Epitope - based Peptide Vaccine Against Glycoprotein GPC Precursor of Lujo Virus using Immunoinformatics Approaches

Arwa A. Mohammed, Mayada E. Elkhalifa, Khadija E. Elamin, Rawan A. Mohammed, Musab E. Ibrahim, Amina I. Dirar, Sara H. Migdar, Maha A. H. Musa, Emeirii H. Elawad, Salam O. Abdelsalam & Mohamed A. Hassan

Africa City of Technology/Applied Bioinformatics Center Biotechnology Park

ABSTRACT

Background: Lujo virus (LUJV) is a highly fatal human pathogen belonging to the Arenaviridae family. Lujo virus causes viral hemorrhagic fever (VHF). An In silico molecular docking was performed on the GPC domain of Lujo virus in complex with the first CUB domain of neuropilin-2.

The aim of this study is to predict an effective epitope-based vaccine against the glycoprotein GPC precursor of Lujo virus using immunoinformatics approaches.

Methods and Materials: A glycoprotein GPC precursor of Lujo virus Sequence was retrieved from NCBI. Different prediction tools were then used to analyze the nominee's epitopes in BepiPred-2.0: Sequential B-Cell Epitope Predictor for B-cell, T-cell MHC class II & I. Later the proposed peptides were docked using the Autodock 4.0 software program.

Keywords: immunoinformatics, glycoprotein GPC precursor, epitope-based vaccine, Lujo virus LUJV, VHF.

Classification: H.3.3

Language: English

Great Britain Journals Press

LJP Copyright ID: 975823
Print ISSN: 2514-863X
Online ISSN: 2514-8648

London Journal of Research in Computer Science and Technology

Volume 23 | Issue 2 | Compilation 1.0

Epitope - based Peptide Vaccine Against Glycoprotein GPC Precursor of *Lujo Virus* using Immunoinformatics Approaches

Arwa A. Mohammed[α], Mayada E. Elkhalifa[σ], Khadija E. Elamin[ρ], Rawan A. Mohammed[ↀ], Musab E. Ibrahim[¥], Amina I. Dirar[§], Sara H. Migdar[χ], Maha A. H. Musa[ν], Emeirii H. Elawad[Θ], Salam O. Abdelsalam[ζ] & Mohamed A. Hassan[£]

ABSTRACT

Background: Lujo virus (LUJV) is a highly fatal human pathogen belonging to the Arenaviridae family. Lujo virus causes viral hemorrhagic fever (VHF). An In silico molecular docking was performed on the GPC domain of Lujo virus in complex with the first CUB domain of neuropilin-2.

The aim of this study is to predict an effective epitope-based vaccine against the glycoprotein GPC precursor of Lujo virus using immunoinformatics approaches.

Methods and Materials: a glycoprotein GPC precursor of Lujo virus Sequence was retrieved from NCBI. Different prediction tools were then used to analyze the nominee's epitopes in BepiPred-2.0: Sequential B-Cell Epitope Predictor for B-cell, T-cell MHC class II & I. Later the proposed peptides were docked using the Autodock 4.0 software program.

Results and Conclusions: The proposed and promising peptides FWYLNHTKL and YMFSVTLCI has shown a very strong binding affinity to MHC class I & II alleles with high population coverage for the world, South Africa, and Sudan. This indicates a strong potential to formulate a new vaccine, especially with the peptide YMFSVTLCI which is likely to be the first proposed epitope-based vaccine against glycoprotein GPC of Lujo virus. This study recommends an in-vivo assessment for the most promising peptides especially FWYLNHTKL, YMFSVTLCI and LPCPKPHRLR.

Keywords: immunoinformatics, glycoprotein GPC precursor, epitope-based vaccine, *Lujo virus* LUJV, VHF.

Author α σ ρ ¥ § χ Θ ζ £: Department of Biotechnology, Africa city of Technology, Khartoum, Sudan.

α: Department of Pharmacy, Sudan Medical Council, Khartoum, Sudan.

σ: Department of Pharmaceutics, Faculty of Pharmacy, Alneelain University, Khartoum, Sudan.

ρ: Department of Botany, Faculty of Science, University of Khartoum, Sudan.

ↀ: Department of Bioinformatics, Faculty of Information & Science Technology, Multimedia University, Malaysia.

¥ χ: Department of Clinical and industrial pharmacy, Faculty of Pharmacy, National University, Khartoum, Sudan.

§: Department of Pharmaceutical Chemistry, Faculty of Pharmacy, University of Khartoum, Sudan.

ν: Department of Clinical Pharmacy, School of Pharmacy, Ahfad University for Women, Khartoum, Sudan.

Θ: Department of Pharmacology, Faculty of Medicine, University of Khartoum, Sudan.

ζ: Department of Clinical Immunology, Sudan Medical Specialization Board, Khartoum, Sudan.

£: Department of Bioinformatics, DETAGEN Genetics Diagnostic Center, Kayseri, Turkey.

I. INTRODUCTION

Arenaviruses are rodent-borne viruses. Where a genetically unique arenavirus called *Lujo virus*, has been discovered as the causal agent of a nosocomial outbreak of acute febrile disease with hemorrhagic manifestations in Zambia and South Africa. The outbreak marked a high case fatality rate of almost 80% [1]. These viruses are genetically and geographically related to the Old World mammarena viruses, endemic to West Africa, and the New World mammarena viruses, endemic to South and North America [2].

Lujo virus causes viral hemorrhagic fever (VHF) which can be caused by five distinct families of viruses: the filo-, arena-, flavi-, rhabdo- and bunya virus family [3].

Viral hemorrhagic fever (VHF) is an acute systemic illness classically involving fever, a constellation of initially nonspecific signs and symptoms, and a propensity for bleeding and shock.

With *Lujo virus* hemorrhagic fever (LVHF) illness typically begins with the abrupt onset of fever, malaise, headache, and myalgias followed successively by a sore throat, chest pain, gastrointestinal symptoms, rash, minor hemorrhage, subconjunctival injection, and neck and facial swelling over the first week of sickness [4]. No major hemorrhage was noted. Whereas neurological signs were sometimes seen in the late stages, shock and multi-organ system failure, often with evidence of disseminated intravascular coagulopathy, ensued in the second week, with death in four of the five cases [4].

There are currently limited preventative and therapeutic options for patients infected with these highly pathogenic viruses [5].

Arenaviruses are enveloped negative-strand RNA viruses with a genome that is bi- segmented into S and L segments. The S segment encodes a nucleocapsid protein (NP) and an envelope glycoprotein precursor (GPC); the L segment encodes a matrix protein (Z) and an RNA-dependent RNA polymerase (L). The GPC is synthesized as a single polypeptide and undergoes processing by the host cell signal peptidase (SP ase) and subtilisin-like kexinisozyme-1/site-1-protease (SKI-1/S1P), yielding typical receptor binding (G1), transmembrane fusion (G2), and stable signal peptide (SSP) subunits, respectively [6-8]. Viral entry into target cells is initiated by the binding of G1 to appropriate cell surface receptors. The first cellular receptor for arenavirus to be identified was_- dystroglycan (_DG), a ubiquitous receptor for extracellular matrix proteins [9].

