Identification of peptides containing epitopes of Ebola virus eliciting immune response

A thesis

Submitted in the partial fulfilment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

IN

BIOTECHNOLOGY

SAHIL JAIN

(Registration No. 901400009)

Department of Biotechnology

Thapar Institute of Engineering and Technology,

Patiala, Punjab

August, 2019

CERTIFICATE

Certified that the thesis "**Identification of peptides containing epitopes of Ebola virus eliciting immune response**" submitted by Mr. Sahil Jain, in the fulfilment of the requirement for the award of the Degree of **Doctor of Philosophy** in the department of Biotechnology, Thapar Institute of Engineering and Technology, Patiala, is a record of candidate's own independent and original research work carried out by himself in my supervision and guidance. The material embodied in this thesis has not been submitted in parts or full in any other university or institute for the award of any degree.

Supervisor

28/05/2020

Dr. Manoj Baranwal

Associate Professor

Department of Biotechnology

Thapar Institute of Engineering

and Technology, Patiala (147004),

Punjab, India

i

DECLARATION

I hereby declare that the work presented in the thesis entitled "**Identification of peptides containing epitopes of Ebola virus eliciting immune response**" in the fulfilment of the requirement for the award of the Degree of **Doctor of Philosophy** in the department of Biotechnology, Thapar Institute of Engineering and Technology, Patiala, is an authentic record of my own work during the period January 2015 to August 2019, under the supervision of Dr. Manoj Baranwal, Associate Professor, Department of Biotechnology, Thapar Institute of Engineering and Technology, Patiala. The material embodied in this thesis has not been submitted in parts or full in any other university or institute for the award of any degree in India or abroad.

Place: Patiala

Date: 28th May 2020

Sahil Jain

Sahil Jain

ii

This work is dedicated to my lifelines:

Father - Dr. Akshey Jain

Mother - Dr. Meenakshi Jain

Brother - Armaan Jain

ACKNOWLEDGEMENT

As a student, i always dreamt of graduating from TIET and today, i feel blessed and honoured to have pursued my post graduation from this prestigious institute. I would like to thank ***Prof. Prakash Gopalan*** *(Director, TIET),* ***Dr. O.P. Pandey*** *(ex-Dean of Research and Sponsored Projects),* ***Dr. Rafat Siddique*** *(Dean of Research and Sponsored Projects),* ***Dr. Seema Bawa*** *(ex-Dean of Student Affairs),* ***Dr. Maneek Kumar*** *(Dean of Student Affairs) and* ***Dr. S.S. Bhatia*** *(Dean of Academic Affairs) for providing me with this grand opportunity.*

On the outset, i would like to express my heartfelt and endless gratitude towards ***Dr. Manoj Baranwal****, my guide and mentor. A gem of a person, he is a true reflection of professionalism and humanity. I shall always be indebted to him for his patience and teachings which helped in honing the naive researcher in me as well as helped in making me a better person. Sir, you shall always be remembered!*

I would also like to express my gratitude towards ***Dr. Moushumi Ghosh****, HOD, BTD, TIET, Patiala. Her support throughout my tenure was a pillar of strength for me and her leadership inspires security and freedom to students to perform in a responsible yet independent manner. I would like to covey my deepest regards to* ***Dr. M.S. Reddy****. He is an ideal figure whose professionalism, coherence and simplicity are astounding. I would also like to thank* ***Dr. Niranjan Das*** *for his valuable and revered guidance from time to time. His intense passion towards guiding students inspires us to take that extra step and deliver more.*

My gratitude goes to ***Dr. Dinesh Goyal,*** *ex-HOD, BTD, TIET as well as* ***Dr. Vikas Handa****, a member of my doctoral committee who taught me to strive for perfection in every field of life. I am also grateful to other members of my doctoral committee,* ***Dr. Siddharth Sharma and Dr. Diptiman Choudhary*** *for their valuable inputs from time to time. I would like to take this opportunity to thank all the other current and ex-faculty members,* ***Dr. Sanjai Saxena, Dr. Anil Kumar, Dr. Shekhar Agnihotri, Mrs. M. Vasundhara and Dr. Manju Anand****, for their guidance and well wishes throughout my study. Special thanks to* ***Dr. Ravi Kiran****, who helped me reinvent myself.*

My endless gratefulness and thankfulness goes towards ***Dr. Neha Lohia*** *(My first friend and support at TIET).*

My seniors acted as the guiding lights which made my journey easier. I would like to thank ***Mr. Nadeem Akhtar, Mrs. Amanpreet Kaur, Ms. Saloni Sharma, Mr. Arkadeep Mukherjee, Ms. Bharti Thakur, Ms. Shikha Khullar, Ms. Purnima Sharma,*** *and* ***Mr. Vagish Dwibedi*** *for their care and support.*

I express my sincerely felt appreciation for the efforts and timely help of all the current and ex-lab associates, ***Mr. Ram Nawal*** *(the one-man army!),* ***Mr. Lallan Yadav, Mr. Surinder, Mr. Mohinder, Mr. Phool Chand, Mr. Prabhat*** *and* ***Mr. Chandan Bhandari.*** *Also, my sincere thanks goes to BTD office personnel,* ***Mr. Anirudh Handa*** *and* ***Mrs. Manjula****, for their support.*

I am highly obliged to ***Dr. Nitin Gupta*** *(Nitin Hospital, Patiala),* ***Dr. Vandana Singla*** *(Rajendra Hospital, Patiala) and* ***Dr. Rimpreet Walia*** *(Lifeline Hospital, Patiala) for providing healthy blood samples. A vote of thanks to* ***Lalit Gas Agency*** *(Patiala),* ***Pooja Life Sciences*** *(Patiala) and* ***Essence Life Sciences*** *(Chandigarh) for their support throughout my study.*

My lab-mates were an integral part of my journey as along with growing professionally in their company, i also spent some valuable and memorable time with them. I would like to thank ***Dibyangana, Navdeep Kaur, Mehendi Goyal*** *(Miss Granger),* ***Harlovleen Kaur*** *(Tonks),* ***Tavleen Saini*** *(Ginny),* ***Vanita Kinra, Garima Batish, Tania Patwal, Baneet Chawla, Khushpreet Virk, Neha Garg, Neha Kaushal, Neha Srivastava and Yogita Kautish*** *for the memories.*

My journey would have been incomplete without the support and well wishes of my friends. Life felt better and pursuing PhD a bit easier in the company of ***Mrs. Gurpreet Kaur*** *and* ***Ms. Harleen Kaur Walia****, two of the most genuine and caring persons i have ever met. Your true friendship is deeply cherished! My gratefulness is also shared by my dear friends,* ***Ms. Preeti Khokra, Mr. Vivek Sharma and Mr. Kulwinder Sran****. A special shout-out to my young friends,* ***Sutirth Baranwal*** *and* ***Urvija Baranwal****; keep rocking!*

Lastly, i would like to mention the most important part of my journey and life; my family: My father ***Dr. Akshey Jain****, my mother* ***Dr. Meenakshi Jain*** *and my brother* ***Mr. Armaan Jain.*** *They are my lifelines and i am forever indebted to them for everything. I am also highly grateful to* ***Dr. Amandeep Aggarwal*** *(Uncle),* ***Mr. Jagmohan Garg*** *(Uncle),* ***Mrs. Madhu Garg*** *(Aunt),* ***Mrs. Aarti Garg*** *(Sister),* ***Mehak*** *(Sister),* ***Jhanvi*** *(Sister),* ***Dr. P.B. Barman***

(Uncle) and ***Renu*** *(Sister). A special shout-out to my rockstar,* ***Aarya*** *(Niece).* ***I*** *was lucky to have support of my friends and family members. This work is dedicated to them.*

Date: 28th May 2020

Place: Patiala

Sahil Jain.

(Sahil Jain)

vi

List of Publications

Published

1. Dhiman, G., Lohia, N., Jain, S. & Baranwal, M. (2016) Metadherin peptides containing $CD4^+$ and $CD8^+$ T cell epitopes as vaccine candidate against cancer. *Microbiology and immunology*, 60(9):646-52. (SCI, IF = 1.44)
2. Jain, S. & Baranwal, M. (2019) Computational analysis in designing T cell epitopes enriched peptides of Ebola glycoprotein exhibiting strong binding interaction with HLA molecules. *Journal of Theoretical Biology*, 465:34-44. (SCI, IF = 1.87)
3. Jain, S. & Baranwal, M. (2019) Conserved peptide vaccine candidates containing multiple Ebola nucleoprotein epitopes display interactions with diverse HLA molecules. *Medical Microbiology and Immunology*, 208(2):227-238. (SCI, IF = 2.96)

Manuscript under preparation

1. Jain S. & Baranwal M. Conserved immunogenic peptides of Ebola glycoprotein illicit immune response
2. Jain S. & Baranwal M. A comprehensive review of the role of immune system and vaccine development strategies against Ebola virus

Participation in Conferences and Symposia

International

1. Jain S. & Baranwal M (2019). *In silico* identified immunogenic Ebola nucleoprotein peptides elicit immune response. "International Conference on Bioinformatics and Systems Biology" Singapore (Mar 20-21, 2019).

National

1. Dhiman G., Lohia N., Jain S. & Baranwal M (2016). *In silico* prediction of immunogenic metadherin peptides containing overlapping T-cell epitopes. "FIMSA Advanced Immunology Course" conducted at PGI, Chandigarh, India (Mar 17-19, 2016).
2. Attended "Computational Drug Design and Molecular Docking" conducted by Schrodinger at Thapar Institute of Engineering and Technology, Patiala, India (Apr 27-29, 2018).
3. Jain S. & Baranwal M (2018). Computational analysis in designing T cell epitopes enriched peptides of Ebola glycoprotein exhibiting strong binding interaction with HLA molecules. "Immunotherapy and advances in immunology (IMMUNOCON 2018)" THSTI, Faridabad (Nov 1-3, 2018).

Abstract

Last 5 years have witnessed two of the worst Ebola virus outbreaks in history which have resulted in significant morbidity and mortality. Rapid spread of the virus across the globe has raised concerns over the safety of world citizens. No licensed vaccine or drug protective against human infecting Ebola species is available as yet. In a panic driven step, incompletely tested vaccines have been approved along with antiviral prophylaxis to counter unforeseen outbreaks. However, production difficulties, safety concerns, high booster dosage requirement, inefficient delivery systems and pre-existing immunity are some of the challenges faced by current vaccine development approaches. Hence, there is a pressing need to develop a vaccine strategy which can offer universal or at least broad protection against current and future Ebola virus strains in populations distributed worldwide. Highly conserved peptide fragments belonging to critical viral proteins and containing multiple epitopes which have the capacity to interact with a wide array of HLA molecules are anticipated to serve as potent candidates for a universal or broadly reactive Ebola vaccine. In light of the above facts, the present study is oriented towards identifying highly conserved promiscuous peptides containing multiple overlapping T ($CD8^+$ and $CD4^+$) and B cell epitopes and devoid of undesirable responses (autoimmunity, toxicity and allergenicity) in glycoprotein (GP) and nucleoprotein (NP) of Ebola virus using immunoinformatics techniques such as epitope prediction tools, HLA and population coverage analysis and molecular docking. Further, peripheral blood mononuclear cells (PBMC) from healthy volunteers were subjected to repetitive stimulation by peptides that presented the best *in silico* results to assess their immunogenic response by measuring cell proliferation and IFN-γ production with the help of MTT and ELISA assays respectively. Fifteen GP and eighteen NP peptide fragments containing multiple $CD8^+$ T cell (HLA class I) epitopes while thirteen GP and fifteen NP fragments containing multiple $CD4^+$ T cell (HLA class II) epitopes were obtained on merging the overlapping epitopes. Six GP and twelve NP peptide fragments common to their respective peptides containing both, $CD8^+$ and $CD4^+$ T cell epitopes were generated. B cell epitope prediction indicated the presence of B cell epitopes in four GP and eight NP peptide fragments. After screening for undesirable responses, four GP and six NP peptide fragments containing both, T and B cell epitopes, were obtained. Majority of the identified peptides were found to be 100% conserved in Zaire ebolavirus, the most pathogenic Ebola species while Gp3 and Np3 were also found to be 100% in other human infecting Ebola virus species (Sudan, Bundibugyo and Taï forest). Rest of the peptides were found to be conserved with

minor variations while in other *Filoviridae* members (*Lloviuvirus* and *Margburgvirus*), these peptides have minor variations or could not be located. HLA coverage analysis based on prediction tools revealed that majority of the identified peptides displayed interaction with several HLA alleles/supertypes. Population coverage analysis indicated that the all peptides covered 69-100% population base in Asian, American, African and European continents with an average population coverage of ≥85% for these continents. Docking results revealed that the binding energy (Autodock Vina) and RMSD (CABS-dock) values obtained for all peptides were mostly found to be in range of native peptide values. Interestingly, results for some peptides outshined those of native peptides for corresponding HLA molecules. Based on the performance of identified peptides during peptide-HLA interaction analysis, six peptides (Gp1, Gp3 and Gp4, Np2, Np3 and Np5) were selected for further evaluation via *in vitro* experimentation. Peptide stimulation assays in PBMC revealed that all six peptides resulted in a significant peptide induced proliferation in healthy blood samples with Np3, Gp4 and Gp1 generating a significant response in ≥ 7 out of ten samples. Also, all peptides (except Np5) displayed a significant response in inducing peptide specific IFN-γ release in ≥ 7 samples. The results suggested that the *in silico* approach applied in combination with *in vitro* experimentation successfully selected peptides of glycoprotein and nucleoprotein which are capable of eliciting immune response, thus, making them potential candidates for global Ebola vaccine development.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ABTS	2,2'-Azinobis[3-ethylbenzothiazoline-6-sulfonicacid]-diammonium salt
Ad	Adenovirus
ADT	AutoDock Tools
ANN	Artificial Neural Network
ANOVA	Analysis of variance
APC	Antigen presenting cells
ASGP-R	Asialoglycoprotein receptor
ATP	Adenosine triphosphate
AVANA	Antigenic variability analyser
BALB	Bagg Albino (inbred research mouse strain)
BB	Branch and bound
BLAST	Basic Local Alignment Search Tool
BSA	Bovine serum albumin
BSL	Biosafety Level
CABS	Carbon alpha (CA) carbon beta (B) side chain (S)
cAd	Chimpanzee adenovirus
CD	Cluster of differentiation
CDC	Centre for Disease Control and Prevention
CLEC4G	C-Type Lectin Domain Family 4 Member G
CMV	Cytomegalovirus
CO_2	Carbon dioxide
Con A	Concanavalin A
cRNA	Complementary ribonucleic acid
CTD	C-terminal domain
CTL	Cytotoxic T cell
DC-SIGN	Dendritic cell-specific ICAM-3-grabbing nonintegrin
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DP	Dynamic programming

DRC	Democratic Republic of the Congo
dsRNA	Double stranded ribonucleic acid
EBI	European Bioinformatics Institute
EBOV	Ebola virus
EDTA	Ethylenediaminetetraacetic acid
EHF	Ebola hemorrhagic fever
ELISA	Enzyme-linked immunosorbent assay
ELISpot	The enzyme-linked immunospot
EMBL	European Molecular Biology Laboratory
EVD	Ebola virus disease
FBS	Foetal bovine Serum
FDA	Food and Drug Administration
FR	Folate receptor
GAS	Group A streptococcus
GP	Glycoprotein
GRFT	Griffithsin
GUI	Graphic user interface
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HIV	Human immunodeficiency virus
HLA	Human leukocyte antigen
hnRNP	Heterogeneous nuclear ribonucleoprotein
HPIV3	Human parainfluenza virus type 3
HR	Heptad region
HRP	Horseradish peroxidase
IC_{50}	Half maximal inhibitory concentration
ICTV	International Committee on Taxonomy of Viruses
IEDB	Immune epitope database
IFN	Interferon
Ig	Immunoglobulin
IL	Interleukin
IRF3	Interferon Regulatory Factor 3

