0.38	Non Antigen	Allergen
DYGVDAAGTTNVWQVNQ	53	0.89	0.89	Antigen	Allergen
QKTINITFSDNKIMVP	493	0.88	0.49	Antigen	Nonallergen
TALPQTGDSSNSSTIF	530	0.87	0.69	Antigen	Nonallergen
TSSNGWQLKWAIKNND	30	0.87	1.77	Antigen	Allergen
PNLKSENFMDAGITT	308	0.86	1.27	Antigen	Allergen
DVVWYDFNGDGIQQDS	214	0.86	1.14	Antigen	Allergen
KIMVPPKPTPTKGSTV	504	0.85	0.69	Antigen	Allergen
SITIFKQDEANKKGLA	416	0.85	0.34	Non Antigen	Allergen
NSDGIGPVYQDIRTIP	106	0.85	0.32	Non Antigen	Allergen
YAIKDNSQSEVAKITT	343	0.84	0.64	Antigen	Allergen
VLPGDYQVKFTLPNND	264	0.84	0.27	Non Antigen	Nonallergen
DGETWGSFEGNYIVPE	158	0.84	-0.22	Non Antigen	Nonallergen
DGIQQDSEEPAPFVKV	223	0.83	0.36	Non Antigen	Nonallergen
TPTKGSTVVKVSGETT	512	0.82	1.22	Antigen	Allergen
GVPKESATTNNIGSYL	245	0.82	0.04	Non Antigen	Allergen
VDLLTKDGVFKESATT	238	0.82	0.12	Non Antigen	Nonallergen
NLTWKFSHRGRMGVDT	124	0.82	0.84	Antigen	Allergen
IEVWQNGNGYNVPAFS	83	0.81	0.79	Antigen	Allergen
DVKSTDGTTLKKVTTN	436	0.81	0.95	Antigen	Nonallergen
YQKNPTPKTFTITYGD	385	0.8	0.26	Non Antigen	Allergen

**Supplementary Table 5**  
**Predicted B cell epitopes from CTC00462 (ABCpred threshold = 0.8)**

Epitope sequence	Start position	ABCpred Score	VaxiJen score	VaxiJen prediction	AllerTOP prediction
AATKGYPVIFGNKNNV	180	0.96	0.24	Nonantigen	Nonallergen
DKTITANKVNVKGDVI	935	0.93	0.87	Antigen	Nonallergen
GIKMDADETHDIKASS	874	0.92	1.45	Antigen	Allergen
DGKVEDTAEDVRNYE	361	0.9	1.46	Antigen	Nonallergen
IAEGKITVKFNTKIDK	1033	0.9	1.21	Antigen	Allergen
VMDDYTATIKGSEVTP	756	0.89	0.78	Antigen	Allergen
KRIAKGADRFDTNLKV	229	0.89	0.64	Antigen	Nonallergen
QKVIDTKVEDKANYTE	1167	0.89	1.18	Antigen	Nonallergen
EYLADASGWIIEKTI	550	0.88	-0.55	Nonantigen	Allergen
SEAIEKDLKGEKESTG	326	0.88	1.75	Antigen	Nonallergen
TVSIELKYTGSLDKFG	909	0.87	1.01	Antigen	Allergen
THDIKASSISTVKMDA	882	0.87	1.19	Antigen	Allergen
DTKITTSKDGKAARLE	822	0.87	2.07	Antigen	Nonallergen
AKERNNYTLRDSKGNK	700	0.87	1.79	Antigen	Allergen
TIEINDKLKDAYGNKI	642	0.87	1.25	Antigen	Nonallergen