The understanding of epitope/antibody interaction is the key to constructing potent vaccines and effective diagnostics. The host defense mechanisms against viruses generally vary from germline-encoded immunity, which present early in the evolution of microorganisms to activation and induction of specific adaptive immune responses by the production of Th-1 andTh-2 cytokines. B-cells recognize antigens via membrane bound antibodies using B-cell receptors (BCRs), resulting in the secretion of antibodies that bind to the antigen and deactivate or remove it. Processing and presentation of peptide epitopes are essential steps in cell-mediated immunity [10]. *Lujo virus* (LUJV) is a highly fatal human pathogen belonging to the *Arenaviridae* family. This virus is unique; as it uses neuropilin-2 (NRP2) as a cellular receptor.

Previous study revealed that the GP1 receptor-binding domain of LUJV (LUJVGP1) recognizes NRP2, where its recognition is metal-ion dependent. The binding of a $Ca2^{+}$ ion stabilizes the conformations of Asp127 and Glu79 from NRP2 pre-organizing them for interaction with Lys110 of LUJVGP1. CUB domain of NRP2 is almost completely conserved among humans, mice, rats and bats, and the only slight variations occur outside of the binding site for LUJV. Hence all of these animal species have a potential to serve as reservoirs for LUJV, considering only the compatibility to NRP2 [2]. *In silico* molecular docking was performed on the GP1 domain of *Lujo virus* in complex with the first CUB domain of neuropilin-2 [2].

The aim of the study is to predict an effective epitope-based vaccine against an envelope glycoprotein precursor (GPC); of *Lujo virus*. The development of immunogenetics approaches will enhance the understanding of the genetic factors impact on the interindividual and interpopulation variations in immune responses to vaccines that could be helpful to progress new vaccine strategies [11]. In silico/reverse vaccinology had replaced conventional culture-based vaccine because it reduces the cost required for laboratory investigation of pathogen, also speeding up the time needed to achieve the results [12,13].

Therefore, using immunoinformatics approaches to predict this new kind of vaccines could be a

magnificently additive in the way forward of preventing *Lujo virus*. Normally, the investigation of the binding affinity of antigenic peptides to the MHC molecules is the main goal when predicting epitopes. The usage of such tools and information leads to the development of new vaccines. While these approaches permit the optimization of a vaccine for a specific population, It's probably can be reformulated to design a "universal vaccine" a vaccine that provides maximum coverage for the whole worlds' population [14-17]. In this study, we focused on both MHC class II and class I with performing of molecular docked in HLA-A0201.

II. MATERIALS AND METHODS

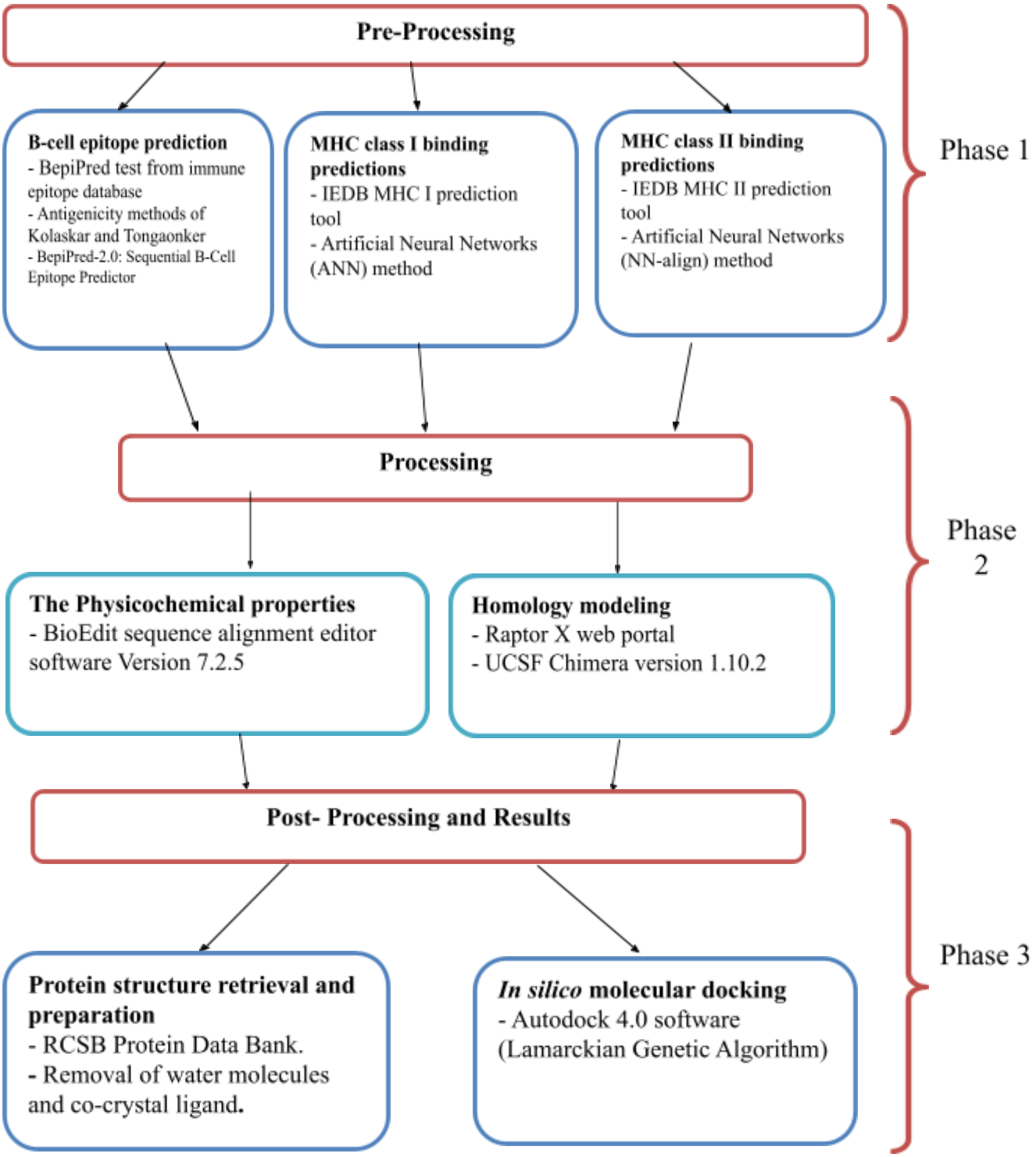

Figure 1: The Three Phases of Material and Method Process

2.1 Sequences Retrieval

The amino acids sequences of Glycoprotein GPC (Glycoside hydrolase family) of *Lujo virus* were retrieved from the NCBI database (https://www.ncbi.nlm.nih.gov/protein) [18] in FASTA format on July 2018. Different prediction tools of Immune Epitope Database IEDB analysis resource (http://www.iedb.org/) [19] were then used to analyze the candidate epitopes as shown on figure (1).