KCl	Potassium Chloride
kDa	Kilodalton
KH_2PO_4	Potassium phosphate monobasic
KISS	Kernel-based Inter-allele peptide binding prediction System
L	Polymerase
mAb	Monoclonal antibody
MAFFT	Multiple Alignment using Fast Fourier Transform
MARV	Marburg virus
MAST	Motif Alignment and Search Tool
MBL	Mannose binding lectin
MDA5	Melanoma Differentiation-Associated protein 5
MEME	Multiple EM for Motif Elicitation
MHC	Major histocompatibility complex
MLD	Mucin like domain
MMR	Measles-Mumps-Rubella
mRNA	Messenger ribonucleic acid
MTT	3-(4,5-Dimethylthiazol-2-yl)-2,5-Diphenyltetrazolium Bromide
MUSCLE	Multiple sequence comparison by log expectation
MVA	Modified vaccinia virus Ankara
Na_2HPO_4	Sodium phosphate dibasic
NaCl	Sodium Chloride
$NaHCO_3$	Sodium Bicarbonate
NC	Nucleocapsid
NCBI	National Center for Biotechnology Information
NF-kB	Nuclear factor-kappa B
NHP	Non-human primates
NIAID	National Institute of Allergy and Infectious Diseases
NK	Natural Killer
NNS	Nonsegmented, negative sense
NP	Nucleoprotein
OPEP	Optimized Potential for Efficient protein structure

ORF	Open reading frame
ORS	Oral rehydrating solution
PBMC	Peripheral blood mononuclear cells
PDB	Protein data bank
PKR	Protein Kinase R
PSSMs	Position Specific Scoring Matrices
QM	Quantitative matrices
RABV	Rabies virus
RBD	Receptor binding domain
RdRp	RNA-dependent RNA polymerase
RIG	Retinoic acid-inducible gene
RLR	RIG-I-like receptor
RMSD	Root mean square deviation
RNA	Ribonucleic acid
RNP	Ribonucleoprotein
RPMI	Rosewell Park Memorial Institute
rVSV	Recombinant vesicular stomatitis virus
SERM	Selective estrogen receptor modulators
sGP	Secreted glycoprotein
SI	Stimulation index
SIV	Simian immunodeficiency virus
ssGP	Small soluble glycoprotein
ssRNA	Single stranded ribonucleic acid
STAT	Signal transducers and activators of transcription
SVM	Support vector machine
TAM	Tyro3 protein kinase
TAP	Transporter associated with antigen processing
T_c	T-cytotoxic
T-Coffee	Tree-based Consistency Objective Function for alignment Evaluation
T_H	T-helper
TLR	Toll like receptor

TNF	Tumour necrosis factor
TPC2	Two-Pore Channel 2
TSG101	Tumor susceptibility gene 101
US	United states
VEEV	Venezuelan equine encephalitis virus
VHF	Viral haemorrhagic fever
VLPs	Virus like particles
VP	Virion protein
VPS4	Vacuolar protein sorting-associated protein 4
WHO	World Health Organisation

LIST OF SYMBOLS

%	Percentage
~	Similar
<	Less than
>	Greater than
±	Plus-minus
≤	Less than or equal to
≥	Greater than or equal to
°C	Degree Celsius
μ	Micron
μg/mL	Microgram per milliliter
Å	Angstrom
h	Hours
mL	Milliliter
α	Alpha
β	Beta
γ	Gamma

Chapter 1: Introduction

Ebola virus (EBOV) presents the scientific community with a daunting challenge in the form of Ebola virus disease (EVD). EVD, earlier known as Ebola hemorrhagic fever (EHF), is a fatal disease and infects humans as well as primates such as monkeys, chimpanzees and gorillas. It has periodically caused major social issues and financial losses. Since the turn of the century, EBOV cases have been on a rise with over 12,000 deaths reported till date (WHO, 2019). After an outbreak, the rapid rate at which EBOV spreads can be attributed to person to person transmission upon direct contact with either body fluids such as stool, urine, semen and saliva or by contact with pierced skin of a victim (Ilesanmi and Alele, 2014). Generally, symptoms appear after an incubation period ranging from 2-21 days, typically within 4-10 days (Goeijenbier et al., 2014). Symptoms include fever (>38.6°C), muscular pain and weakness, vomiting, anorexia, abdominal pains, vomiting and diarrhoea. Persistent infection results in unexplained haemorrhage, electrolyte disorders and multiple vital organ (liver, respiratory and renal) failure (Safari et al., 2015).

Till date, six Ebola virus species are known out of which, four species are known to infect humans. The human infecting species are *Zaire ebolavirus* (commonly called as Ebola virus), *Sudan ebolavirus*, *Taï Forest ebolavirus* (earlier known as Côte d'Ivoire ebolavirus) and *Bundibugyo ebolavirus* (Groseth et al., 2012). *Reston ebolavirus* and *Bombali ebolavirus* (discovered on 27th July, 2018) are the other two species. *Reston ebolavirus* is known to infect non-human primates (Elisha and Adegboro, 2014; Geisbert et al., 2009). Out of all the species, Zaire is the most pathogenic in humans. EBOV first surfaced on June 27, 1976 in the Democratic Republic of the Sudan and northeast Zaire where it claimed 150 and 280 lives respectively (WHO report, 1978). *Taï Forest ebolavirus* first surfaced in 1994 when it infected an ethnologist who was performing an autopsy on a chimpanzee found dead in Côte d'Ivoire (Le Guenno et al., 1995). *Bundibugyo ebolavirus* came to limelight in 2007 in Uganda where it infected 116 people with 26% mortality (Towner et al., 2008). 2014 outbreak in West Africa was the worst in the history as it claimed over 11,000 lives (Okoror et al., 2018). As of July, 2019, more than 30,000 EBOV cases have been reported since 1976 and the second worst outbreak in history is ongoing in North Kivu and has already claimed more than 1600 lives (WHO, 2019).

An effective prophylaxis, approved treatment or licensed vaccine for EVD is not yet available (Dhama et al., 2018; Mathebula et al., 2019). Post exposure treatment and care includes maintenance of electrolyte balance with the help of oral rehydrating solution (ORS), providing healthy nutritional diet and comforting with the help of analgesics (Folayan et al., 2016). To control the after effects such as diarrhoea and vomiting, the administration of antidiarrhoeal and antiemetics is recommended (Fowler et al., 2014; Schieffelin et al., 2014). Also, the victims are particularly prone to secondary bacterial infections and therefore, antimicrobial agents with cephalosporins are administered in such cases (Plachouras et al., 2015). Moreover, the susceptibility of victims towards parasitic infections such as malaria further complicates the recovery process (O'Shea et al., 2015).

The infamous West African epidemic has spurred efforts to develop effective counter measures against EBOV virus. As of May, 2019, atleast thirteen vaccines candidates are in different phases of clinical trials (Mathebula et al., 2019). These vaccine candidates include recombinant adenoviruses, recombinant human para influenza viruses, virus-like particles and recombinant vesicular stomatitis viruses. Until approval of a vaccine candidate, rVSVΔG-ZEBOV-GP has been recommended to be used in cases of emergency while two vaccines viz., GamEvac-Combi and Ad5-EBOV have been licensed regionally in Russia and China respectively (Mathebula et al., 2019) even though they are still in nascent stages of testing (Pharmaintelligence, 2019).

The current vaccine development approaches suffer from some limitations such as requirement of initial priming with another vaccine, need for repeated dosages to achieve the recommended antibody titer level and the issue of pre-existing immunity (Sridhar, 2015). To overcome some of these limitations, the trend is shifting towards development of peptide based vaccines (Martins et al., 2015). These vaccines consist of a peptide fragment composed of multiple immunogenic epitopes with the ability to generate an immune response against one or multiple pathogens. Peptide vaccines are capable of conferring both, cell mediated and humoral immunity. They help in reducing risk of infectious agent leaks, are easily modified by addition of lipids or phosphates to enhance immunogenicity and eliminate the risk of reversion or genetic recombination (Fosgerau and Hoffmann, 2015). Several studies have been conducted to decipher potent peptide vaccine candidates against cancer (Li et al., 2014), influenza (Francis et al., 2015; Lohia and Baranwal, 2015) and HIV (Huang et al., 2018).

Elucidation of plausible peptide vaccine candidates can be fastened with the help of computational immunology or immunoinformatics which is a branch that aims to investigate

genomic and proteomic data related to immunology by means of various statistical models and machine learning algorithms to obtain immunologically relevant results. Epitope prediction tools help in making a knowledge based selection of probable epitopes which can then be validated via *in vitro/in vivo* experiments. A variety of tools for predicting HLA class I and class II binders are available. Numerous researchers have worked on different pathogens and diseases such as *Influenza virus* (Lohia and Baranwal, 2015), *Human papilloma virus* (Kumar et al., 2015), hepatitis C (Sirskyj et al., 2011) and cancer (Naik et al., 2012) by using various computational approaches for elucidation of epitopes. This has dramatically reduced the time and efforts required to identify the potential epitopes based on their capability to bind HLA in contrast to the traditional approach which involves the synthesis and assessment of all possible overlapping epitopes. Therefore, it offers a promising alternative method which is less expensive as compared to the experimental determination.

Further, a vaccine candidate must exhibit interaction with multiple HLA alleles belonging to different populations of the world to be effective globally (Ling and Whitton, 1997). Molecular docking (Patronov and Doytchinova, 2013; Singh et al., 2014) and population coverage analysis (Dewi et al., 2019; Pritam et al., 2019) have been recognized as valuable techniques in computer aided drug and vaccine design and help to predict the peptide-HLA interaction for a wide array of HLA molecules. The identification of epitope enriched peptides by epitope prediction tools and peptide-HLA interaction analysis is a screening step and not a stand-alone technology to identify immunogenic peptides. These peptides are required to be validated for elicitation of immune response in *in vitro* and *in vivo* systems. Successful *in vitro* and/or *in vivo* validation of eptiopes predicted via computational approaches has been carried out for *Influenza virus* (Lohia and Baranwal, 2017), *Leishmania donovani* (Dikhit et al., 2017), *Moraxella catarrhalis* (Yassin et al., 2016) and *Brucella abortus* (Afley et al., 2015).

In light of these facts, present study focused on computational identification of conserved peptide fragments containing multiple overlapping T ($CD8^+$ and $CD4^+$) and B cell epitopes in Ebola glycoprotein and nucleoprotein by employing conservation studies, epitope prediction tools and different computational algorithms to screen peptides for undesirable (autoimmune, allergic and toxic) responses. The identified peptides were analyzed for interaction with diverse HLA (human MHC) molecules based on HLA coverage analysis, population coverage analysis and molecular docking. Further, the selected peptides were evaluated for their immunogenic response by measuring peptide induced cell proliferation and interferon-

gamma (IFN-γ) release in peripheral blood mononuclear cells (PBMC) isolated from blood samples of healthy volunteers.

Chapter 2: Review of Literature

2.1. Ebola virus disease (EVD)

Periodic Ebola virus (EBOV) outbreaks have garnered significant attraction since their inception in Africa, 1976. The disease severity, fear of devastating effects rendered by EBOV when used as a biological weapon and the possible spread into the western world have led to its classification as a Category A Priority pathogen by NIAID and a bioterrorism Category A Agent by the CDC (La Vega et al., 2015; Hoenen et al., 2012).

EBOV zoonotic hosts and their role in Ebola virus disease (EVD) escalation is yet to be clearly identified (Jun et al., 2015). EBOV transmission from its zoonotic reservoir to susceptible human recipients is thought to take place via the spread of infected reservoir hosts and human exposure to their blood or carcasses (Plowright et al., 2015). The transmission occurs from a victim to a healthy person via close contact with the victim, their pierced skin or contact with body fluids of the victim such as urine, saliva, semen, blood and stools (Ilesanmi and Alele, 2014).

EVD is a hemorrhagic fever resulting in high fever, muscular pain and weakness, malaise and headache followed by nausea and diarrhoea. The victim shows signs of hemorrhagic symptoms such as hematemesis, hematochezia and melena 5-7 days after the initial infection. In later stages, multi-organ failure and a state of shock may lead to death (Feldman and Geisbert, 2011; Hoenen et al., 2007). 40% of the survivors continue to show symptoms like inflammation at the "immune privileged" sites (sites with attenuated immunity such as eyes, testis, synovium and meninges) even after passing medical tests for viral presence (Matzinger and Kamala, 2011).

Six Ebola virus species are known so far viz., a) *Taï Forest ebolavirus* b) *Zaire ebolavirus* c) *Sudan ebolavirus* d) *Reston ebolavirus* e) *Bundibugyo ebolavirus* and f) *Bombali ebolavirus*. All species except Reston and Bombali have been reported to cause human infections. Zaire is the most pathogenic species for humans (Groseth et al., 2012) and architected the infamous 2014-2016 epidemic in West Africa, the biggest EBOV outbreak in history which registered more than 30,000 cases and claimed >11000 lives (CDC, 2019; Brannan et al., 2019).

2.2. Ebola virus

Ebola virus (EBOV) is an enveloped, non-segmented and negative–stranded RNA virus belonging to the filoviridae family (Khan et al., 2015).

2.2.1. Classification and nomenclature

The genus *Ebolavirus* belongs to the family *Filoviridae* and has six species (Dhama et al., 2018). The sixth species, *Bombali ebolavirus*, was discovered in 2018. A schematic taxonomical classification of EBOV done by International Committee on Taxonomy of Viruses (ICTV) (Kuhn, 2017) has been depicted in Figure 2.1, page 6.

Members of this family are pleomorphic and might appear as either circles, filaments or in a number six formation. They cause fulminant hemorrhagic fever, sporadically in humans and principally in nonhuman primates (NHPs) (Miraglia, 2018). Possible routes of zoonotic transmission are general lack of basic precautions, exposure to reservoir, handling contaminated meat and/or being bitten by the infected animal (Azarian et al., 2015).