KIYINFDRAMDKKSLK	588	0.87	1	Antigen	Allergen
DGDAIADLQDDERTVI	392	0.87	-0.11	Nonantigen	Allergen
EVEKIAGDTYDVISSE	720	0.86	0.58	Antigen	Allergen
GESKKGEKDGDAIADL	384	0.86	1.73	Antigen	Allergen
EKTIDFKIGEVTGKPK	562	0.85	1.54	Antigen	Allergen
SGWIIKEKTIDFKIGE	556	0.85	0.72	Antigen	Allergen
KGEKESTGKADVQSIE	334	0.85	1.65	Antigen	Allergen
LAVGGEGVLPDAVIK	210	0.85	0.7	Antigen	Allergen
STVKMDAEPMTLKKDT	891	0.84	0.74	Antigen	Nonallergen
VTPIVDEVVNMGKADK	769	0.84	-0.47	Nonantigen	Allergen
VKTVKGGTDGKIYINF	578	0.84	1.62	Antigen	Allergen
GFEADYSEVKNSIKAK	502	0.84	1.12	Antigen	Nonallergen
KELAKNYKVERIGGNS	125	0.84	0.61	Antigen	Allergen
SEKNNDKLFKVTVTE	1132	0.84	1.48	Antigen	Nonallergen
SQRIESKEDLDGNSKK	1111	0.84	1.8	Antigen	Allergen
VEKGNAAGDKDWAVNI	1018	0.84	0.77	Antigen	Nonallergen
KFGEEKDYQFKAGDKT	922	0.83	1.25	Antigen	Allergen
IRLNEDKVSADKIELK	607	0.83	1.76	Antigen	Allergen
LGAKNIYIVGGKGVVT	105	0.82	0.7	Antigen	Nonallergen
DFAGRTIEKEDLSKDK	979	0.81	0.68	Antigen	Nonallergen
IGKKGEKVANKSASNI	844	0.81	1.16	Antigen	Allergen
MKTSTTNKVENYKFKD	796	0.81	1.18	Antigen	Nonallergen
KKERIAILFGQEMKTS	784	0.81	0.24	Nonantigen	Nonallergen
DVIDLAGNVMDDYTAT	748	0.81	0.27	Nonantigen	Allergen
IRVIYSKDVSEKVAKE	687	0.81	0.89	Antigen	Nonallergen
PSGTNKLYTPDGKDE	535	0.81	0.9	Antigen	Nonallergen
YYLEEKLPSGTNKLYT	528	0.81	0.16	Nonantigen	Nonallergen
LNLNQIKVVFDDGKVD	351	0.81	1.08	Antigen	Nonallergen
VERIGGNSRYETNAEI	133	0.81	0.93	Antigen	Nonallergen
TIYTLTDEGTERLQK	1077	0.81	0.48	Antigen	Nonallergen
ILLTDASDKPSADLTA	85	0.8	0.65	Antigen	Nonallergen
DGKVRDLPSDTKITTS	813	0.8	1.09	Antigen	Nonallergen
GVENVRTGSNTIEIND	632	0.8	1.35	Antigen	Allergen
KVTVTEADTKVDQASK	1142	0.8	1.33	Antigen	Nonallergen

Figure 4 represents the graph of the visual output from DeepTMHMM, a predictive model for protein structures. It showcases the probabilities of protein segments being located within the cell membrane (transmembrane), or on the inside or outside of the cell. A pronounced peak signifies a high probability transmembrane region, suggesting this part of the protein integrates into the lipid bilayer. Conversely, the absence of a peak in the inside area indicates a low probability of intracellular localization, while a consistent line at the top of the outside area denotes a high probability extracellular domain. This predictive model is crucial for understanding protein orientation and interactions within cellular membranes, informing research and therapeutic strategies.

Figure 5 represents the graph of the pronounced peak suggesting a segment is likely embedded in the membrane, essential for functions like signalling or transport. The absence of peaks in the inside region implies low intracellular presence, while a consistent line in the outside region denotes a probable extracellular stretch. Figure 6

represents the graph displaying a high-probability peak in the transmembrane region, indicating a segment likely spanning the membrane, integral for cellular functions such as signalling. The flat line in the outside region suggests a continuous extracellular domain while the absence of a peak in the inside region implies a low probability of intracellular localization.

Figure 7 represents the peak in the transmembrane section indicating a high probability of a helix spanning the membrane, essential for cellular functions like signalling. The flat line in the outside section suggests a continuous extracellular domain, while the absence of a peak in the inside section implies low intracellular localization.

Figure 8 represents the pronounced peak in the transmembrane region suggesting a high probability of a helical segment spanning the lipid bilayer, essential for cellular functions such as signalling or substance transport. The consistent line at the top of the outside region indicates a segment likely exposed to the extracellular environment

while the absence of a peak in the inside region implies a low probability of intracellular localization.