2.2 Conservation Region and Physicochemical Properties

Conservation regions were determined using multiple sequence alignment with the help of Clustal-W in the BioEdit software version 7.2.5 [20]. Epitope conservancy prediction for individual epitopes was then calculated using the IEDB analysis resource. Conservancy can be defined as the portion of protein sequences that restrain the epitope measured at or exceeding a specific level of identity [21]. The physicochemical properties of the retrieved sequence; molecular weight and amino acid composition; were also determined by BioEdit software version 7.2.5 [20].

2.3 B Cell Epitope Prediction Tools

Candidate epitopes were analyzed using several B-cell prediction methods to determine their antigenicity, flexibility, hydrophilicity, and surface accessibility. The linear prediction epitopes were obtained from the Immune epitope database (http://tools.iedb.org/bcell/result/) [22] by using the BepiPred test with a threshold value of 0.149 and a window size of 6.0.

Moreover, surface-accessible epitopes were predicated with a threshold value of 1.0 and window size of 6.0 using the Emini surface accessibility prediction tool [23].

Kolaskar and Tongaonker antigenicity methods (http://tools.iedb.org/bcell/result/) were proposed to determine the sites of antigenic epitopes with a default threshold value of 1.030 and a window size of 6.0 [24].

2.4 T cell epitope prediction tools

2.4.1 Peptide binding to MHC class I molecules

The peptide binding was assessed by IEDB MHC class I prediction tool at http://tools.iedb.org/mhc1. This tool employs different methods to determine the ability of submitted sequence to bind to a specific MHC class I molecule. The artificial neural network (ANN) method [25, 26] was used to calculate IC50 values of peptide binding to MHC- class I molecules. For both frequent and non-frequent alleles, peptide length was set to 9 amino acids earlier to the prediction. The alleles having binding affinity IC50 equal to or less than 500 nM were considered for further analysis. The affinity of 500 nM is routinely used as a threshold for peptide selection and it captures 92% of the epitopes [27- 29].

2.4.2 Peptide Binding to MHC Class II Molecules

MHC class II prediction tool http://tools.iedb.org/mhcII provided by Immune Epitope Database (IEDB) analysis resource and human allele references set was used to predict the peptide binding to MHC class II molecules. Where the Artificial Neural Network prediction method was chosen to identify the binding affinity to MHC class II grooves and MHC class II binding core epitopes. All epitopes that bind to as many alleles at score equal to or less than 1000 half-maximal inhibitory concentration (IC50) were selected for further analysis. [30].

2.5 Population Coverage

Population coverage for each epitope was calculated by the IEDB population coverage tool at http://tools.iedb.org/tools/population/iedb_input [31]. This tool was targeted in order to determine the fraction of individual alleles predicted to respond to a given set of epitopes with known MHC restrictions. For every single population coverage, the tool computed the following information: (1) predicted population coverage, (2) HLA combinations recognized by the population, and (3) HLA combinations recognized by 90% of the population (PC90). All epitopes and their MHC class I and MHC class II

molecules were assessed against a population coverage area selected prior to the submission.

2.6 Homology Modeling

The 3D structure of glycoprotein GPC of *Lujo virus* was predicted using Raptor X web portal (http://raptorx.uchicago.edu/) [32]. The reference sequence was submitted in FASTA format on 14/9/2018 and the structure was received on 15/9/2018. Subsequently the structure was treated with UCSF Chimera 1.10.2 [33] to visualize the position of the proposed peptides.

2.7 In Silico Molecular Docking

2.7.1 Ligand Preparation

In order to estimate the binding affinities between the epitopes and the molecular structure of MHC class I & MHC class II, in silico molecular docking was utilized. The proposed epitopes sequences were then selected from the *Lujo virus* reference sequence using Chimera 1.10 and saved as a "pdb" file. The obtained files were later optimized and energy minimized. The HLA-A0201 was selected as the macromolecule for docking; as HLA-A0201 is considered as the most popular MHC allele and most MHC-I epitopes were nonapeptides [34].

Its crystal structure (4UQ3) was downloaded from the RCSB Protein Data Bank (http://www.rcsb.org/pdb/home/home.do), which was in complex with an azobenzene-containing peptide [35].

The crystal structure of LUJVGP1/NRP2 was retrieved from protein databank (PDB ID: 6GH8) [2].

2.7.2 Molecular Docking

Molecular docking was performed using Autodock 4.0 software, based on Lamarckian Genetic Algorithm; which combines energy evaluation through grids of affinity potential to find the suitable binding position for a ligand on a given protein [36]. Polar hydrogen atoms were added to the protein targets and Kollman united atomic charges were computed. All hydrogen atoms were added to the ligands before the Gastiger partial charges were assigned. The co-crystal ligand was removed and the bond orders were checked. The target's grid map was calculated and set to 60×60×60 points with a grid spacing of 0.375 Å. The grid box was then allocated properly in the target to include the active residue in the center. The default docking algorithms were set in accordance with standard docking protocol [37]. Finally, ten independent docking runs were carried out for each ligand and results were retrieved as binding energies. Poses that showed the lowest binding energies were visualized using the UCSF chimera [38].

III. RESULTS

3.1. Lujo Virus Glycoprotein GPC Physical and Chemical Parameters

The physicochemical properties of the *Lujo virus* glycoprotein GPC protein was assessed using BioEdit software version 7.0.9.0. The protein length was found to be 454 amino acids. The amino acids that form *Lujo virus* glycoprotein GPC protein is shown in Figure (2) along with their numbers and molar percentages in (Mol%).