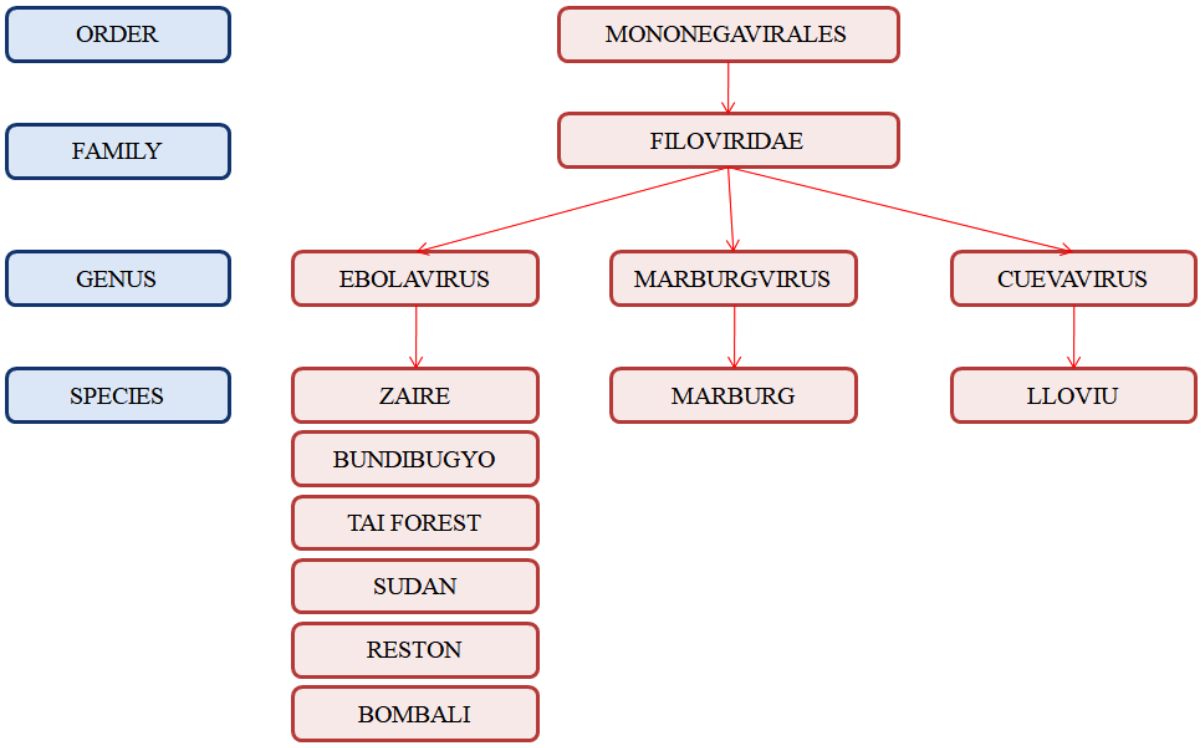


Figure 2.1: Taxonomical classification of Ebola virus

2.2.2. Structure of Ebola virus

A simplified structure of Ebola virus has been provided in Figure 2.2, page 7. EBOV is a thread-like, thin virus with the ability to alter its shape (pleomorphism). It has an 80 nm diameter, 1200 nm mean length and 1000 nm virion filament length (Geisbert and Jahrling, 1995). A 19kb linear, non-segmented, single-stranded RNA serves as the viral genome (Grifoni et al., 2016). It encodes for seven genes positioned successionally in the order: a) nucleoprotein (NP) b) viral protein (VP) VP35 c) VP40 d) glycoprotein (GP) e) VP30 f) VP24 and lastly g) polymerase (L) genes (Sanchez et al., 2007) as depicted in Figure 2.3, page 8. All four virion proteins comprise of the structural proteins and along with NP form the nucleocapsid-associated proteins (Radzimanowski et al., 2014). Nucleocapsid (NC) formation is necessary for EBOV existence as its genome, like other ~ssRNA genomes, cannot exist in naked form (Sun et al., 2012). Also, NC protects the viral genome from deterioration as well as serves as a template for genome replication (Ruigrok et al., 2011). Nucleoprotein and polymerase proteins along with VP30 and VP35 interact strongly with RNA resulting in the formation of ribonucleoprotein (RNP) complex which is enclosed on the outside by VP40. Membrane-anchored glycoprotein trimers are present throughout the surface of the virion and protrude from this outer envelope (Emanuel et al., 2018).

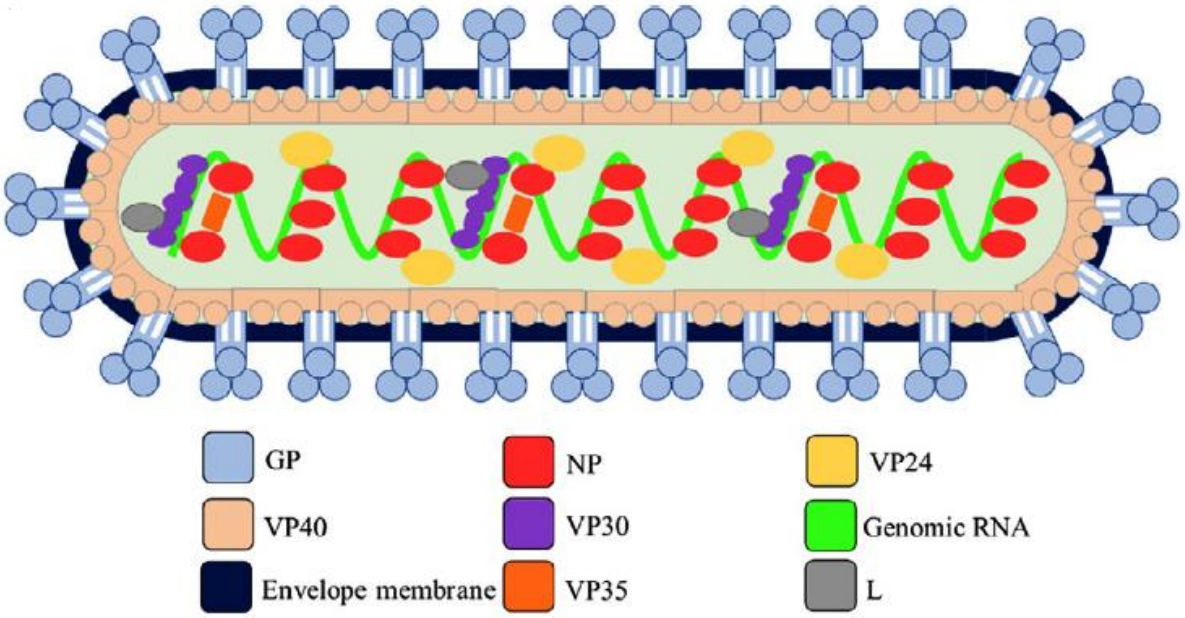


Figure 2.2: Ebola virus structure (Emanuel et al., 2018)

Figure 2.3: The sequential order of Ebola virus genes (Cantoni and Rossman, 2018)

Each gene is characterized by one open reading frame (ORF) except the fourth gene viz., GP gene which consists of three overlapping ORFs (Mehedi et al., 2011; Ikegami et al., 2001). Each of the genes is surrounded by highly preserved sites required for starting and ending the transcriptional process (Feldmann et al., 1992; Groseth et al., 2002). A long intergenic region is present between VP30 and VP24 genes (Crary et al., 2003) while short, intergenic regions consisting of 4-7 highly conserved nucleotides separate the rest of the genes (Bukreyev et al., 1995). A feature unique to *henipaviruses* (*Paramyxoviridae* family) and viruses belonging to *filoviridae* is the presence of long, untranslated regions (Wang et al., 2001). Conserved and complimentary sequences form stem-loop structures at the terminal ends of the EBOV genome. A 60bp leader sequence is present at the 3' end while a trailer sequence of variable length (as large as 676 bp) has been documented at the 5' end (M€uhlberger et al., 1992; Sanchez and Rollin, 2005).

2.2.2.1. Glycoprotein

Owing to the three overlapping ORFs present in the fourth EBOV gene, it results in the formation of three proteins viz., a) A protein consisting of 676 amino acids and interacting with transmembrane (termed as GP) b) A shorter, 364 amino acid containing secreted glycoprotein (termed as sGP) and c) Small soluble glycoprotein (termed as ssGP) (Sanchez et al., 1996; Sanchez et al., 1998) (Figure 2.4, page 9). sGP and ssGP are the only EBOV proteins that are completely non-structural (Emanuel et al., 2018).

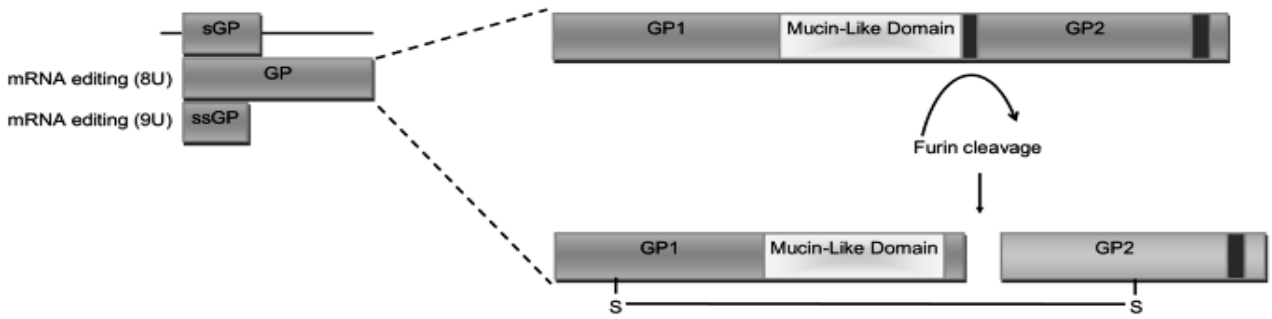


Figure 2.4: Transcriptional editing of Ebola glycoprotein (Cantoni and Rossman, 2018)

EBOV GP can be found on the exterior of the virus and is an assembly of a stable trimer formed of heterodimers (Lee et al., 2008). Post-translational cleavage of GP is the reason behind the formation of disulphide-linked GP1 and GP2 subunit (Jeffers et al., 2002). Three monomers, each consisting of one disulphide-linked GP1 and GP2 subunit, attach non-covalently to form a EBOV GP trimer. This trimer is shaped in the form of a chalice with approximate dimensions of 95A° * 395A° * 370A°. A bowl shaped structure is formed by the three ectodomains of GP1 while the helices of the three GP2 subunits surround the bowl (Lee et al., 2008).

GP1 is accountable for coupling with host cells. Residues 54-201 are thought to be involved in this process (Kuhn et al., 2006). GP1 can be split into three regions viz., a) Base b) Head and C) Glycan cap (Lee et al., 2008). The base region comprises of two sets of β-sheets in the form of a semicircular surface. This surface interacts with and binds one of the helices of GP2 subunit and an interior fusion loop with the help of its hydrophobicity. Cys 53 present in the base region is thought to be responsible for the formation of an intermolecular disulphide bridge leading to Cys 609 present in the GP2 subunit (Jeffers et al., 2002). The glycan cap region forms no contacts with the neighbouring monomers and is completely in the open on the exterior of the chalice (Lee et al., 2008).

GP2 is accountable for membrane fusion of the virus and the host cell. It consists of a) The internal fusion loop and b) HR1 and HR2, together known as the heptad repeat regions. The internal fusion loop comprises of 511-556 amino acids and is found as a scaffold of anti-parallel β-strands. Due to this confirmation, a fusion peptide, which is partly helical and partly hydrophobic, comprising of residues 529, 531, 532, 533, 534 and 535 is visible (Lee et al., 2008). The side chains belonging to the aforementioned amino acid moieties interact with

the head region of a neighbouring monomer GP1 subunit; a feature that draws similarities with packaging of fusion peptide witnessed in the pre-fusion *Human parainfluenza virus* (Yin et al., 2006). A disulphide bond between Cys 511 residue of the internal fusion loop and the Cys 556 residue of the HR1 helix covalently links the anti-parallel β-sheet. This bond is completely conserved in all *filoviridae* members and presents an analogy to the cysteine pair in *Avian sarcoma leukosis virus* (Delos et al., 2008).

2.2.2.2. Nucleoprotein

Each EBOV consists of ~3200 NP moieties (Bharat et al., 2012) with each of the molecules comprising of 739 residues. NP N-terminal end is hydrophobic and forms a condensed helix and a variable hydrophilic C-terminal which binds to VP40 (Niikura et al., 2001; Bharat et al., 2012). Residues 1-450 are conserved amongst paramyxoviruses and filoviruses while residues 451-739 are conserved only amongst the *filoviridae* members (Su et al., 2018; Leung et al., 2015).

The core region of EBOV NP comprises of residues 36-351 and is highly conserved amongst all *filoviridae* members. It has dimensions of 53 Å * 40 Å * 75 Å and the intramolecular interactions help in its packaging. The NP core region is significant for RNA binding and NP oligomerization for RNP formation. It can be divided into two regions viz., a) N-lobe (residues 36-240) made up of α-helices (thirteen) and β-strands (two) and b) C-lobe (residues 241-351) made up of seven α-helices. A highly positively charged groove, common to most virus NP's, is bound by these two regions. The positive charge is majorly contributed by residues 160, 171, 174 and 248 (Dong et al., 2015). This groove helps in RNA encapsidation as lysine, arginine and histidine present in the groove exhibit communication with the ribose and phosphate molecules belonging to the RNA (Ariza et al., 2013; Raymond et al., 2012; Reguera et al., 2013).

Oligomerization is carried out by residues 38-383 as they form oligomers with a back-to-back double ring configuration (Leung et al., 2015). The last two helices of each Ebola NP N-terminus molecule interact with the corresponding last two helices of neighboring NP molecules. This results in the formation of oligomeric rings (Su et al., 2018). Due to the high conservation observed in this region amongst viruses, it is thought that such molecular contacts are crucial for NP-NC interactions (Reguera et al., 2014).

2.2.2.3. Virion protein 35

EBOV VP35 is a multifunctional protein which consists of 340 amino acid residues (Grifoni et al., 2016) and has a tetrameric configuration in solution (Bruhn et al., 2017). It is functionally analogous to the phosphoprotein of other non-segmented, negative sense (NNS) RNA viruses (Mühlberger et al., 1999) and is made up of an N-terminal coiled region significant for its oligomerization and a C-terminal region which binds to dsRNA as well as helps in IFN-γ inhibition (Leung et al., 2009; Bale et al., 2013). Oligomerization domain is marked by residues 82 and 145 at either end. Residues 118 and 145 present inside the C-terminal region belonging to the oligomerization domain are immensely preserved amongst all Ebola virus species (Zinzula et al., 2018).

VP35 acts as a non-enzymatic co-factor of the L protein and thus, is crucial for viral replication (Becker et al., 1998). The homo-oligomers formed by VP35 help in genome replication as well as IFN inhibition (Reid et al., 2005). VP35 C-terminal region stops type I α/β interferon (IFN) induction by blocking normal functionality of dsRNA and via interaction with members of the retinoic acid inducible gene I (RIG-I) like receptors pathway (Zinzula and Tramontano, 2013).

Hence, as VP35 is crucial for viral replication, NC formation as well as host IFN inhibition (Zampieri et al., 2007). Lack of a functional VP35 impedes viral amplification and results in reduced viral infectivity prowess (Hartman et al., 2008).

2.2.2.4. Virion protein 24

VP24 is a secondary matrix protein consisting of 251 amino acids (UNIPROT, 1998). It makes up around 7.5% of total EBOV protein concentration and weighs approximately 24 kDa (Elliot et al., 1985). No homologs of VP24 have been indentified in NNS RNA viruses and hence, its presence is unique to *filoviridae* members (Watt et al., 2014).

VP24 weakly associates with NP and VP35 to form the nucleocapsid (NC) (Huang et al., 2002). Residues 170 and 171 are crucial for its interaction with NP (Banadyga et al., 2017). Along with VP35, it induces various conformational changes in NP required for RNA encapsidation (Huang et al., 2002). Han et al observed that a truncated VP24 results in loss of formation of a viable NC, thus, confirming the role of VP24 in NC formation (Han et al., 2003).

Like VP35, VP24 helps in host innate immune response inhibition but via different mechanisms (Ilinykh et al., 2015). Interaction of VP24 with karyopherin nuclear transporters arrests movement of STAT1 to the nucleus (Mateo et al., 2010). VP24 also remoulds the translocation of hnRNP C1/C2 (a heterogenous protein complex involved in interaction with karyopherin) by binding to it (Shabman et al., 2011). p38 is responsible for phosphorylation of various transcription factors responsible for mediating IFN response. VP24 blocks phosphorylation of p38, essentially blocking the p38 function and hence, indirectly, inhibiting IFN response (Halfmann et al., 2011; Uddin et al., 1999). Further, it also results in blockage of the activation of nuclear factor-kappa B (NF-kB) following stimulation by tumour necrosis factor alpha (TNF-α) which again leads to IFN response inhibition (Guito et al., 2017).