## Conclusion

Our research marks a significant stride in vaccine development against *Listeria monocytogenes* and *Clostridium tetani*. Through rigorous computational analysis, we have identified promising vaccine candidates and unravelled the intricacies of protein-peptide interactions crucial for vaccine efficacy. Leveraging State-of-the-Art tools and predictive models, we have pinpointed antigenic epitopes located in the outer membrane regions, ensuring accessibility to the immune system. Molecular dynamics simulations have affirmed the stability and structural integrity of the protein-peptide complexes, with RMSD and RMSF analyses confirming a robust interaction conducive to vaccine potency. Visual inspection using Pymol and the Protein-Ligand Interaction Profiler (PLIP) tools has underscored the significant hydrophobic interactions contributing to complex stability.

Our findings offer valuable insights into the rational design of multi-epitope vaccines, poised to provide broad protection against these formidable pathogens. The novelty of our approach lies in the comprehensive *in silico* methodology employed, which has not only identified potential vaccine targets but has also elucidated the underlying mechanisms governing protein-peptide interactions. This research represents a pivotal step towards advancing immunisation strategies, offering a promising avenue for combating infectious diseases on a global scale.

This sets the stage for further experimental validation and clinical trials, with the ultimate goal of translating these computational predictions into tangible solutions for public health challenges. By bridging the gap between computational biology and vaccine development, our work holds immense potential to revolutionise the field and contributes to the ongoing efforts to safeguard human health. As we continue to refine and validate our findings, we remain committed to advance the frontier of vaccine research.

## References

1. Alvarez-Dominguez C. et al, Characterization of a *Listeria monocytogenes* protein interfering with Rab5a, *Traffic*, **9(3)**, 325-337, <https://doi.org/10.1111/j.1600-0854.2007.00683.x> (2008)
2. Alvarez-Dominguez C. et al, Epitopes for multivalent vaccines against *Listeria*, *Mycobacterium* and *Streptococcus* spp: A novel role for glyceraldehyde-3-phosphate dehydrogenase, *Frontiers in Cellular and Infection Microbiology*, **10**, 573348 (2020)
3. Boradia V.M. et al, *Mycobacterium tuberculosis* acquires iron by cell-surface sequestration and internalization of human holo-transferrin, *Nature Communications*, **5**, 4730, <https://doi.org/10.1038/ncomms5730> (2014)
4. Farah H. and Steve S., Biotechnology and the transformation of vaccine innovation: The case of the hepatitis B vaccines 1968-

2000, *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, **64**, 11-21, <https://doi.org/10.1016/j.shpsc.2017.05.004> (2017)

5. Galassie A.C. and Link A.J., Proteomic contributions to our understanding of vaccine and immune responses, *Proteomics Clin Appl*, **9**, 972-989, <https://doi.org/10.1002/prca.201500054> (2015)

6. Hallgren J. et al, Deep TMHMM predicts alpha and beta transmembrane proteins using deep neural networks, *bioRxiv*, <https://doi.org/10.1101/2022.04.08.487609> (2022)

7. Ireton K., Rigano L.A. and Dowd G.C., Role of host GTPases in infection by *Listeria monocytogenes*, *Cellular Microbiology*, **16(9)**, 1311-1320, <https://doi.org/10.1111/cmi.12324> (2014)

8. Khan M.A., Amin A., Farid A., Ullah A., Waris A., Shinwari K., Hussain Y., Alsharif K.F., Alzahrani K.J. and Khan H., Recent advances in genomics-based approaches for the development of intracellular bacterial pathogen vaccines, *Pharmaceutics*, **15**, 152, <https://doi.org/10.3390/pharmaceutics15010152> (2023)

9. Los F.C.O., Randis T.M., Aroian R.V. and Ratner A.J., Role of pore-forming toxins in bacterial infectious diseases, *Microbiology and Molecular Biology Reviews*, **77(2)**, 173-207, <https://doi.org/10.1128/MMBR.00052-12> (2013)

10. Malhotra H. et al, *Mycobacterium tuberculosis* glyceraldehyde-3-phosphate dehydrogenase (GAPDH) functions as a receptor for human lactoferrin, *Frontiers in Cellular and Infection Microbiology*, **7**, 245, <https://doi.org/10.3389/fcimb.2017.00245> (2017)