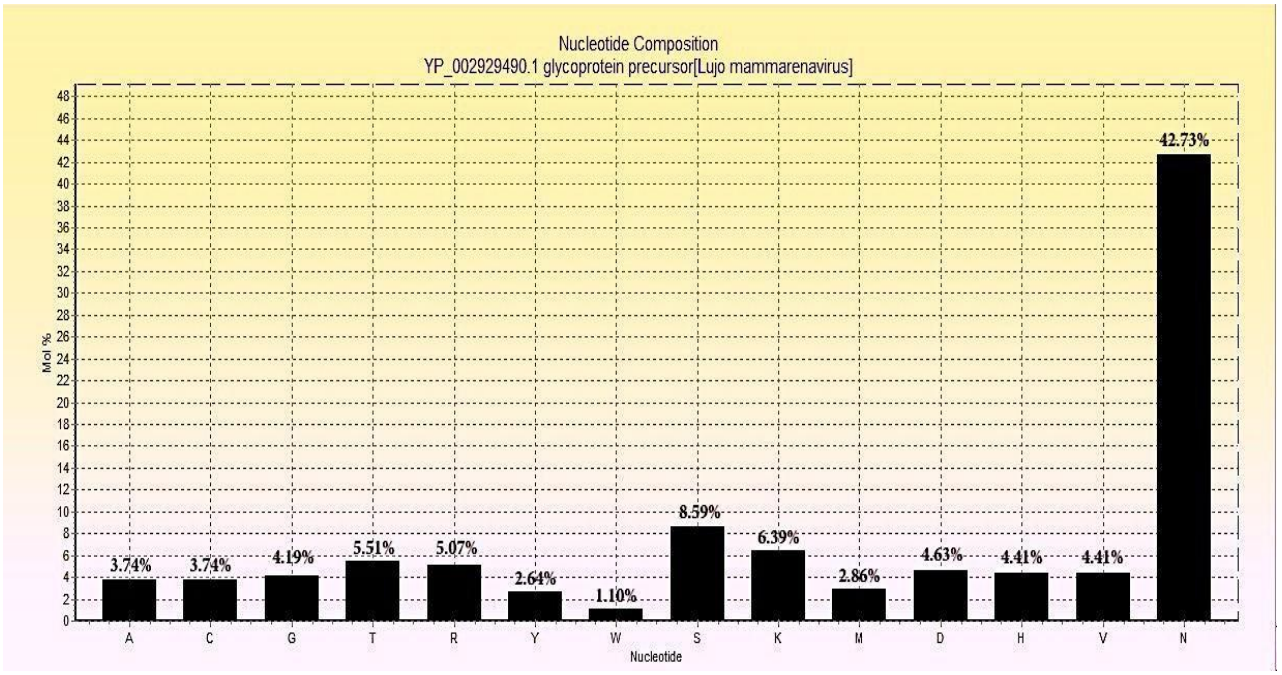

Figure 2: Amino acids composition of *Lujo virus* glycoprotein GPC using Bioedit

3.2 B-Cell Epitope Prediction

The ref sequence of the *Lujo virus* glycoprotein GPC was subjected to a Bepipred linear epitope prediction. Emini surface accessibility, Kolaskar and Tongaonkar antigenicity methods in IEDB were used to determine bindings to the B cell, testing its surface and immunogenicity. The results are shown in Table1.

Table 1: Most Linear Epitopes, High Surface Accessibility and Immunogenicity Bindings to the B Cell

No.	Start	End	Peptide	Length	Emini Surfce	scor	Kolasker & Tongankar	scor
33	423	432	LPCPKP HRLR	10	pass	1.371	pass	1.088
35	423	430	LPCPKP HR	8	pass	1.378	pass	1.095

3.3 Prediction of T Helper Cell Epitopes and Interaction With MHC Class I Alleles

Lujo virus glycoprotein GPC sequence was analyzed using IEDB MHC class I binding prediction tool based on ANN-align with half-maximal inhibitory concentration (IC_{50}) ≤**500**; the least most promising epitopes that had a binding affinity with the Class I alleles along with their positions in the *Lujo virus* glycoprotein GPC are shown in Table 2.

Table 2: Most Potential T-Cell Epitopes With Interacting MHC- Class I Alleles, Their Positions, IC50, Rank and Conservancy

Peptide	Start	End	Allele	IC50	Rank	Conservancy
Ymfsvtlci	403	411	HLA-A*02:01	3.96	0.02	
	403	411	HLA-A*02:06	43.36	0.48	
	403	411	HLA-A*23:01	121.22	0.38	
	403	411	HLA-A*29:02	77.5	0.45	
	403	411	HLA-A*32:01	116.01	0.15	

	403	411	HLA-B*15:01	92.12	0.49	100
	403	411	HLA-B*39:01	56	0.08	
	403	411	HLA-B*58:01	314.96	0.77	
	403	411	HLA-C*03:03	420.85	0.68	
	403	411	HLA-C*07:01	486.45	0.14	
	403	411	HLA-C*12:03	13.35	0.03	
	403	411	HLA-C*14:02	45.86	0.08	
Lmfsvsfym	396	404	HLA-A*02:01	4.29	0.03	100
	396	404	HLA-A*02:06	22.33	0.27	
	396	404	HLA-A*11:01	482.09	2.3	
	396	404	HLA-A*29:02	9.94	0.08	
	396	404	HLA-A*30:02	130.57	0.49	
	396	404	HLA-A*31:01	46.07	0.58	
	396	404	HLA-B*15:01	28.4	0.17	
	396	404	HLA-B*35:01	76.06	0.25	
	396	404	HLA-B*58:01	32.61	0.16	
	396	404	HLA-C*03:03	271.14	0.54	
	396	404	HLA-C*12:03	254.08	0.33	
Rlqeavstl	165	173	HLA-A*02:01	66.37	0.69	100
	165	173	HLA-A*02:06	239.83	1.7	
	165	173	HLA-A*32:01	114.21	0.15	
	165	173	HLA-B*15:01	227.04	0.93	
	165	173	HLA-C*12:03	434.14	0.5	
	165	173	HLA-C*14:02	130.18	0.22	
Fqlviflll	44	52	HLA-A*02:01	160.67	1.5	100
	44	52	HLA-A*02:06	10.35	0.13	
	44	52	HLA-B*27:05	223.21	0.83	
	44	52	HLA-B*39:01	37.81	0.06	
	44	52	HLA-B*40:01	207.49	0.39	
	44	52	HLA-B*48:01	160.64	0.02	
Msllssipm	74	82	HLA-A*02:06	287.92	1.9	100
	74	82	HLA-A*30:01	213.76	0.53	
	74	82	HLA-B*15:01	221.62	0.91	
	74	82	HLA-B*35:01	13.96	0.06	
	74	82	HLA-B*39:01	364.24	0.26	
	74	82	HLA-B*58:01	150.48	0.52	
	74	82	HLA-C*03:03	99.8	0.32	
	74	82	HLA-C*14:02	337.92	0.43	

Vifdlfref	122 122 122 122 122	130 130 130 130 130	HLA-A*02:06 HLA-A*29:02 HLA-A*32:01 HLA-B*15:01 HLA-C*12:03	362.08 430.15 329.32 50.88 424.4	2.2 1.3 0.33 0.3 0.5	100
Itfslltnk	97 97 97 97 97	105 105 105 105 105	HLA-A*03:01 HLA-A*11:01 HLA-A*30:01 HLA-A*31:01 HLA-A*68:01	16.97 9.43 173.26 95.4 14.88	0.06 0.03 0.48 1.2 0.1	100
Ilmfsvsfy	395 395 395 395 395 395 395	403 403 403 403 403 403 403	HLA-A*03:01 HLA-A*11:01 HLA-A*29:02 HLA-A*30:02 HLA-A*68:01 HLA-B*15:01 HLA-B*15:02	25.13 69.52 16.35 23.88 208.18 31.98 289.75	0.1 0.48 0.13 0.04 1.3 0.2 0.11	100
Fwylnhtkl	331 331 331 331 331	339 339 339 339 339	HLA-A*23:01 HLA-C*03:03 HLA-C*07:02 HLA-C*12:03 HLA-C*14:02	278.63 68.25 370.98 413.98 14.88	0.65 0.25 0.09 0.49 0.03	100

3.4 Prediction of T Helper Cell Epitopes and Interaction With MHC Class II Alleles

Lujo virus glycoprotein GPC sequence was analyzed using IEDB MHC class II binding prediction tool based on NN-align with half-maximal inhibitory concentration (IC_{50}) ≤1000. The list of the most promising epitopes, that had a strong binding affinity to MHC class II alleles and depending on the number of their binding alleles are shown in Table 3.