2.2.2.5. Virion protein 30

VP30 is a phosphoprotein consisting of 288 amino acids and weighing approximately 30 kDa (Modrof et al., 2002). It lends filoviruses distinction amongst the mononegaviruses as it is associated with NC only in the former (Kirchdoerfer et al., 2016) via direct binding with other NC components viz., NP and L (Kirchdoerfer et al., 2015). Residues 179, 180, 183 and 197 are crucial for this interaction. A study suggested that the VP30 C-terminal domain (CTD) is enough on its own for VP30-NC interaction (Hartlieb et al., 2007). Another report indicated that N-terminal phosphorylation regulates this interaction and effects viral transcription and RNA replication (Modrof et al., 2001).

EBOV transcription depends significantly on VP30 as it is essential for viral mRNA synthesis (Muhlberger et al., 1999; Hartlieb et al., 2007). The N-terminal region consists of a RNA-binding site (Modrof et al., 2001). Interestingly, a weakly or partially phosphorylated VP30 acts as a transcription factor while a fully phosphorylated VP30 acts as a transcription inhibitor (Modrof et al., 2002). The N-terminal region also consists of a zinc binding site. Zinc binding enhances the abilities of VP30 as a transcription cofactor (Modrof et al., 2003).

Another role of VP30 in regulation of co-transcriptional editing of glycoprotein mRNAs has also been reported (Martinez et al., 2008).

2.2.2.6. Virion protein 40

VP40 is made up of 326 amino acids (Madara et al., 2015), weighs around 40 kDa and is the most expressed protein in filoviruses (Adu-Gyamfi et al., 2012). It is located at multiple intracellular regions and can take up multiple oligomeric confirmations such as plasma-membrane localized hexameric confirmation, perinnuclear octameric ring confirmation etc. The N-terminal region of VP40 is credited with formation of a dimer in the cytoplasm where as the C-terminal region is accountable for the communication between the protein and the plasma membrane (Bornholdt et al., 2013).

VP40 coordinates with the viral cellular components and helps the virus assemble at the plasma membrane (Adu-Gyamfi et al., 2012). Residues 213, 295 and 298 are a part of the C-terminal domain and play a significant part in deep penetration of hexameric VP40 inside the the plasma membrane, an occurrence significant for building viral matrix, virion assembly and egress. Also, residues 221, 224, 225, 270, 274 and 275 are thought to be critical for VP40 interaction with plasma membrane. The importance of these interactions in the larger picture presents them as plausible therapeutic targets for deranging viral budding (Madara et al., 2015).

2.2.2.7. RNA polymerase

L is the largest EBOV protein and is made up of 2,212 amino acids (Ayub and Waheed, 2016). It weighs around 253 kDa (Volchkov et al., 1999). Unable to function independently, it performs as a part of RNA-dependent RNA polymerase (RdRp) complex with NP, VP30 and VP35 as the other members (Ayub and Waheed, 2016). Residues 1-505 are self-sufficient for binding to VP35 (Prins et al., 2010). Interestingly, VP30 is a dispensable member of the RdRp complex for viral replication (Trunschke et al., 2013).

The L protein is thought to perform a variety of functions including transcription and replication (mRNA guanylyltransferase, poly (A) synthetase activity etc) (Poch et al., 1990). Editing of mRNA by L is the root cause of GP protein production as this editing deviates from the normal sGP production and hence, mRNA editing by L regulates the expression of GP, sGP and ssGP (Volchkov et al., 1995).

Despite the fact that L protein is present at the heart of essential viral mechanisms, little has been deduced about the structure or functioning of this protein owing to its large size and

unavailability of specific antibodies against it (Cantoni and Rossman, 2018; Schmidt and Hoenen, 2017).

2.2.3. Life cycle of Ebola virus

The different phases of EBOV life cycle have been portrayed in the Figure 2.5, page 14.

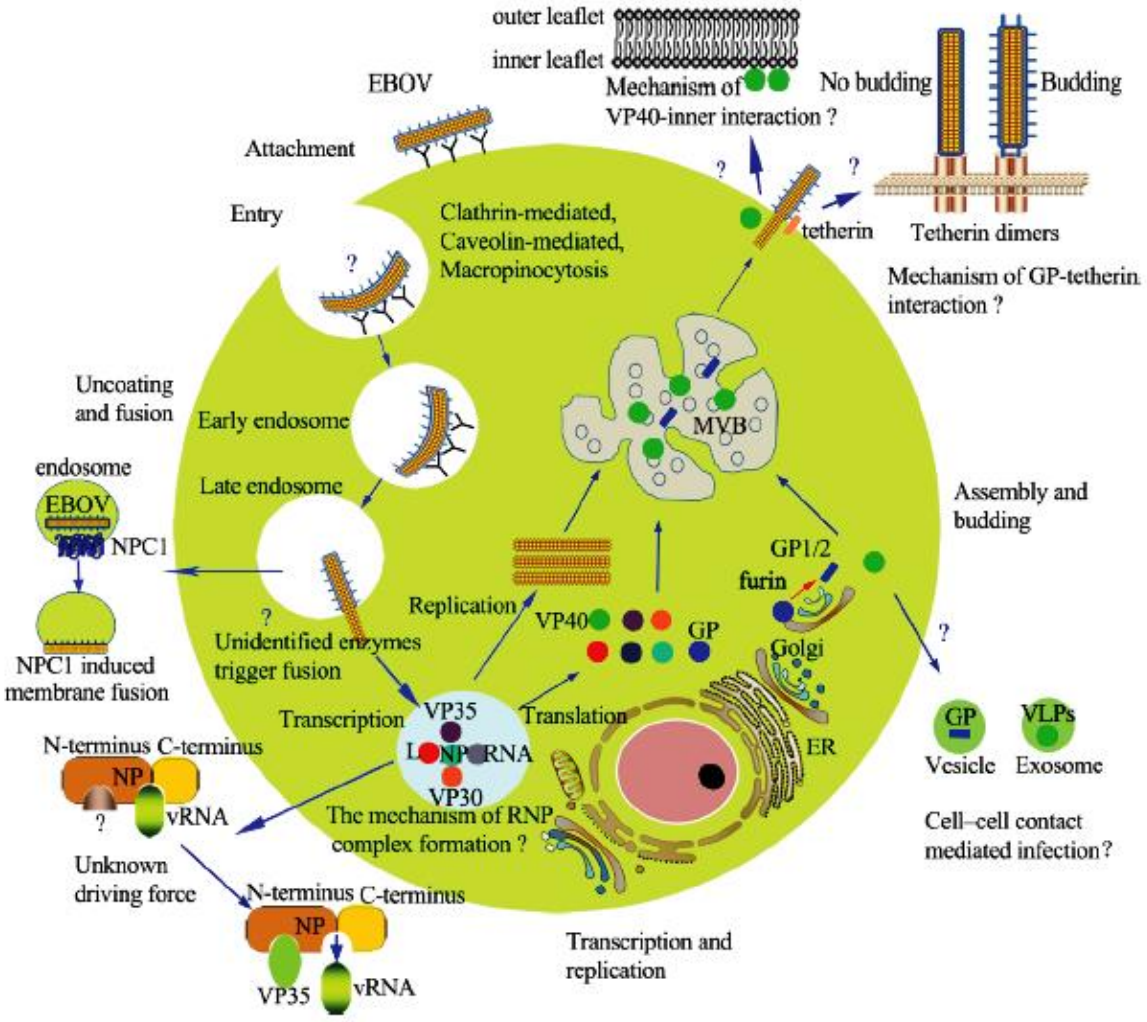


Figure 2.5: Ebola virus life cycle (Yu et al., 2017)

2.2.3.1. Viral entry (attachment and uptake)

Various stages of EBOV entry comprise of binding, internalization, trafficking to and fusion with a late endosome membrane (Figure 2.5, page 14). Currently, the mechanism of EBOV entry is partially known (Yu et al., 2017). Trimeric GP spikes present all around the EBOV surface mediate all stages of EBOV entry. EBOV GP cleavage within the golgi apparatus

results in GP1 and GP2 domains which remain in contact via non-covalent interactions (Lee and Saphire, 2009). But unlike haemagglutinin of Influenza virus, the site of EBOV GP cleavage is important for neither viral entry (Wool-Lewis and Bates, 1999) nor pathogenesis (Neumann et al., 2007). EBOV GP1 can be subdivided into three regions viz., a) The receptor binding domain (RBD) b) The glycan cap and c) Mucin-like domain (MLD). RBD affects the EBOV coupling to host cells by interacting with multiple cellular receptors (Chan et al., 2001; Kuhn et al., 2006).

Amongst the multiple receptors, C-type lectin family is prominent with its various members such as asialoglycoprotein receptor (ASGP-R), acetylgalactosamine-specific C-type lectin (hMGL) etc being linked to EBOV attachment (Yu et al., 2017). Tyro3 protein kinase (TAM) family (Hunt et al., 2011) and folate receptor-α (FR-α) (Chan et al., 2001) have also been reported to effect viral attachment. Tyrosine kinase receptor Axl (Hunt et al., 2011) mediates viral binding while presence of plasma membrane sphingomyelin and activity of sphingomyelinase have been reported to be crucial to viral binding efficiency (Miller et al., 2012).

Various endocytic pathways lead to viral uptake after its entry. A study reported that macropinocytosis is essential to the same where as clathrin and caveolin-mediated endocytosis is non-significant (Nanbo et al., 2010). In another study, cathepsin B/L (endocytic enzyme) and cholesterol (lipid-raft component) were thought to be responsible for viral uptake (Saeed et al., 2010). GP-mediated fusion is independent of acidic condition though low pH conditions do catalyse the process. GTPases, such as RhoB, Rac1, and CDC42 play a vital part in EBOV GP-dependent transduction (Bavari et al., 2002).

2.2.3.2. Uncoating and fusion

Endocytosis precedes uncoating and fusion between viral and endosomal membranes (Yu et al., 2017). Internalization of the virus into a macropinosome is followed by its transfer to an endosomal compartment consisting of cysteine proteases such as cathepsin B and cathepsin L. The latter helps in digestion of GP which initiates viral and endosomal membrane fusion (White and Schornberg, 2012). Processed GP1 interacts with Neimann-Pick C1 protein of the late endosome resulting in GP2-dependent coupling of the envelope belonging to the virus and membrane of the endosome (Carette et al., 2011). It has been found that absence of Two-

Pore Channel 2 (TPC2) has an adverse effect on the fusion process (Sakurai et al., 2015). The operational workflow of TPC2 is still unknown (Simmons et al., 2016).

Furin cleaves the EBOV GP into GP1, GP2, sGP and ssGP inside the golgi apparatus (Falzarano et al., 2007). GP2 consists of five domains viz., a) N-terminal heptad region b) C-terminal heptad region c) Transmembrane region d) Fusion loop consisting of a core 16 residue hydrophobic sequence and e) Cytoplasmic tail (Lee and Saphire, 2009). The fusion loop helps in initiation of membrane fusion process by inserting into host endosomal membranes (Bar et al., 2006). Lysosomal thiol reductase is thought to trigger the fusion process. It can be blocked with the help of cysteine protease determents and acidic pH conditions (Schornberg et al., 2006). Amongst various GTPases that play a crucial role in viral fusion, Rab7 GTPase, related to the late endosomes, is at the forefront (Simpson et al., 2012).

After fusion, trimeric HR regions of GP2 subunit result in the genesis of a hexa-helical transmembrane structure which sets off an opening through the membrane (Malashkevich et al., 1999). Further, the viral proteins are discharged into the host cytoplasm where they participate in the replication process.

2.2.3.3. Transcription and replication

EBOV RNA (encapsulated by NP) and RNA-dependent RNA polymerase (RdRp) form a complex together known as ribonucleoprotein (RNP) complex (Yu et al., 2017). This complex is released in the host cytoplasm and serves as a template for replication. Throughout the replication process, complementary positive stranded RNA (cRNA) genesis is observed in the way of RNP which further serves as template or is packaged as genomic RNA into virions (Ruigrok et al., 2011). This process stresses on the significance of correct RNP formation, a mechanism whose knowledge still eludes the scientific community (Yu et al., 2017).

N-lobe and C-lobe present in the core region of EBOV nucleoprotein result in the formation of a RNA binding groove (Ruigrok et al., 2011). Both the terminals of EBOV NP extend towards this groove and help in RNA binding as well as NP oligomerization for RNP formation (Zhou et al., 2013). VP35 blocks IRF3 signalling which inhibits IFN Type I production, thus, clearing the passage for smooth viral replication (Forrester, 2018). *In vitro* experiments have shown that VP35 N-terminal peptide performs an essential role in release

of RNA from the RNP by coupling with a hydrophobic region present on the C-terminal of nucleoprotein. This also results in inhibition of NP oligomerization (Dong et al., 2015). The exact mechanism by which VP35 carries out these functions remains largely unknown.

2.2.3.4. Assembly and budding

Various EBOV proteins play different roles throughout the viral assembly and budding process (Figure 2.5, page 14).

Nucleocapsid (NP, VP35 and VP24) formation and accumulation at the perinuclear region triggers the viral assembly process (Yu et al., 2017). NC migrates to cell surface with the help of microtubules supervised by VP40 and is lastly assimilated into virions by NP-VP40 interaction (Baker et al., 2016). Moreover, flexible NP-NP interactions during NP oligomerization enable dense packing of RNP into the viral particles (Dziubanska et al., 2014).

GP precursor is synthesized in the endoplasmic reticulum. It follows the classical secretory pathway wherein it arrives at the plasma membrane via golgi apparatus. GP experiences several post-translational modifications such as acylation before undergoing proteolytic cleavage by furin while being inside golgi apparatus (Ito et al., 2001). Lastly, GP is carried to late endosomes where it communicates with VP40 for assembly and budding (Neil et al., 2008).

VP40 protein is made up of two domains viz., a) N-terminal domain and b) C-terminal domain (Bornholdt et al., 2013). The former has a surface which is hydrophobic and results in the genesis of homodimers by interacting with various cell proteins. This results in the shipping of VP40 dimers towards plasma membrane (Dessen et al., 2000). The C-terminal domain exists as a pliant, hydrophobic loop whereas the N-terminal domain carries out the process of oligomerization (Adu-Gyamfi et al., 2015). The interaction of VP40 and the interior surface of the plasma membrane is vital for the budding process to ensue. Both, electrostatic and hydrophobic components, present in VP40 interact with the phosphatidylserine of the plasma membrane. Phosphatidylserine supervises the localization and oligomerization of VP40 onto the inner leaflet of the plasma membrane (Soni and Stahelin, 2014). Further information regarding the mechanism of VP40 association and the budding process is still unclear.

The role of EBOV VP24 in assembly and budding process is yet to be realised. A decrease in the quantity of virions released via budding was reported upon silencing of VP24 gene though viral transcription and replication remained unaffected (Huang et al., 2002).

Lastly, all the viral proteins and the ribonucleoprotein complex (RNA plus nucleoprotein (NP)) assemble at the plasma membrane and result in genesis of virions which bud out from the cell surface. sGP and ssGP are also secreted out along with the virions (White and Schornberg, 2012).

2.3. Epidemiology of Ebola

In the last 45 years, EBOV outbreaks of different proportions and varying lethality have been witnessed in various geographical regions (Table 2.1, page 20-22). After initial zoonotic infection, spread to humans has been documented several times for three EBOV species (*Sudan ebolavirus*, *Zaire ebolavirus* and *Bundibugyo ebolavirus*) (King et al., 2012). Only one human infection case has been registered for *Taï Forest ebolavirus* species so far (Le Guenno et al., 1995) while the remaining two species (*Reston ebolavirus* and *Bombali ebolavirus*) are not known to infect humans.