11. Miyasaka M., Is BCG vaccination causally related to reduced COVID-19 mortality?, *EMBO Mol Med*, **12(6)**, e12661, <https://doi.org/10.15252/emmm.202012661> (2020)

12. Myllymäki H. et al, Identification of novel antigen candidates for a tuberculosis vaccine in the adult zebrafish (*Danio rerio*), *PLoS One*, **12(7)**, e0181942, <https://doi.org/10.1371/journal.pone.0181942> (2017)

13. Nguyen B.N., Peterson B.N. and Portnoy D.A., Listeriolysin O: A phagosome-specific cytolysin revisited, *Cell Microbiol.*, **21(3)**, e12988, <https://doi.org/10.1111/cmi.12988> (2019)

14. Oli A.N., Obialor W.O., Ifeanyichukwu M.O., Odimegwu D.C., Okoyeh J.N., Emechebe G.O., Adejumo S.A. and Ibeanu G.C., Immunoinformatics and vaccine development: An overview. *Immunotargets Ther*, **9**, 13-30, <https://doi.org/10.2147/ITT.S241064> (2020)

15. Pollard A.J. and Bijker E.M., A guide to vaccinology: from basic principles to new developments, *Nat Rev Immunol*, **21**, 83-100, <https://doi.org/10.1038/s41577-020-00479-7> (2021)

16. Qazi O., Brailsford A., Wright A., Faraar J., Campbell J. and Fairweather N., Identification and characterization of the surface-layer protein of *Clostridium tetani*, *FEMS Microbiology Letters*, **274(1)**, 126-131, <https://doi.org/10.1111/j.1574-6968.2007.00834.x> (2007)

17. Qiu C., Xiao C., Wang Z., Zhu G., Mao L., Chen X., Gao L., Deng J., Su J., Su H., Fang E.F., Zhang Z.J., Zhang J., Xie C., Yuan

- J., Luo O.J., Huang L., Wang P. and Chen G., CD8+ T-Cell Epitope Variations Suggest a Potential Antigen HLA-A2 Binding Deficiency for Spike Protein of SARS-CoV-2, *Frontiers in Immunology*, **12**, <https://doi.org/10.3389/fimmu.2021.764949> (2022)
18. Smith J. et al, Evidence for pore formation in host cell membranes by ESX-1-secreted ESAT-6 and its role in *Mycobacterium marinum* escape from the vacuole, *Infection and Immunity*, **76(12)**, 5478-5487, <https://doi.org/10.1128/IAI.00614-08> (2008)
19. Szeto C., Nguyen A.T., Lobos C.A., Chatzileontiadou D.S.M., Jayasinghe D., Grant E.J., Riboldi-Tunncliffe A., Smith C. and Gras S., Molecular Basis of a Dominant SARS-CoV-2 Spike-Derived Epitope Presented by HLA-A\*02:01 Recognised by a Public TCR, *Cells*, **10**, 2646, <https://doi.org/10.3390/cells10102646> (2021)
20. Takagi A. and Matsui M., Identification of HLA-A\*02:01-restricted candidate epitopes derived from the non-structural polyprotein 1a of SARS-CoV-2 that may be natural targets of CD8+ T cell recognition *in vivo*, *Journal of Virology*, **95(5)**, JVI.01837-20, <https://doi.org/10.1128/jvi.01837-20> (2021)
21. Terrasse R. et al, Human and pneumococcal cell surface glyceraldehyde-3-phosphate dehydrogenase (GAPDH) proteins are both ligands of human C1q protein, *Journal of Biological Chemistry*, **287(51)**, 42620-42633, <https://doi.org/10.1074/jbc.M112.423731> (2012)
22. Zhang C.X., Brooks B.W., Huang H., Pagotto F. and Lin M., Identification of surface protein biomarkers of *Listeria monocytogenes* via bioinformatics and antibody-based protein detection tools, *Applied and Environmental Microbiology*, **82(17)**, 5465-5476, <https://doi.org/10.1128/AEM.00774-16> (2016).

(Received 23<sup>rd</sup> October 2024, accepted 04<sup>th</sup> December 2024)