Table 3: Most potential T-cell Epitopes with Interacting MHC- Class II Alleles

Peptide	Allele	No. of Alleles
Fwylnhtkl	HLA-DRB1*01:01, HLA-DRB1*04:04, HLA-DRB1*04:05, HLA-DRB1*07:01, HLA-DRB1*08:02, HLA-DRB1*09:01, HLA-DRB1*15:01, HLA-DRB4*01:01, HLA-DRB5*01:01, HLA-DPA1*01/DPB1*04:01, HLA-DPA1*01:03/DPB1*02:01, HLA-DPA1*02:01/DPB1*01:01, HLA-DPA1*02:01/DPB1*05:01, HLA-DPA1*03:01/DPB1*04:02	14

Fnmsllssi	HLA-DRB1*04:01, HLA-DRB1*04:04, HLA-DRB1*04:05, HLA-DRB1*07:01, HLA-DRB1*09:01, HLA-DRB1*11:01, HLA-DRB3*01:01, HLA-DRB5*01:01, HLA-DPA1*01/DPB1*04:01, HLA-DPA1*01:03/DPB1*02:01, HLA-DPA1*02:01/DPB1*01:01, HLA-DPA1*03:01/DPB1*04:02, HLA-DQA1*05:01/DQB1*03:01	13
Inaiisdtl	HLA-DRB1*01:01, HLA-DRB1*03:01, HLA-DRB1*04:01, HLA-DRB1*04:04, HLA-DRB1*04:05, HLA-DRB1*07:01, HLA-DRB1*09:01, HLA-DRB1*13:02, HLA-DRB1*15:01, HLA-DRB4*01:01, HLA-DQA1*01:02/DQB1*06:02, HLA-DQA1*05:01/DQB1*02:01, HLA-DQA1*05:01/DQB1*03:01	13
Lmklfqwsl	HLA-DRB1*01:01, HLA-DRB1*04:01, HLA-DRB1*04:05, HLA-DRB1*07:01, HLA-DRB1*09:01, HLA-DRB1*15:01, HLA-DRB4*01:01, HLA-DRB5*01:01, HLA-DPA1*01/DPB1*04:01, HLA-DPA1*01:03/DPB1*02:01, HLA-DPA1*02:01/DPB1*01:01, HLA-DPA1*03:01/DPB1*04:02, HLA-DQA1*01:01/DQB1*05:01	13
Vfqaipeil	HLA-DRB1*03:01, HLA-DRB1*07:01, HLA-DRB1*09:01, HLA-DRB1*13:02, HLA-DRB4*01:01, HLA-DRB5*01:01, HLA-DPA1*01/DPB1*04:01, HLA-DPA1*01:03/DPB1*02:01, HLA-DPA1*02:01/DPB1*01:01, HLA-DPA1*03:01/DPB1*04:02, HLA-DQA1*01:01/DQB1*05:01, HLA-DQA1*05:01/DQB1*02:01, HLA-DQA1*05:01/DQB1*03:01	13
Ymfsvtlci	HLA-DRB1*01:01, HLA-DRB1*04:01, HLA-DRB1*07:01, HLA-DRB1*13:02, HLA-DRB1*15:01, HLA-DRB3*01:01, HLA-DRB5*01:01, HLA-DPA1*01/DPB1*04:01, HLA-DPA1*01:03/DPB1*02:01, HLA-DPA1*02:01/DPB1*05:01, HLA-DPA1*03:01/DPB1*04:02, HLA-DQA1*01:02/DQB1*06:02, HLA-DQA1*05:01/DQB1*03:01	13

3.5 Population Coverage

A population coverage test was performed to detect all the epitopes that bind to MHC class I alleles and MHC class II alleles available in the database in relation to the world, South Africa, and Sudan.

Table 4: APopulation Coverage for All Epitopes That Bind to MHC Class I and II Alleles From Different Parts of the World

MHC classes	Population	World	South Africa	Sudan
Class I	Coverage[a]	99.83%	99.4%	99.41%
	Average_hit[b]	32.5	25.45	28.25
	PC90[c]	13.28	8.99	8.76
Class II	Coverage[a]	68.23%	32.1%	56.38%
	Average_hit[b]	55.22	10.89	34.14
	PC90[c]	-6.57	4.71	5.57

[a] projected population coverage
[b] average number of epitope hits / HLA combinations recognized by the population
[c] minimum number of epitope hits / HLA combinations recognized by 90% of the population.

Table 5: Population Coverage of the Proposed Peptides in MHC Class I and MHC Class II in Five Areas

Peptide	Population coverage %/ Area								
	World			South Africa			Sudan		
	MHC I	MHC II	MHC I & II	MHC I	MHC II	MHC I & II	MHC I	MHC II	MHC I & II
Ymfsvtlci	73.92%	56.92%	88.77%	63.56%	5.91%	65.72%	75.17%	21.85%	80.6%
Fwylnhtkl	42.99%	55.84%	74.82%	41.97%	1.79%	43.01%	35.12%	25.35%	51.56%