Even though many candidates have been contemplated as EBOV zoonotic reservoirs, it remains a topic of debate (Peterson et al., 2004). EBOV infection has been reported in various bat species such as *Hypsignathus monstrosus*, *Myonycteris torquata* and *Epomops franqueti* (Leroy et al., 2005) and a few human infections have been associated with bat exposure (Leroy et al., 2009). Dead end hosts such as chimpanzees and gorillas are also thought to be potent candidates for transfer of EBOV infection from animals to humans (Bermejo et al., 2006; Formenty et al., 1999).

Secondary transmission results after primary infection upon close contact with either the dead bodies, the infected individuals or their body fluids (Feldmann et al., 2011). The source of human infection has been difficult to identify in many cases (Nkoghe et al., 2011). Activities such as hunting and bushmeat trading might bring humans in contact with an infected animal reservoir and cause problems (Nkoghe et al., 2011). Long-established and customary funeral habits which include contact with corpse and corpse washing further complicate the matter. Rapid spread of the disease in terms of both, space and time, is observed in case of lack of optimum protective measures in heath care facilities (Georges et al., 1999).

EBOV first surfaced on June 27, 1976 in the Democratic Republic of the Sudan where it infected a man living in a rural area. The patient died within two weeks of contracting the disease. 23% people who touched the man and 81% of health governors became infected (WHO report, 1978). Overall, 284 people were reported to be infected in this outbreak and around 150 people died.

The world witnessed the second EBOV epidemic in September, 1976 in northeast Zaire. Many hospital personnel died upon transmission of disease from victims. 55 villages had reports of EBOV infection. An estimated 20% people coming in close contact of infected individuals were thought to contract the disease. Middle aged women (17-31 years) were the worst affected age group. A total of 318 cases were reported with around 88% mortality (WHO report, 1978).

First infection by *Reston ebolavirus* was recorded in 1989 in *Macaca fascicularis*, a monkey species imported from Philippine Islands to Reston. As of now, no human cases caused by this species have been reported (Dhama et al., 2018). Taï Forest *ebolavirus* first surfaced in 1994 when it infected an ethnologist who was carrying out an autopsy on a chimpanzee found dead in Côte d'Ivoire (Le Guenno et al., 1995). The fifth EBOV species, Bundibugyo *ebolavirus* came to limelight in 2007 in Uganda where it infected 116 people with 26% mortality (Towner et al., 2008).

The West African epidemic of 2013-2014 is the largest recorded EBOV outbreak till date. It is thought to have started in December, 2013 from a child in Guinea who came in contact with an infected bat (Murray, 2015). Amongst various reasons responsible for rapid spread of this disease such as resource limitation and lack of proper health care facilities, fear factor played a significant role. Lack of awareness among the locals, including the urban population, led them to think that health care workers were responsible for the spread of disease. Consequently, unwarranted attacks on medical staff led to several deaths (McMahon et al., 2016).

The latest EBOV outbreak began on August 1, 2018 in DRC in North Kivu and Ituri provinces. Response to the outbreak was impeded due to conflict between locals and the medical staff (Petesch, 2018). As a result, a secondary wave of EBOV attack was witness in September, 2018 (Anna et al., 2018). Nearly 700 cases have been reported so far (WHO report, 2019). Till date, more than 25 outbreaks have been registered with Zaire and Sudan species being responsible for most of them (Tariq et al., 2019) (Table 2.1, page 20-22).

Table 2.l: All human Ebola outbreaks since the inception of Ebola virus disease (CDC, 2019) a) From 1976-2000 b) From 2000-2013 c) 2014 onwards.

a) 1976-2000

Year	Area/Locality	Ebola species	Number of cases	Number of deaths	Reference
1976	England	Sudan	1	0	Emond et al., 1977
1976	Sudan	Sudan	284	151	WHO report, 1978
1976	Democratic republic of the Congo (DRC)	Zaire	318	280	WHO report, 1978
1977	DRC	Zaire	1	1	Heymann et al., 1980
1979	Sudan	Sudan	34	22	Baron et al., 1983
1994	Cote D'Ivoire	Taï Forest	1	0	Le Guenno et al., 1995
1994	Gabon	Zaire	52	31	Georges et al., 1999
1995	DRC	Zaire	315	250	Khan et al., 1999
1996	Russia	Zaire	1	1	Borisevich et al., 2006
1996	South Africa	Zaire	2	1	WHO records, 1996
1996	Gabon	Zaire	97	66	Georges et al., 1999
2000	Uganda	Sudan	425	224	Okware et al., 2002

b) 2000-2013

Year	Area/Locality	Ebola species	Number of cases	Number of deaths	Reference
2001	Gabon	Zaire	65	53	WHO report, 2003
2002	DRC*	Zaire	143	128	Formenty et al., 2003
2003	DRC	Zaire	35	29	WHO record, 2004
2004	Russia	Zaire	1	1	Akinfeyeva et al., 2005
2004	Sudan	Sudan	17	7	WHO record, 2005
2005	DRC	Zaire	12	10	CDC, 2019
2007	Uganda	Bundibugyo	131	42	MacNeil et al., 2011
2007	DRC	Zaire	264	187	WHO record, 2007
2008	DRC	Zaire	32	15	WHO, 2009
2011	Uganda	Sudan	1	1	Shoemaker et al., 2012
2012	Uganda	Sudan	17	7	Albarino et al., 2013
2012	DRC	Bundibugyo	36	13	Albarino et al., 2013

* DRC stands for Democratic republic of the Congo

c) 2014 onwards

Year	Area/Locality	Ebola species	Number of cases	Number of deaths	Reference
2014	DRC*	Zaire	69	49	Maganga et al., 2014
2014	Guinea, Liberia, Sierra Leone	Zaire	28,610	11,308	Bell et al., 2016
2014	Italy	Zaire	1	0	WHO, 2015
2014	Mali	Zaire	8	6	Bell et al., 2016
2014	Nigeria	Zaire	20	8	Bell et al., 2016
2014	Senegal	Zaire	1	0	Bell et al., 2016
2014	Spain	Zaire	1	0	WHO, 2014.
2014	United States	Zaire	4	1	Bell et al., 2016
2017	DRC	Zaire	8	4	CDC, 2019
2018	DRC	Zaire	54	33	CDC, 2019

* DRC stands for Democratic republic of the Congo

2.4. Ebola and immune system

2.4.1. Innate immune response

Innate as well as adaptive (humoral and cellular) immunity is triggered at every step of EBOV infection (Marcinkiewicz et al., 2014).

Innate responses against viral infections are mediated by Type I IFNs (Ivashkiv and Donlin, 2014). Multiple IFN-α and a single IFN-β gene encodes for Type I IFNs protein family. Activation of various pattern recognition receptor pathways can lead to IFN–α/β gene expression. Examples of such pathways include the RIG-I-like receptor (RLR) pathways, STING/cGAS pathway etc. (Brubaker et al., 2015). Two RIG-I-like receptors, RIG-I and MDA5, have been reported to be relevant in case of filoviral infections (Basler, 2015). Upon activation, cells secrete IFN–α/β which couples with the IFN-α/β receptor on the cell surface ensuing in autocrine or paracrine signalling. Consequently, a JAK-STAT signaling cascade develops which activates several hundred genes responsible for conferring resistance to viral infections and viral replication blockage ability to cells (Basler, 2015) (Figure 2.6, page 23).

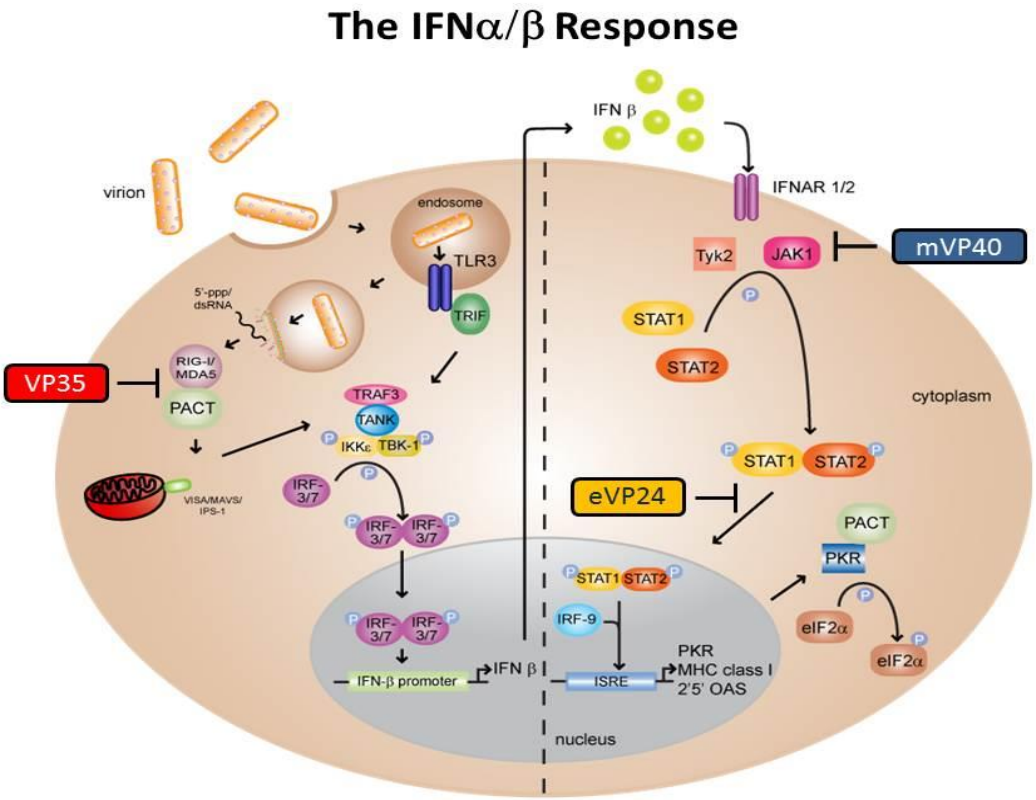


Figure 2.6: Innate immune response generated against Ebola virus (ICAHN, 2019)

IFN-γ or Type II IFN is vital for macrophage activation and expression of HLA class I molecules (Ben-Asouli et al., 2002). It can also trigger type I IFN response via cross talk (Barkhouse et al., 2015). IFN-λ or Type III IFN constitutes of IL-28A, IL-28B and IL-29. They are secreted by diverse human cells, especially respiratory epithelial cells (Wei et al., 2016).

2.4.2. Adaptive immune response

Different haemorrhagic fever viruses trigger varied responses by T cell priming and activation, thus, resulting in the activation of adaptive immunity (Figure 2.7, page 25). EBOV infection results in haemorrhagic fever marked by onset of type-I effector mechanisms (cytotoxicity and interferon production), tremendous T cell activation and differentiation to effector profiles (Perdomo-celis et al., 2019).

Protection against viral haemorrhagic fevers (VHF), whether natural or vaccine mediated, requires the stimulation of T_c as well as T_h cells mediated by immunodominant epitopes. Activation of lytic and non-lytic effector mechanisms of both cell types is also essential. T_h cells perform numerous other functions such as enhancement of T_C response (Wherry and Ahmed, 2004) and B cell activation which results in antibody production (Farooq et al., 2016) and hence, are indispensable for a robust immune response generation. Prodigious activation and terminal differentiation of T_C cells has been reported in human survivors of EBOV disease. Unusually high numbers of HLA-DR^+ $CD38^+$, Ki-67, $CD45RA^+$ $CCR7^-$ and $CD45RA^+$ $CCR7^-$ cells were detected even after two months of viral infection (McElroy et al., 2015). In another study, no change in T_h cell count was observed (an observation made during severe DENV infection as well) and this has been contemplated as a reason for the tremendous T cell activation and disease severity (Jayaratne et al., 2018).

Cytokine circulation follows a specific pattern in case of VHF. Massive cytokine production is a trademark of VHF and is known as “cytokine storm” (Khaiboullina et al., 2017). Cells primarily involved in production of these factors include endothelial cells, antigen-presenting cells, NK cells, polymorphonuclear cells etc (Yiu et al., 2012). An increase in IFN-α, IFN-β, TNF-α and IL-6 (Hutchinson and Rollin, 2007) is observed during acute phase of infection. This phase, witnessed between 1^{st} - 4^{th} days after symptom detection, is also characterized by development of fever. IL-2 (Rathakrishnan et al., 2012) and IL-10 (Hutchinson and Rollin, 2007) levels peak between 5^{th} – 10^{th} days of infection, a period which reflects the worst

consequences of VHF. Other mediators/factors observed include lipid mediators (Bennett and Gilroy, 2016) and oxygen and nitrogen reactive species (Miller et al., 2011). Together, these factors result in T cell activation and delimitation into effector profiles by causing their polarization (Kaiko et al., 2008).

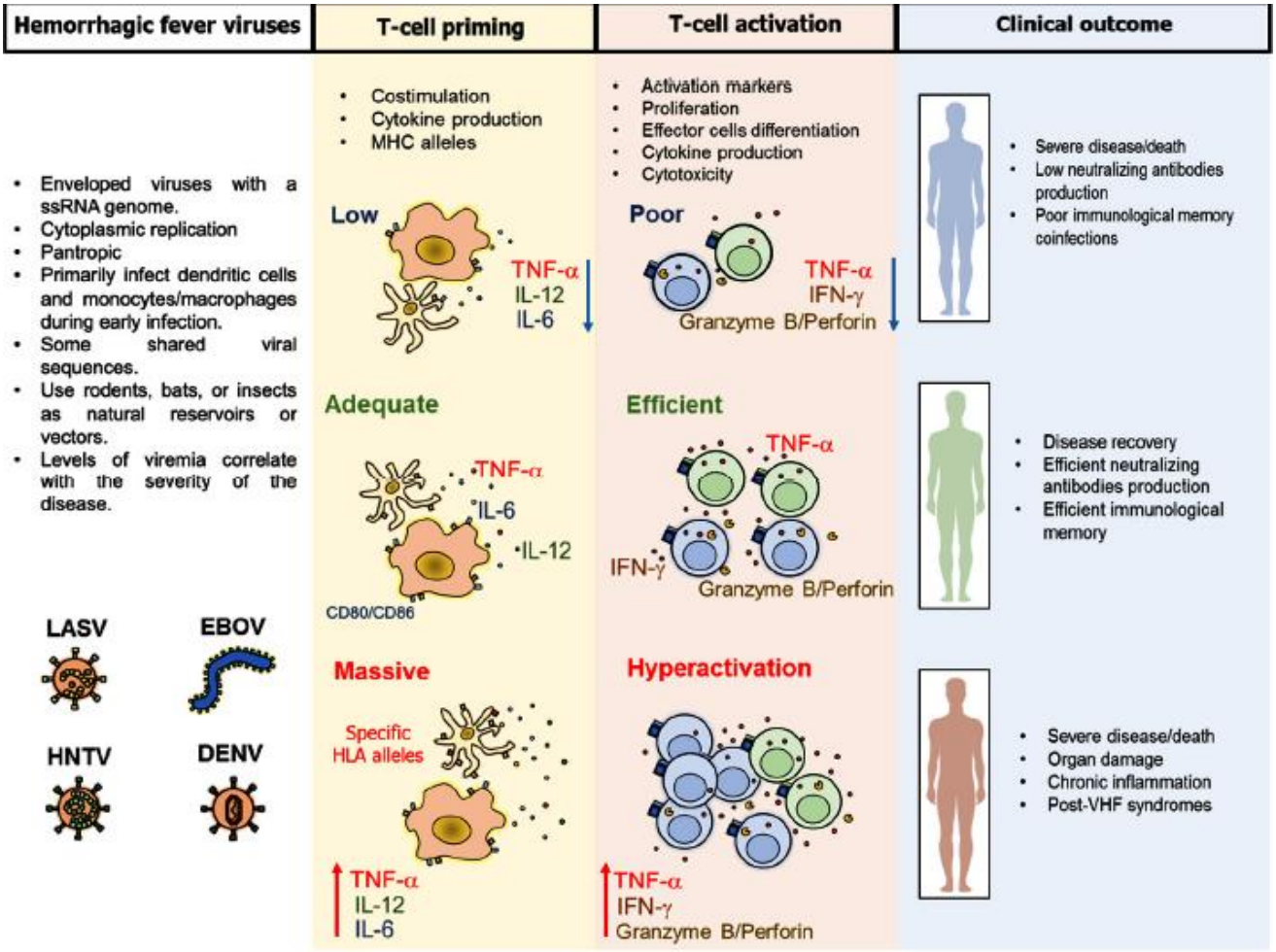


Figure 2.7: Activation of adaptive immune response by various haemorrhagic fever viruses (Perdomo-Celis et al., 2019)

2.4.3. Evasion of host immune response by Ebola virus

Filoviridae members employ an array of mechanisms to effectively counter the host protective mechanisms. Table 2.2 enlists the various ways of host immune evasion observed during different studies. Antiviral defences such as IFN responses are counteracted by suppressing Type I IFN and obstructing IFN signalling. This is indicated by the rapid replication rate displayed by these viruses inside the host (Bray and Geisbert, 2005).

EBOV VP35 plays a central role in host immune response evasion. Not only does it impede the normal functioning of both IFN's (α and β) (Basler, 2015), it also blocks the activation of IFN-induced kinase PKR which exhibits antiviral activity (Schumann et al., 2009). Interferon regulatory factor 3 (IRF-3) is a transcription factor vital for IFN-β promoter induction. VP35 prevents phosphorylation of this factor (Basler and Amarasinghe, 2009). In 2006, an interesting correlation was found by Cardenas et al within VP35's ability to bind to dsRNA and to restrict RIG-I signalling (Cardenas et al., 2006).

The suppression of innate immune response by EBOV VP35 has far reaching consequences. It impairs dendritic cell maturation which results in inhibition of IFN-α and IFN-β production as well as a failure of dendritic cells to activate T cells (Basler, 2015).

Mohan and coworkers discovered in mice that sGP acts as a roadblock between the virus GP and anti-GP 12 antibodies as it can efficiently compete for these antibodies. Hence, it does not allow the antibodies and the virus GP to come in contact. It concentrates the attention of host onto its epitopes which are identical to those on virus GP and hence, absorbs anti-GP1,2 antibodies. This event is known as antigenic subversion (Mohan et al., 2012). Contrary to the popular belief that viruses reduce the number of immune recognition molecules present on the victim's cell surface, Francica and colleagues found that virus GP prevents the recognition of HLA class I and II molecules, a phenomenon termed as steric occlusion. Not only does this occur for host cell, steric occlusion also prevents the recognition of Ebola virus GP itself (Francica et al., 2010). Hence, steric occlusion is another way in which viruses can elude host immunity.

Further, GP1 glycosylation, although not required for viral entry, is thought to guard GP from immune system, thus, enabling GP to function normally (Collar et al., 2017).

Table 2.2: Various mechanisms employed by Ebola virus to evade host immune response

Mechanism	Description	Reference
Antigenic variation	Mucin like domain of GP varies between species and forms a part of the GP spike after cleavage of GP to GP1-GP2. This lends the virus protection against antibodies as cross-protection between viral species is nullified	Hevey et al., 1997
Abundant N and O linked glycosylation	Present in mucin-like domain of GP and leads to viral tropism for dendritic cells and monocytes. Also results in epitope masking and variations in phenotypes.	Jeffers et al., 2002
GP gene editing	Amongst filoviruses, this phenomenon is observed only in EBOV. A truncated, frame-shifted version of sGP is produced in this case.	Sanchez et al., 1996
Enshrouding of receptor-binding and fusion domains	The effect of antibodies is reduced drastically and probably completely upon exposure of vital functional domains only transiently upon entry.	Manicassamy et al., 2005
Promiscuous binding to cells	GP has the ability to bind with different cell receptors. This proves to be problematic for antibodies as they are unable to block so many interactions.	Feldmann et al., 1994
Cytokine release	GP reigns havoc on the immune system by activating multiple immune measures such as intracellular cascades, granule and cytokine release leading to an unbalanced immune response	Leroy et al., 2000

2.5. Ebola: Prevention and control

2.5.1. Antiviral drugs

It takes at least three weeks for a vaccine to stimulate immunity in a person post-exposure to a pathogen. Owing to this and the consistent EBOV threat increase and rapid transmission of the disease, post-exposure therapeutic drug development is in order. The need was further stressed in 2014 when during the largest EBOV outbreak till date, WHO permitted the use of untested drugs to somehow control the high death count (Jerebtsova and Nekhai, 2015).

A summary of various works done with respect to therapeutic drug development has been presented in table 2.3, page 30. Common approaches include targeting viral proteins and restriction of viral entry. The latter can be achieved in two ways viz., disrupting normal functioning of proteins crucial to the process and/or blocking cell receptors accountable for attachment of virus to host cell (Jerebtsova and Nekhai, 2015).

Pradimicin A and benanomicin A are non-peptidic antibiotics that target EBOV GP (Balzarini, 2007). Enzon Pharmaceuticals provides commercial-grade human mannose-binding lectin (MBL), a C-type lectin that identifies various EBOV surface glycan structures, such as glucose, mannose etc. Regrettably, complex MBL quaternary structure makes its large-scale production cumbersome and expensive (Jerebtsova and Nekhai, 2015). Griffithsin (GRFT) is a red algae derived lectin which is devoid of mitogenic activity and interacts with the terminal mannose molecules of N-linked glycans on the outside of EBOV. In contrast to many lectins, GRFT doesn't actuate proinflammatory cytokine production in human PBMC. Periodic subcutaneous dosages of GRFT exhibited an insignificant danger in two rodent species. Further validation of GRFT is being done (Barton et al., 2014).

BioCryst Pharmaceuticals developed a viral RNA polymerase inhibitor, adenosine analog (nucleoside drug) named BCX4430 which functions as a non-obligate RNA-chain terminator. Upon being injected in the subject 48h post infection, BCX4430 successfully stopped the flow of viral infection in human cells and non human primates. Dismal pharmacokinetic properties and short half life (<5 min) are the major drawbacks of this drug (Warren et al., 2014). Another nucleoside drug capable of viral RNA synthesis inhibition is named Favipiravir. It can be obtained via chemical modification of pyrazine (Furuta et al., 2017).

Antibodies are known to inhibit viral entry into cells as well as encourage antibody-directed cell-mediated cytotoxicity by NK cells. Also, along with the complement system, antibodies help in neutralization of the virus (Keller and Stiehm, 2000). Polyclonal IgG antibody

obtained from EBOV survivors and injected into non-human primates two days post-exposure protected nearly 65% of test individuals (Dye et al., 2012). Though polyclonal antibodies confer better protection when compared to monoclonal antibodies, production difficulties, high lethality and low survivor numbers are the major limitations of polyclonal antibodies that render them unsuitable for human usage. This led the scientists to look at recombinant mAbs as a plausible alternative for post-exposure antibody treatment (Jerebtsova and Nekhai, 2015).

Failure of human monoclonal antibody KZ52 in securing non-human primates from EBOV infection even when injected 1 day prior to infection (Oswald et al., 2007) paved the pathway for development of ZMapp, an antibody cocktail manufactured in collaboration by US Mapp Biopharmaceutical, Inc. and Defyrus, Inc. This cocktail conferred protection in all test non-human primates when injected five days post-exposure (Qiu et al., 2014) though efficacy in humans is still under debate (Dixon and Schafer, 2014). In another case, Russian federation approved the temporary treatment of EBOV infected humans with hyperimmune serum (containing immunoglobulins) obtained from EBOV infected goats as the serum proved to be successful in tests on various human volunteers (Kudoyarova-Zubavichene et al., 1999).

Therefore, mAb-based therapies in non-human primates have proven their metal at choking EBOV spread and reducing the fatality rate (Qiu and Kobinger, 2014). Still, the phenomenal measures employed by filovirus to evade the antibody based treatments (Kajihara et al., 2013) present a challenge to the scientific community, especially to counter EBOV infection in humans (Jerebtsova and Nekhai, 2015).

Table 2.3: Various therapeutic drug development experiments performed in different studies

Drug Name	Target	Description	Reference
Mannose-binding lectin (MBL)	GP	Post infection treatment protects 40% mice from lethal injection with EBOV	Michelow et al., 2011
Commercial-grade MBL	GP	Chimeric fusion protein L-FCN/MBL76 reduced infection	Michelow et al., 2010
Griffithsin (GRFT)	Glycan structures	GRFT binds N-linked glycans present on the viral surface.	Barton et al., 2014
BCX4430	RNA polymerase	Protects NHP from MARV; Obstructs RNA polymerase	Warren et al., 2014
T-705 (Favipiravir)	RNA polymerase	Acts as a nucleotide analog. Treatment at day 6 postinfection protects 100% of mice	Oestereich et al., 2014
Aptamers	VP35	Bind to VP35 protein and disrupt its interaction with NP	Binning et al., 2013
AVI-6002 (PMO)	VP24 and VP35	Postinfection treatment protects 60% NHP	Iversen et al., 2012
TKM-Ebola	RNA polymerase, VP24 and VP35	Combination of siRNAs in lipid particles protects guinea pigs and monkeys	Geisbert et al., 2010
Small molecule inhibitor of VP40	VP40	Inhibited Nedd4–PPxY interaction and PPxY-dependent budding	Han et al., 2014

Currently, 80 drugs consented by the FDA are known to display anti-EBOV activity (Schneider-futschik et al., 2018) out of which 45 are cationic amphiphiles (Fan et al., 2017). Ease of availability of FDA-approved drugs makes them a reliable candidate for treatment in case of unprecedented EBOV breakouts. These drugs include antidepressants, selective estrogen receptor modulators (SERMs), antihistamines etc (Schneider-futschik et al., 2018). Amiodarone, dronedarone and verapamil drugs block viral entry effectively though their mechanism of action is unknown (Gehring et al., 2014). Positive results in *in vitro* and/or *in vivo* experiments and presence of a tertiary amine in these and other drugs such as chloroquine and amodiaquine have led to the belief of a similar action mechanism for all these drugs (Ekins and Coffee, 2015). A list of various FDA-approved drugs has been presented in the table 2.4, page 32.

Table 2.4: Food and drug administration (FDA) approved drugs known to display anti-Ebola virus activity

FDA approved drug	Description	Reference
Amiodarone	Block viral entry at concentrations used for antiarrhythmic therapy	Gehring et al., 2014
Dronedarone	Block viral entry at concentrations used for antiarrhythmic therapy	Gehring et al., 2014
Verapamil	Block viral entry at concentrations used for antiarrhythmic therapy	Gehring et al., 2014
Miglustat, acarbose and miglitol (Imino sugars)	Postinfection treatment protects 70% mice from lethal injections with EBOV and MARV	Hensley et al., 2002
IFN-beta	Increases time of survival in NHP after lethal EBOV exposure	Smith et al., 2013
Toremifene	Protects 50% of mice from lethal exposure with EBOV	Johansen et al., 2013
Clomiphene	Protects 90% of mice from lethal exposure with EBOV	Johansen et al., 2013

2.5.2. Vaccines

Despite many showing successful results against non-human primates, no vaccines have been licensed for human usage as of 2019 (Mathebula et al., 2019; Feldmann et al., 2018). In a panic driven step to counter the possible and unprecedented presence of EBOV in their countries, Russia and China licensed two EBOV vaccines (STAT, 2017) though the efficacy of these vaccines is under debate as sufficient phase III trial data is not available (PharmaIntelligence, 2019).

The first ever vaccine candidates against EBOV were tested in the 1970s and 1980s via traditional vaccine development strategy of injecting the attenuated virus in animal models (Lupton et al., 1980). Lack of efficacy and concern for safety led scientists to consider subunit and DNA vaccines in the 1990s. 21st century saw expedited efforts in designing an anti-EBOV vaccine with various vectors such as alphavirus replicons (Herbert et al., 2013; Hevey et al., 1998; Pushko et al., 2000), virus-like particles (Warfield et al., 2007), human adenoviruses (Geisbert et al., 2011; Sullivan et al., 2000; Swenson et al., 2008), a biologically contained EBOV lacking VP30 (Marzi et al., 2015), chimpanzee adenovirus (Stanley et al., 2014), DNA (Grant-Klein et al., 2015), paramyxoviruses (Bukreyev et al., 2009; Bukreyev et al., 2006), cytomegalovirus (CMV) vectors (Marzi et al., 2016), rabies virus (Blaney et al., 2013; Willet et al., 2015), modified vaccinia virus Ankara (MVA) (Domi et al., 2018) and different strategies with recombinant vesicular stomatitis virus (rVSV) (Jones et al., 2005; Matassov et al., 2015) under consideration. Many experiments with such vectors conferred protection in non-human primates (Stanley et al., 2014; Sullivan et al., 2000).

A list of various vaccines under consideration has been provided in the table 2.5, page 34-35.

Table 2.5: Various vaccines against Ebola virus under consideration and their clinical trial phase. a) Vaccines in clinical trial phase I and II and b) Vaccines in clinical trial phase III and pre-registration

a) Vaccines against Ebola virus in clinical trial phase I and II

Vaccine	Clinical trial phase	References
Adenoviral vector vaccine	I	Ledgerwood et al., 2017
INO-4212	I	ClinicalTrials, 2019
Nucleic acid vaccine	I	Martin et al., 2006
Recombinant adenoviral vector vaccine	I	Wang et al., 2017
HPIV3/EboGP vaccine	I	Meyer et al., 2015
Recombinant EBOV GP vaccine	I	Bengtsson et al., 2016
MVA-EBOZ	I	ClinicalTrials, 2017
rVSVN4CT1-EBOVGP1	I	Matassov et al., 2015
GamEvac-combi vaccine	II	Dolzhikova et al., 2017
interferon beta- 1a	II	Konde et al., 2017
ZMapp	II	Group PIW et al., 2016
SRC VB Vector vaccine	II	Konde et al., 2017

b) Vaccines against Ebola virus in clinical trial phase III and pre-registration

Vaccine	Clinical trial phase	References
rVSV-EBOV	III	Wang et al., 2017
VRCEBOADC076-00-VP	III	Martins et al., 2016
MVA-BN Filo + Ad26-ZEBOV	III	ClinicalTrials, 2017
Adenovirus based Ebola vaccine (freeze dried) PLA/Tianjin CanSino	Pre-registration	Wang et al., 2017

2.5.2.1. Virus like particles

Viral like particles (VLPs) are protein subunit–based vectors obtained by transfecting cells with EBOV GP and VP40 encoding plasmid resulting in production of particles shaped like EBOV virus (Warfield et al., 2007). Plasmids encoding other EBOV proteins as well such as NP and VP24 increase VLP production efficiency (Licata et al., 2004). These particles are non-infectious and strongly immunogenic as they are capable of inducing humoral as well as cell mediated immunity (Warfield et al., 2005). Production difficulties and economic setbacks are the major limitations of using this platform (Feldmann et al., 2018).