3.6 3D Structure

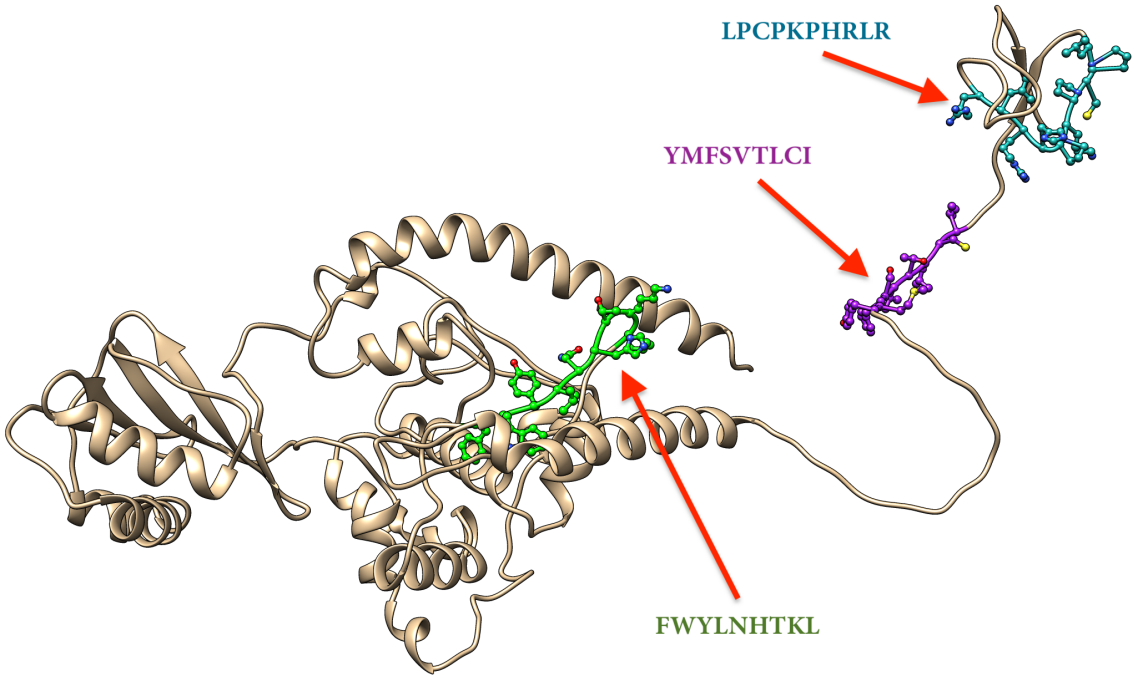

Figure 3: The Four Potential Peptides Bound to MHC Class I, MHC Class II and B Cell Visualized by Chimera 1.10.2

3.7 Molecular Docking

Three peptides; FWYLNHTKL, LPCPKPHRLR and YMFSVTLCI were docked onto protein target of GP1 domain of *Lujo virus* in complex with the first CUB domain of NRP2.

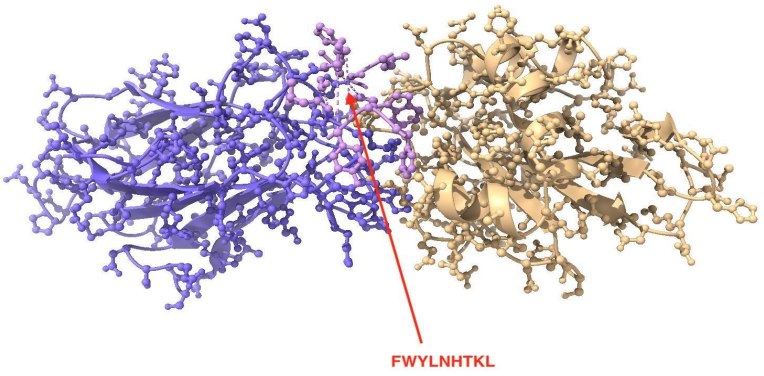

Figure 4: Peptide-1 FWYLNHTKL

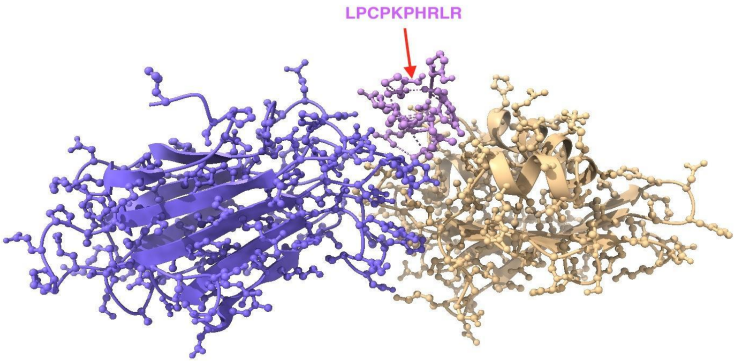

Figure 5: Peptide-2 LPCPKPHRLR

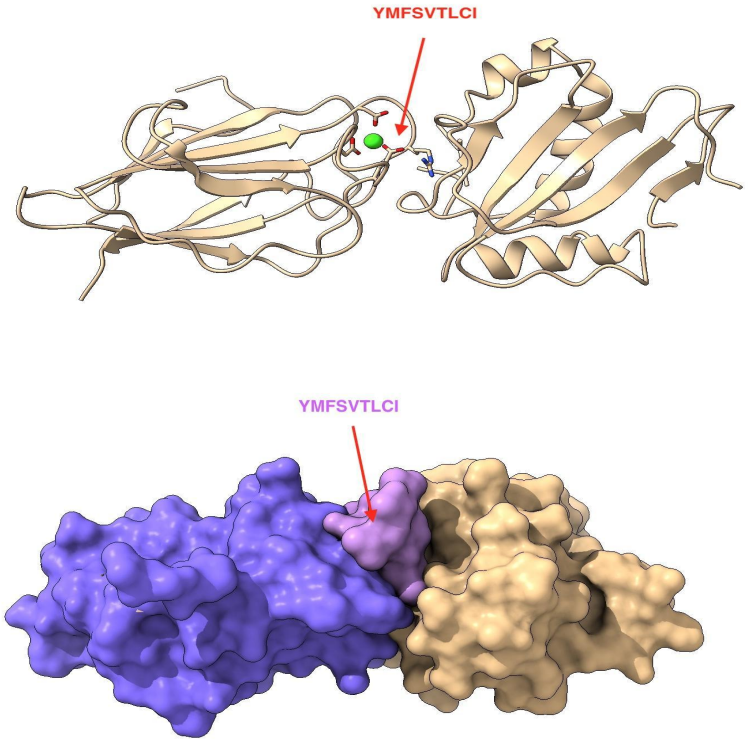

Figure 6: Peptide-3 YMFSVTLCI

IV. DISCUSSION

In this computational immunoinformatic study we suggest a new promising highly selective peptides vaccine against *Lujo virus* for the first time according to our findings. We expect to obtain a peptide-based vaccine which implies a high antigenicity and a minimum allergic outcome that is more accurate than the currently used vaccines. The analytical process started after having adequate information on the protein structure of *Lujo virus* according to the literature review. Simultaneously, though the 3D structure was previously available on the database with all its prospects, we produced our own structure using Raptor X web portal to utilize its complete physiochemical properties information file to confirm our results, and it's a technique we have pursued. The reference sequence of *Lujo virus* glycoprotein GPC was obtained from the NCBI database. To determine the binding affinity of the

conserves epitopes to B-cell and to examine the immunogenicity several tests on the IEDB database were used; the Bepipred linear epitope prediction test, Emini surface accessibility test, and Kolaskar and Tongaonkar antigenicity test were examined. For the Bepipred test of B-cell, the total number of epitopes was 39. For Emini surface accessibility prediction, 29 conserved epitopes were passing the default threshold of 1.0.