2.5.2.2. DNA vaccines

DNA vaccines are economic and versatile with an uncomplicated production procedure making them a great alternative to VLPs. They also result in induction of both, humoral as well as cell mediated immunity and are safe to use. Experiments with rodent models yielded encouraging results (Riemenschneider et al., 2003; Vanderzanden et al., 1998) though hundered percent efficacy in non-human primates is yet to be achieved (Grant-Klein et al., 2015). Moreover, high requirement of regular booster dosages presents a limitation in this vaccination approach (Grant-Klein et al., 2011). A combination of DNA vaccination with other approaches may lead to an optimized and efficient EBOV vaccine.

2.5.2.3. Venezuelan equine encephalitis virus replicons

Alphavirus replicons are well accepted and viable vaccine vectors. Replacement of alphavirus structural genes by a foreign antigen results in alphavirus replicons with the ability to transcribe and replicate normally but lacking the ability to make viral particles (Feldmann et al., 2018). The challenge in the use of these replicons is the need for booster dosages. Inconsistent results were obtained upon evaluation of EBOV-GP and NP expressing Venezuelan equine encephalitis virus (VEEV) replicons in non-human primates (Geisbert et al., 2002; Herbert et al., 2013). However, use of adjuvants or better delivery systems might improve vaccine efficacy.

2.5.2.4. Recombinant EBOVΔVP30

EBOV VP30 gene is essential for viral transcription. Removal of VP30 gene from EBOV genome results in a recombinant, replication incompetent and biologically contained EBOV named EBOVΔVP30 (Halfmann et al., 2008; Halfmann et al., 2009). EBOVΔVP30 vector proved to be a success upon its assessment in non-human primates. It conferred complete protection in a single dose administered using varying strategies such as prime-boost approach etc. (Marzi et al., 2015). However, safety concerns regarding the use of EBOVΔVP30 in humans remain. The inactivated prime-boost vaccine strategy might help in this aspect and hence, has emerged as the leading administration strategy.

2.5.2.5. Recombinant cytomegalovirus-based vectors

Recently, cytomegalovirus (CMV) based vectors were reported to generate desirable antibodies levels in rhesus macaques (Marzi et al., 2016). This finding was contrary to earlier studies based on CMV vectors (Hansen et al., 2011; Hansen et al., 2009) and advocated selective transgene promoter usage. A major advantage of CMV vectors is that they are not hindered by pre-existing immunity. Hence, this vaccine platform is a highly potent alternative to EBOV vaccine development for non-human primates as well as humans.

2.5.2.6. Modified vaccinia virus Ankara

In the past decade, many experiments have been conducted by utilizing modified vaccinia virus Ankara (MVA) vaccine alone or in combination with other vaccines. A few candidates have made it to human trials and are showing encouraging results. A combination of adenovirus-based vaccine and MVA based vaccine was reported to induce protection against deadly EBOV doses (Mire et al., 2016) and is undergoing clinical trials. Another MVA based candidate entering clinical trials conferred protection in non-human primates after a single dose. It has been modified to express EBOV GP and VP40 proteins leading to release of virus like particles from infected cells (Domi et al., 2018).

2.5.2.7. Recombinant human parainfluenza virus

Human parainfluenza virus 3 (HPIV3) leads to replication competent, live-attenuated vectors. These vectors have been under consideration since a long time (Durbin et al., 2000; Karron et al., 2003) but due to their ability to replicate and pre-existing immunity concerns, subject safety remains an issue and robust validation methods are required to authorize human usage. A vaccine vector based on HPIV3 and expressing EBOV-GP was shown to induce an immune response even under pre-existing immunity conditions (Bukreyev et al., 2010). In another study, a similar vaccine vector conferred protection in non-human primates upon intranasal, intratracheal and aerosol vaccination (Meyer et al., 2015). Another vector based on HPIV type 1 has been under consideration off late (Lingemann et al., 2017).

2.5.2.8. Recombinant rabies virus–based vectors

Rabies virus (RABV) infection has been observed in many regions endemically affected by EBOV infection. Hence, recombinant rabies virus based vector present an opportunity for protection against both the diseases. A point mutation in RABV glycoprotein of RABV vaccine strain SAD B19 followed by induction of ability to express EBOV-GP results in a live-attenuated, recombinant RABV based vector (rRABV/EBOV-GP) (Blaney et al., 2011). This vector has been reported to induce complete protection in non-human primates against EBOV in a single dose (Blaney et al., 2013). Contrary to these results, administration of live-attenuated rRABVΔG/EBOV-GP followed by inactivated rRABV/EBOV-GP displayed only 50% efficacy (Blaney et al., 2013). Modification of the inactivated RABV vaccine vector using a codon-optimized antigen (Willet et al., 2015) and use of adjuvants showed encouraging results in non-human primates. This strategy is now moving towards human clinical trials (Johnson et al., 2016).

2.5.2.9. Recombinant adenovirus-based vectors

Adenovirus (Ad) based vectors are replication incompetent vectors and hence, provide a safer alternative as compared to replication-competent vectors. Though adenovirus (Ad) based vectors such as human adenovirus serotype 5 (Ad5) vector have been under consideration (Marzi and Feldmann, 2014), pre-existing immunity against the vector is a major challenge in vaccine development (Mast et al., 2010) as highlighted by low vaccine efficacy in *in vivo* experiments performed by administration of Ad5-based vaccines in subjects with pre-existing

immunity (Croyle et al., 2008; Richardson et al., 2013). Interestingly, modification in the vaccine delivery system to mucosal delivery nullified the pre-existing immunity effect as well as generated better immune response (Croyle et al., 2008; Richardson et al., 2013). Administration of a DNA-prime and Ad5-boost vaccination conferred protection in non-human primates though only humoral immunity was induced (Sullivan et al., 2000). In another study, administration of high volume of Ad5 expressing EBOV-GP showed encouraging results (Sullivan et al., 2003). Clearly, pre-existing immunity and high titer requirement are two major concerns in Ad5 vaccine development. Replacement vectors such as Ad26, Ad35 (Geisbert et al., 2011), chimpanzee adenovirus (cAd) and simian Ad21 (Roy et al., 2006; Stanley et al., 2014) have displayed inconsistent results. A single dosage of cAd3 vector expressing EBOV-GP protected all the subjects (Stanley et al., 2014). Optimization and improvements in this platform are a work in progress.

2.5.2.10. Recombinant vesicular stomatitis virus–based vectors

Amongst the most successful of EBOV vaccine candidates is recombinant vesicular stomatitis virus (VSV) based vector. VSV based vaccines have been generated against EBOV as well as against Marburgvirus species (Marzi et al., 2011). VSV Indiana serotype expressing EBOV GP is used as a live-attenuated, replication incompetent vector for filovirus vaccine (Rose et al., 2001; Rose et al., 2000). VSV-GP, responsible for ingress of the virus as well as its pathogenicity (Garbutt et al., 2004; Rose et al., 2000), is removed from the vector beforehand. This virus based vaccine is highly immunogenic (Agnandji et al., 2015; Huttner et al., 2015) as proven by experiment on non-human primates who remained protected even when injected with lethal EBOV dose 28 days after a solitary dose of vaccine administration (Jones et al., 2005). Another study reported complete protection in non-human primates even on administration of the vaccine a week after EBOV infection (Marzi et al., 2015). High antibody generation has been identified as the reason for high immunogenicity (Marzi et al., 2013; Geisbert et al., 2008). Safety concerns were put to rest by testing on susceptible pigs and SIV-infected non-human primates (de Wit et al., 2015; Geisbert et al., 2008), authenticating the usage of VSV-EBOV vaccine as an emergency solution (Feldmann et al., 2018). Another vaccine candidate in this category which successfully protected non-human primates is VSV-N4CT1-ZEBOVGP. It expresses VSV-N and EBOV-GP and also carries a mutant form of VSV-G (Mire et al., 2015).

2.6. Peptide based vaccines

Even though protein and DNA vaccines cater to many challenges presented by traditional means of vaccination, peptide based vaccines have emerged as an alternative to further improve safety, specificity, selectivity and stability. The antigen presenting cells internalize (phagocytosis) and digest the pathogen leading to its proteolysis. The peptides thus obtained are presented to naive T cells via HLA molecules present on the cell surface. This leads to stimulation of cell mediated and/or humoral immunity (Neefjes et al., 2011). Therefore, antigenic peptides are responsible for generating immune response. Such peptides are known as epitopes. In other words, epitopes are the minimal requirement for recognition of a pathogen by immune system (Skwarczynski and Toth, 2016) and based on this molecular mechanism, peptide based vaccine development has picked up pace in recent times (Nevagi et al., 2018).

Betterment in gene expression understanding and reverse immunology studies coupled with progress in immunological techniques has played a significant role in peptide based vaccine development (Nevagi et al., 2018). A list of various therapeutic and prophylactic peptide based vaccines under different clinical trial stages has been presented in the table 2.6, page 41.

Table 2.6: Peptide based vaccines under consideration for various therapeutic indications and their clinical trial phase

Peptide Vaccine	Therapeutic indication	Clinical trial stage	Reference
J8	GAS (Group A streptococcus) infections	I	Steer et al., 2016
$(T_1B)_4$ MAP	Malaria	I	Nardin et al., 2000
MSP3 SLP	Malaria	II	Sirima et al., 2009; Malaria Site, 2019
AFO-18	HIV	I	Karlsson et al., 2013
Peptides derived from B cell epitopes of HA	Influenza	I	Berlanda Scorza et al., 2016
HerpV	Herpes simplex virus	II	Wald et al., 2011
E75, HER2/neu peptide	Breast cancer	II	Mittendorf et al;., 2008
E6/E7 peptides	Cervical cancer	II	Vici et al., 2016
CDCA1, CDH3 and KIF-20A	Advanced biliary tract cancer	I	Aruga et al., 2014
SurVaxM (SVN53-67/M57-KLH)	Recurrent malignant glioma	II	Fenstermaker et al., 2016
VEGFR1-770/1084 peptides	Renal cell carcinoma	I	Yoshimura et al., 2013
WT 1 Peptides	Solid malignancy	II	Ohno et al., 2012
R3 peptide	Acute myeloid leukemia	I	Schmitt et al., 2008
GV1001	Pancreatic cancer	II	Staff et al., 2014; ClinicalTrials, 2017

2.6.1. Advantages of Peptide based vaccines

Peptide-based vaccines present many advantages over traditional vaccines. Peptides can be acquired in profoundly unadulterated structure and are effectively characterized utilizing analytical techniques, for example, fluid chromatography and mass spectrometry. They can be lyophilized into powder form, removing cold chain requirement amid transport and storage. Peptide size and properties can be customized without much difficulty to produce the ideal immune response. Also, peptides can be covalently conjugated to adjuvants or protein carriers in a highly desirable and controllable manner to attune their immunogenicity.

Peptide based vaccines are safer to produce as handling of dangerous pathogens can be avoided. Also, they help in nullifying autoimmune or allergic response to vaccination (Skwarczynski and Toth, 2016). Some immunogenic regions (epitopes) take the shelter of other epitopes and hence, are able to evade the immune system. Such epitopes are called as cryptic epitopes. Peptide based vaccines help to overcome such immune evasion by directing humoral immunity against cryptic epitopes as well. Moreover, peptide-based vaccines potentially effective against multiple pathogens can be created by amalgamating immunogenic stretches of peptides obtained from their proteins.

2.7. Immunoinformatics

Computational immunology or immunoinformatics is a branch of science that addresses the organisation of massive, raw immunological data with the help of computational approaches into an understandable form, enabling the researcher to generate meaningful interpretations (Korber et al., 2006). It is the bridge that links mathematical and computational model results to experimental validation, thus, preventing the tiresome and daunting routine of validating thousands of candidates experimentally (Bui and Steiner, 2016; Fradera and Babaoglu, 2017). Success in protein and genome sequencing laid the foundations of this field. *In silico* approaches have far-ranging applications in the study of evolutionary patterns (Singh and Pardasani, 2009), structure prediction (Singh, 2004), host-pathogen interactions, peptide-HLA interactions, epitope prediction, therapeutic and diagnostic technique representation and vaccine (both, prophylactic and therapeutic) design (Backert and Kohlbacher, 2015). Computational study of interaction between *Mycobacterium tuberculosis* enolase and human plasminogen suggested the possibility of selective targeting of enolase to hamper human plasminogen binding (Rahi et al., 2018). In an interesting study, computational analysis of

binding affinity of freely available natural herbal lead compounds towards prostate cancer effector molecule revealed the potential of such compounds in chemotherapy (Bhattacharjee et al., 2011).

Escalation of peptide-based vaccine development has given rise to a number of databases specifically focussing on peptides and their therapeutic aspects such as anti-cancerous, anti-microbial etc. This compilation of otherwise dispersed data has led to development of various tools, softwares and algorithms which help to predict plausible peptide vaccine candidates and/or novel peptide analogs with acceptable accuracy (Usmani et al., 2018).

Numerous works on different pathogens and diseases such as *Influenza virus* (Lohia and Baranwal, 2015), *Human papilloma virus* (Kumar et al., 2015), hepatitis C (Sirskyj et al., 2011), adenovirus (Hossain et al., 2018), tuberculosis (Vani et al., 2006; Monterrubio-López et al., 2015), Leishmaniasis (Agallou et al., 2014), *West Nile virus* (Hossain et al., 2018) and cancer (Dhiman et al., 2016; Naik et al., 2012) have been carried out using various computational approaches for elucidation of epitopes. Successful *in vitro* and/or *in vivo* validation of eptiopes predicted via computational approaches was carried out for *Influenza virus* (Lohia and Baranwal, 2017), *Brucella abortus* (Afley et al., 2015), *Mycobacterium tuberculosis* (Kovjazin et al., 2013), *Schistosoma mansoni* (Oliveira et al., 2016), *Leishmania donovani* (Dikhit et al., 2017), H1N1 virus (Duvvuri et al., 2013) and *Moraxella catarrhalis* (Yassin et al., 2016). An influenza vaccine, concept named FluV, comprising of peptides predicted via immunoinformatics approach has entered phase II clinical trials (Pleguezuelos et al., 2012; van Doorn et al., 2017). Recently, *in silico* prediction and *in vivo* validation of fluoro-benzimidazole derivatives was carried out in an effort to curb Alzheimer's disease (Ali et al., 2019).

2.7.1. Epitope prediction tools

Immunoinformatics tools for epitope prediction have a strong and varied statistical and machine learning system core and cater to various parameters such as proteosomal cleavage and molecular interactions such as antigen processing and presentation. They help in making a knowledge based selection of probable epitopes which can then be validated via *in vitro/in vivo* experiments. The predicted epitopes have varying ability to generate immune response as some epitopes might be dominant over others (Sun et al., 1991) and therefore, peptide

vaccines containing multiple epitopes result in the generation of a greater immune response, thus, serving the purpose of conferring enhanced immunogenicity (Almeida et al., 2012).

HLA class I molecules, present on the exterior of antigen presenting cells, are accountable for presenting antigenic peptides to T_C cells. A wide number of tools for predicting HLA class I binders are available. These tools predict epitopes for a stupendous assemblage of HLA alleles with high efficacy and 90-95% precision (He et al., 2010; Hoof et al., 2009; Lundegaard et al., 2008).

RANKPEP is a popular server based on Position Specific Scoring Matrices (PSSMs) for predicting both, MHC class I and II epitopes. Currently, it is the leading tool in terms of MHC allele coverage as it offers prediction for 118 and 67 MHC class I and II alleles respectively. The algorithm is written in Python language and functions by scoring all protein parts with length of the PSSM width. Also, the prediction is biased towards those MHC class I ligands whose C-terminal end is probably going to be the consequence of proteasomal cleavage (Reche and Reinherz, 2007).