In Kolaskar and Tongaonkar antigenicity, 7 epitopes provided a score above the default threshold 1.045. However, there are only two epitopes that passed our three tests, which were (LPCPKPHRLR, LPCPKPHR). The reference glycoprotein GPC strain was analyzed using IEDB class MHC- class I binding prediction tool to predict T cell epitope. 165 peptides were predicted to interact with different class MHC- class I alleles. For class MHC- class II binding prediction, there were 315 epitopes found to interact with class MHC- class II alleles. The peptides YMFSVTLCI, LMFSVSFYM, RLQEAVSTL, FQLVIFLLL, MSLLSSIPM, VIFDLFREF, ITFSLLTNK, ILMFSVSFY and FWYLNHTKL had the affinity to bind with the highest number of class MHC- class I alleles. The peptides FWYLNHTKL, YMFSVTLCI, FNMSLLSSI, INAIISDTL, LMKLFQWSL and VFQAIPEIL had the affinity to bind with the highest number of class MHC- class II alleles.

The most promising three peptides for both class MHC- class I and MHC- class II were FWYLNHTKL, LPCPKPHRLR and YMFSVTLCI as shown on figure (3). On the other hand, the world Population coverage of all epitopes that bind to MHC- class I were found to be 99.83%, while the world population coverage of all epitopes that bind to MHC- class II were 68.23% as presented in table 4. For the binding affinity to MHC- class I and MHC- class II the peptide FWYLNHTKL was found to bind 14 different alleles of MHC- class II & five alleles of MHC- class I, that gave a world population coverage of 74.82% , 43.01% for South Africa and 51.56% for Sudan of both MHC class I and II as shown on table 5. This finding shows a very strong potential to formulate an epitopes-based peptide vaccine for *Lujo virus*. The binding affinity of the peptide YMFSVTLCI to both class MHC- class I and MHC- class II alleles were found to be 13 different alleles with world population coverage 88.77%, 65.72% for South Africa and 80.6% for Sudan of both MHC class I and II as shown on table 5.

According to these interesting findings, a very promising vaccine against *Lujo virus* can potentially be formulated. The most promising three peptides; FWYLNHTKL, LPCPKPHRLR and YMFSVTLCI were docked on to protein target of GPC domain of *Lujo virus* in complex with the first CUB domain of NRP2 as shown on figure (4 – 6). All peptides were docked on the interface of *Lujo virus* GPC/NRP2 and scored binding energies of -5.84, -3.88 and -8.20 Kcal/mol for peptides 1, 2 and 3, respectively. As the docking results from the analysis of peptide-1 showed binding hydrogen bonding with two amino acid residues: SER-51 and GLN-131 of NRP2. While, Peptide-2 showed a bonding affinity to hydrogen with residues HIS-131 and PHE-137 of *Lujo virus* GPC and ARG-432 of NRP2, whereas peptide-3 formed hydrogen bonds with PHE-137 of *Lujo virus* GPC.

Only Peptide-2 and 3 interact by forming hydrogen bonding with residues on *Lujo virus* GPC (HIS-131 and PHE-137). These residues are located at a hydrophobic pocket and adjacent to both residues Val139 and Thr140 of the α2β4 loop which participates in Van der Waals interactions with NRP2 residues. In addition, histidine residues in the *Lujo virus* GPC/NRP2 complex are obvious candidates for controlling pH-dependent protein-protein interactions.

As for peptide-1, it has formed hydrogen bonds with GLN-131 of NRP2, which is adjacent to the key residue Arg130 that is important for NRP2-fc to recognize *Lujo virus* GPC-bearing cells and cell entry of *Lujo virus* [2]. The overall docking results analysis has revealed that the peptides are docked at *Lujo virus* GPC/NRP2 binding surfaces, in which these peptides would serve as potential inhibitors for blocking binding to NRP2 and thus may neutralize the virus.

V. CONCLUSIONS

To the best of our knowledge, this study is considered to be the first to propose an epitope-based peptide vaccine against glycoprotein GPC of *Lujo virus*, which is expected to be highly antigenic with a minimum allergic impact. Furthermore, this study proposes a promising peptide FWYLNHTKL with a very strong binding affinity to MHC1 and MHC11 alleles. This peptide shows exceptional population coverage results for both MHC1 and MHC11 alleles.

In-vivo and in-vitro assessments for the most promising peptides namely, FWYLNHTKL, LPCPKPHRLR and YMFSVTLCI are recommended to be explored and studied on their ability to be developed into vaccines against *Lujo virus* glycoprotein GPC.

Competing Interests

The authors declare that they have no competing interests.

REFERENCES

1. Bergeron, E., et al. (2012). "Reverse genetics recovery of Lujo virus and role of virus RNA secondary structures in efficient virus growth." J Virol86(19): 10759-10765.
2. Cohen-Dvashi, H., et al. (2018). "Structural basis for receptor recognition by Lujo virus." Nat Microbiol 3 (10): 1153-1160.
3. Kunz, S. and J. C. de la Torre (2017). "Breaking the Barrier: Host Cell Invasion by Lujo Virus." Cell Host Microbe 22(5): 583-585.
4. Sewlall, N. H., et al. (2014). "Clinical features and patient management of Lujo hemorrhagic fever." PLoS Negl Trop Dis 8(11): e3233.
5. Shao, J., et al. (2015). "Human hemorrhagic Fever causing arenaviruses: molecular mechanisms contributing to virus virulence and disease pathogenesis." Pathogens 4(2): 283-306.
6. Beyer WR, Popplau D, Garten W, von Laer D, Lenz O. Endoproteolytic processing of the lymphocytic choriomeningitis virus glycoprotein by the subtilase SKI-1/S1P. J Virol 2003;77(5):2866–72. [PubMed:125843 10]
7. Pinschewer DD, Perez M, Sanchez AB, de la Torre JC. Recombinant lymphocytic choriomeningitis virus expressing vesicular stomatitis virus glycoprotein. Proc Natl Acad Sci U S A 2003;100(13): 7895–900. [PubMed: 12808132]
8. Rojek JM, Lee AM, Nguyen N, Spiropoulou CF, Kunz S. Site 1 protease is required for proteolytic processing of the glycoproteins of the South American hemorrhagic fever viruses Junin, Machupo, and Guanarito. J Virol 2008a;82(12):6045–51. [PubMed: 18400865]
9. Tani, H., et al. (2014). "Analysis of Lujo virus cell entry using pseudotype vesicular stomatitis virus." J Virol88(13): 7317-7330.
10. Omar Hashim Ahmed, Arwa Abdelhalim, Sahar Obi, Khoubieb Ali Abdelrahman, Ahmed Hamdi, Mohammed A. Hassan: Immuno informatic Approach for Epitope-Based Peptide Vaccine against Lagos Rabies Virus Glycoprotein G. Immunome Research 08/2017;, DOI:10.4172/1745-7580. 1000137.
11. Mohammed AA, Hashim O, Elrahman KAA, Hamdi A, Hassan MA (2017) Epitope-Based PeptideVaccine Design Against Mokola Rabies Virus Glycoprotein G Utilizing In Silico Approaches. Immunome Res 13: 144. Doi: 10.4172/1745-7580.1000144.
12. Mohammed, A. A., et al. (2018). "Epitope-Based Peptide Vaccine Against Fructose-Bisphosphate Aldolase of Madurella mycetomatis Using Immunoinformatics Approaches." Bioinform Biol Insights12: 1177932218809703.
13. Skwarczynski M, Toth I. Peptide-based synthetic vaccines. Chemical science. 2016;7(2):842-54.
14. Mohammed, A. A., Shantier, S. W., Mustafa, M. I., Osman, H. K., Elmansi, H. E., Osman, I. A. A., Hassan, M. A. (2020). Epitope-Based Peptide Vaccine against Glycoprotein G of Nipah Henipavirus Using Immunoinformatics Approaches. Journal of Immunology Research, *2020*. https://doi.org/10.1155/2020/2567957.

15. Badawi MM, Osman MM, Alla AAF, Ahmedani AM, Abdalla Mh, Gasemelseed MM, et al. Highly Conserved Epitopes of ZIKA Envelope Glycoprotein May Act as a Novel Peptide Vaccine with High Coverage: Immunoinformatics Approach. American Journal of Biomedical Research. 2016; 4(3):46-60.
16. Li, W., Joshi, M., Singhania, S., Ramsey, K., & Murthy, A. (2014). Peptide Vaccine: Progress and Challenges. *Vaccines*, *2*(3), 515–536.
17. Oyarzun, P., Ellis, J. J., Gonzalez-Galarza, F. F., Jones, A. R., Middleton, D., Boden, M., & Kobe, B. (2015). A bioinformatics tool for epitope-based vaccine design that accounts for human ethnic diversity: Application to emerging infectious diseases. *Vaccine*, *33*(10), 1267–1273.
18. "NCBI protein sequence database," 2018, https://www.ncbi.nlm.nih.gov/protein.
19. Vita R, Overton JA, Greenbaum JA, Ponomarenko J, Clark JD, Cantrell JR, Wheeler DK, Gabbard JL, Hix D, Sette A, Peters B. The immune epitope database (IEDB) 3.0. Nucleic Acids Res. 2014 Oct 9. pii: gku938. [Epub ahead of print] PubMed PMID: 25300482.
20. Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser 41: 95-98.
21. Bui, H.H., Sidney, J., Li, W., Fusseder, N., Sette, A., 2007. Development of an epitope conservancy analysis tool to facilitate the design of epitope-based diagnostic sand vaccines. BMC Bioinform. 8 (1), 361.
22. Larsen JE, Lund O, Nielsen M (2006) Improved method for predicting linear B-cell epitopes. Immunome Res 2: 2.
23. Emini, E.A., Hughes, J.V., Perlow, D.S., Boger, J., 1985. Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. J. Virol. 55 (3), 836–839.
24. Kolaskar, A.S., Tongaonkar, P.C., 1990. A semi-empirical method for prediction of anti-genic determinants on protein antigens. FEBS Lett. 276 (1–2), 172–174.
25. Nielsen M, Lundegaard C, Worning P, Lauemøller SL, Lamberth K, et al. (2003) Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. Protein Sci 12: 1007-1017.
26. Kim Y, Ponomarenko J, Zhu Z, Tamang D, Wang P, et al. (2012) Immune epitope database analysis resource. Nucleic Acids Res 40: W525-W530.
27. Sinu Paul, Daniela Weiskopf, Michael A. Angelo, John Sidney, Bjoern Peters and Alessandro Sette J Immunol December 15, 2013, 191 (12) 5831-5839; DOI: https://doi.org/10.4049/jimmunol.1302101.
28. Zeynep Koşaloğlu-Yalçın, Manasa Lanka, Angela Frentzen, Ashmitaa Logandha Ramamoorthy Premlal, John Sidney, Kerrie Vaughan, Jason Greenbaum, Paul Robbins, Jared Gartner, Alessandro Sette & Bjoern Peters (2018) Predicting T cell recognition of MHC class I restricted neoepitopes, OncoImmunology, 7:11, e1492508, DOI: 10.1080/2162402X.2018.1492508.
29. Laimer J, Lackner P. MHCII3D—Robust Structure Based Prediction of MHC II Binding Peptides. *International Journal of Molecular Sciences*. 2021; 22(1):12. https://doi.org/10.3390/ijms22010012.
30. Zhang Q, Wang P, Kim Y, Haste-Andersen P, Beaver J, et al. (2008) Immune epitope database analysis resource (IEDB-AR). Nucleic Acids Res 36: W513-W518.
31. Choo, J.A.; Thong, S.Y.; Yap, J.; van Esch, W.J.; Raida, M.; Meijers, R.; Lescar, J.; Verhelst, S.H.; Grotenbreg, G.M. Bioorthogonal cleavage and exchange of major histocompatibility complex ligands by employing azobenzene-containing peptides. Angew. Chem. 2014, 53, 13390–13394.
32. Källberg M, Wang H, Wang S, Peng J, Wang Z, Lu H, Xu H. Temple-based protein structure modelling using the RaptorX website server. Nature Protocols. 2012; 7: 1511–1522.
33. Goddard TD1, Huang CC, Ferrin. TE Software extensions to UCSF chimera for interactive visualization of large molecular assemblies. J Structure (2005); 13(3):473-82.
34. Zheng J, Lin X, Wang X, Zheng L, Lan S, Jin S, Ou Z, Wu J. In Silico Analysis of Epitope-Based Vaccine Candidates against Hepatitis B Virus Polymerase Protein. *Viruses*.

2017; 9(5):112. https://doi.org/10.3390/v9050112.<http://www.rcsb.org/pdb/home/home.do>.
35. Morris GM, Goodsell DS, Halliday RS, et al (1998) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. J Comput Chem 19:1639–1662. doi:10.1002/(SICI)1096-987X(19981115)19:14<1639::AID-JCC10>3.0.CO;2-B.
36. Morris GM, Ruth H, Lindstrom W, et al (2009) Software news and updates AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. J Comput Chem 30:2785–2791. doi: 10.1002/jcc.21256.
37. Pettersen EF, Goddard TD, Huang CC, et al (2004) UCSF Chimera - A visualization system for exploratory research and analysis. J Comput Chem 25:1605–1612. doi:10.1002/jcc.20084.