The Immune Epitope database (IEDB) is a free resource since 2004, funded by a contract from the NIAID, USA (Vita et al., 2015). It provides easy search of antibody and T cell epitope data related to infectious diseases, allergies, autoimmunity and transplants in humans and other species reported by scientists across the globe. It also hosts various tools for prediction and analysis of B cell and T cell epitopes (Nielsen et al., 2007). Amongst the various algorithms available to be used for prediction, NetMHCpan is popular for supporting atypical number of organisms (humans, chimpanzee, cow, pig, gorilla, mouse and macaque). It has been in use since 2011 and helps in assessing the affinity of peptides for MHC molecules in a quantitative manner (Hoof et al., 2009).

Another widely used MHC class I epitope prediction server is nHLAPred which provides the user an option to chose from 67 MHC class I alleles and two methods viz., Compred and ANNPred. It is based on Artificial Neural Networks (ANNs) and quantitative matrices (QM) while proteosomal matrices help in selecting potent candidates amongst all the predicted binders. Compred method utilizes a hybrid approach consisting of artificial neural networks and quantitative matrices and hence, is favored by the researchers (Bhasin and Raghava, 2007).

Kernel-based Inter-allele peptide binding prediction SyStem (KISS) is based on support vector machine (SVM) algorithm and predicts MHC class I epitopes for 64 alleles. It has

been coached based on databases having knowledge of epitopes already known to the scientific community obtained from various databases such as SYFPEITHI, IEDB etc (Soria-Guerra et al., 2015).

Apart from MHC class I specific T cell epitope prediction servers, various TAP (transporter associated with antigen processing) binding prediction softwares such as Epijen and TAPPred are available. TAP proteins are heterodimeric transporters included in the ABC transporter family and are associated with MHC class I restricted antigen processing. They help in translocation of peptides beyond the membrane of endoplasmic reticulum with the help of energy generated via ATP hydrolysis (Soria-Guerra et al., 2015).

Unlike MHC class I predictions, variable success has been attained in prediction of MHC class II specific T cell epitopes (Lin et al., 2008; Gowthaman and Agrewala, 2008). Inadequate training data, snag in spotting core binding regions of long peptides, ignorance of flanking peptide affect and the tolerance of MHC class II molecule binding groove towards epitope binding region, thus, allowing peptides to bind with freedom are some of the factors responsible for inconsistent results and modest accuracy of MHC class II prediction servers (He et al., 2010; Lin et al., 2008).

ProPred (Institute of Microbial Technology, India) is based on quantitative matrix algorithm and helps in epitope prediction for 51 MHC class II alleles (Singh and Raghava, 2001). It employs amino acid /position coefficient table derived from pocket profile database described by Sturniolo and coworkers (Sturniolo et al., 1999).

A server highly useful in vaccine design, immunodiagnostics and cellular immunology is MHC2Pred. It is based on SVM algorithm and performs prediction for 42 alleles with an average accuracy of 80%. Small datasets are chiefly responsible for its poor performance for some alleles (Gowthaman and Agrewala, 2008).

A list of various T cell epitope prediction tools has been provided in table 2.7, page 46.

Table 2.7: T cell epitope prediction tools

Tool	Method	MHC		Reference
		I	II	
EpiJen	Muti-step algorithm	24		Doytchinova et al., 2006
SYFPEITHI	Published motifs	42	7	Rammensee et al., 1999
ANNPRED	ANN-regression	30		Bhasin and Raghava, 2007
BIMAS	Published coefficient tables	41		Parker et al., 1994
PROPRED I	Quantitative matrix	47		Soria-Guerra et al., 2015
PROPRED	Quantitative matrix		51	Singh and Raghava, 2001
MHCPred	Additive method	14	11	Guan et al., 2003
MHC2Pred	SVM-based method		42	Gowthaman and Agrewala, 2008
NetMHC	ANN based method	57		Lundegaard et al., 2008
PREDEP	Published coefficient tables	13		Soria-Guerra et al., 2015
RANKPEP	PSSM	118	62	Reche and Reinherz, 2007
SVMHC	SVM-based method	33	51	Donnes and Elofsson, 2001
IEDB binding	ANN and SMM method	77		Nielsen et al., 2007
EpiVax	Epimatrix algorithm	6	8	Soria-Guerra et al., 2015
MMBPred	Quantitative matrix	46		Soria-Guerra et al., 2015
NetCTL	ANN-regression	12		Larsen et al., 2007
nHLAPred	Artificial neural networks	67		Bhasin and Raghava, 2007
KISS	SVM-based method	64		Jacob and Vert, 2008
SVRMHC	SVM-based method	36	6	Soria-Guerra et al., 2015
IMTECH	Quantitative matrix		3	Soria-Guerra et al., 2015

B cell receptors recognize two types of epitopes viz., linear (continuous) and conformational (discontinuous) epitopes. Both these types are responsible for generation of humoral immune response (Soria-Guerra et al., 2015). Prediction of continuous B cell epitopes is very similar to that of T cell epitopes as both predict epitopes based on various amino acid properties such as charge, secondary structure, unshielded surface area etc. Conformational B cell epitope prediction necessitates the knowledge of antigenic 3D structure (Greenbaum et al., 2007; Yao et al., 2013).

Various tools are available for linear/continuous B cell epitope prediction. Bcepred tool is composed of 1029 epitopes procured from Bcipep database and 1029 non-epitopes acquired from Swiss-Prot database. It predicts epitopes with 52.92 - 57.53% accuracy and caters to various properties such as pliability, hydrophilicity, polarity and exposed surface area (Saha and Raghava, 2006). On the other hand, ABCpred server predicts continuous B cell epitopes with 65.93% accuracy and is based on neural network algorithm (Sun et al., 2013). Another server called BepiPred is based on an amalgamation of propensity scale and hidden Markov model methodology (Larsen et al., 2006).

DiscoTope is at the forefront of tools used for conformational B cell epitope prediction. It can be used for epitope prediction in multiple chain complexes. It ascertains surface accessibility and a novel epitope propensity amino acid score based on 3D protein structure followed by the determination of final score by consolidating the residue propensity scores in dimensional juxtaposition and the contact numbers (Kringelum et al., 2012).

A list of various B cell epitope prediction tools is given in the table 2.8, page 48.

Table 2.8: B cell epitope prediction tools

Tool	Method	Reference
Bcepred	Parker, Karplus, Emini and Kolaskar method	Saha and Raghava, 2006
BepiPred	Random Forest algorithm	Jespersen et al., 2017
ABCPred	ANN	Saha and Raghava, 2006
BEST	Support Vector Machine (SVM) method	Soria-Guerra et al., 2015
EPCES	Consensus Scoring	Soria-Guerra et al., 2015
Discotope	Spatial neighborhood definition and half-sphere exposure	Kringelum et al., 2012
BEPro (PEPITO)	Multiple distance thresholds and half sphere exposure	Sweredoski and Baldi, 2008
SEPPA	Unit patch of residue triangle concept	Sun et al., 2009
EpiSearch	Frequency distribution	Negi and Braun, 2009
MimoPro	Dynamic programming (DP) and branch and bound (BB) optimization	Chen et al., 2011
MIMOX	Statistical method	Huang et al., 2006
Pep-3D-Search	Mimotope analysis	Huang et al., 2008
Epitopia	Naive Bayes classifier	Rubinstein et al., 2009
Pepitope	PepSurf and Mapitope algorithms	Soria-Guerra et al., 2015
ElliPro	Thornton's method	Ponomarenko et al., 2008

2.7.2. Role of immunoinformatics in the development of Ebola vaccine

It is significant to counter EBOV spread at a fast pace. Immunoinformatics helps to save both, time as well as cost in developing anti-EBOV measures. Availability of computationally mined EBOV data, EBOV genome and protein sequences, comparative genome analysis procedures, 3-D modelling tools, molecular docking approaches and the possibility of creating a dedicated EBOV database are reasons enough to believe in the potential of immunoinformatics in modern EBOV drug/vaccine development (Karp et al., 2015).

In an earlier study, MHC class I restricted epitopes in EBOV NP were reported based on experimentations on murine models immunized with EBOV NP encoding vectors. (Simmons et al., 2004). In 2007, Sundar et al reported five computationally identified EBOV NP epitopes specific for HLA-A2 allele. Three conserved peptides (FLSFASLFL, RLMRTNFLI and KLTEAITAA) showed encouraging results in MHC stabilisation assay conducted based on flow cytometry. *In vivo* studies conducted in transgenic mice using interferon-γ ELISPOT assay confirmed the ability of these reported peptides to generate cytotoxic T cell response (Sundar et al., 2007).

In another study, computational prediction of H-2d-specific T cell epitopes and their identification in BALB/c mice infected with replication-incompetent, EBOV GP expressing adenovirus vectors was done for Sudan and Zaire EBOV species (Wu et al., 2012). Changula et al worked on filoviral NP and reported two distinctly conserved antigenic regions (residues 421-440 and 601-620) with the help of mouse monoclonal antibodies (Changula et al., 2013).

In 2015, a 100% conserved NP B cell epitope (GEQYQQLR) was reported to aid antibody based therapeutic development against EBOV. It was identified after conducting multiple sequence alignment and conservation analysis studies on various structural and non-structural EBOV proteins (Ali and Islam, 2015). In the same year, 28 EBOV 9-mer epitopes with no apparent human homologs were identified using computer assisted approach (Dikhit et al., 2015). Further, TLASIGTAF was identified as a highly potent B and T cell activating epitope with the ability to be a part of globally protective epitope-based vaccine as it showed positive interaction with 12 HLA alleles and population coverage of 80.99%. It was reported after positive results were obtained in various immunoinformatics based and molecular docking studies (Khan et al., 2015). One of the studies was focussed on EBOV RNA-dependent RNA polymerase-L. After initial sequence and conservancy analysis with the help of various

immunoinformatics tools such as BioEdit and CLC Sequence Viewer, epitope prediction was carried out with the help of IEDB tools followed by population coverage analysis. FRYEFTAPF was reported as the most potent epitope amongst all identified candidates and was found to be non-allelergic in nature. It was 100% conserved in the considered EBOV strains and exhibited a cumulative population coverage of 99.87 % for both MHC classes (Oany et al., 2015).

Off late, three B cell epitopes (PPPPDGVR, ETFLQSPP, LQSPPIRE), four MHC class I specific T cell epitopes (FLYDRLAST, IIIAIIALL, MHNQNALVC and RTYTILNRK) and four MHC class II specific T cell epitopes (FAEGVIAFL, FLRATTELR, FLYDRLAST and FVWVIILFQ) were identified in 21 *Sudan ebolavirus* GP sequences with the help of computational analysis (Abu-Haraz et al., 2017). Also, successful *in vivo* validation of *in silico* derived DNA vaccine containing T cell epitopes was reported (Bounds et al., 2017).

2.8. Gaps in study

Despite numerous efforts to develop therapeutic and preventive measures against EBOV, the virus continues to pose a huge global threat as no prophylactic procedure or treatment course for Ebola virus disease has been licensed as yet. Palliative care is the only solution after an Ebola virus attack. Limited studies have been done to identify potential peptide vaccine candidates through computational approaches followed by simultaneous validation of the *in silico* selected peptides in the *in vitro* system.

Therefore, objectives of the present investigations are:

1. Immunoinformatics based prediction of peptides containing epitopes of Ebola virus

2. *In silico* analysis of peptide MHC interaction

3. Estimation of immunogenic response of *in silico* screened peptides in the *in vitro* system

Chapter 3: Materials and Methods

3.1. Retrieval of glycoprotein and nucleoprotein sequences

Full length Ebola virus (EBOV) glycoprotein (GP) and nucleoprotein (NP) sequences belonging to all human infecting EBOV species were downloaded from NCBI GenBank database and viprbrc (https://www.viprbrc.org/brc/) database. The sequences were manually screened for redundancy.

3.2. Determination of conserved peptide sequences

Multiple sequence alignment of the unique protein sequences was carried out by applying multiple sequence comparison by log expectation (MUSCLE) tool. It aligns the sequences based on iteration and offers a large array of options to the user. MUSCLE has proved to be more efficient as compared to other currently available multiple sequence alignment tools. Its accuracy is statistically equivalent to T-Coffee and MAFFT and it has been rated as fastest among T-Coffee, MAFFT and CLUSTALW for a high sequence number input (Edgar, 2004).

The .fasta file obtained after multiple sequence alignment was used as input for antigenic variability analyser (AVANA) tool to identify conserved regions showing >90% conservancy amongst the GP and NP sequences. AVANA functions by measuring the sequence variability (entropy) in the input file at a given position. AVANA creates a graphical presentation of entropy profile for multiple sequence alignments, thus enabling users to examine position specific variants and their frequencies (Miotto et al., 2008).

3.3. Prediction of $CD8^+$ (HLA class I) and $CD4^+$ (HLA class II) T cell epitopes

A consensus approach was applied for epitope prediction wherein three tools viz., NetCTL v1.2, IEDB analysis resource and SYFPEITHI which are developed based on different algorithms and immunological parameters were employed for $CD8^+$ T cell epitope prediction (Larsen et al., 2007; Nielsen et al., 2003; Rammensee et al., 1999). NetCTL 1.2 uses artificial neural network and weight matrix based method for predicting the HLA class I binding epitope. The server integrates the HLA class I epitope binding prediction with antigen processing steps including C-terminal proteasomal cleavage and TAP (transporter associated

with antigen processing) transport efficiency (Larsen et al., 2007). It predicts epitopes for 12 HLA class I supertypes. Each supertype is a cluster of functionally related HLA alleles that share binding specificities towards the same panel of peptides owing to similar structural features of HLAs peptide binding groove (Lund et al., 2004). IEDB ANN (Nielsen et al., 2003) tool is based on artificial neural network algorithm and calculates the IC_{50} value which shows the binding ability of peptides to specific HLA molecules. SYFPEITHI relies on published motifs for T cell epitope prediction. It considers the positions of amino acids and scores them based on whether they are anchor or auxiliary anchor position preferring residues (Rammensee et al., 1999).

Continuing with the consensus approach, three different tools viz., MHC2pred, IEDB analysis resource and ProPred were employed to increase the stringency for $CD4^+$ T cell epitope prediction. MHC2Pred is based on support vector machine (SVM) method and allows the user to make predictions for 42 MHC class II alleles (Bhasin and Raghava, 2004). IEDB NN Align is based on artificial neural network algorithm and calculates the IC_{50} value which shows the binding ability of peptides to specific HLA molecules (Nielsen et al., 2009). ProPred facilitates predictions based on matrix algorithm for 51 HLA class II alleles (Singh and Raghava, 2001). It employs amino acid/position coefficient table derived from pocket profile database described by Sturniolo and coworkers (Sturniolo et al., 1999). The threshold taken and algorithm used in each tool is mentioned in table 3.1, page 53. All the HLA alleles/supertypes available in each prediction tool were taken into consideration for predicting the epitopes (Table 3.1, page 53).

The epitopes commonly predicted by all tools belonging to a specific HLA class were considered and then searched for overlapping epitopes. The overlapping epitopes were combined to get peptide fragments containing more than one epitope. In this way, peptides containing multiple epitopes were generated for both, HLA class I and II. Further, the peptide fragments which were present commonly in both, HLA class I and II peptide fragments were selected with the intention to consider peptides with a potential to induce both, $CD4^+$ and $CD8^+$ T cell response.

Table 3.1: $CD8^+$ (HLA class I) and $CD4^+$ (HLA class II) T cell epitope prediction tools employed in the current study

T cell	Tools	Algorithm	HLA molecules	Threshold	Reference
$CD8^+$	SYFPEITHI	Motifs search approach	7 HLA-A and 26 HLA-B	20	Rammensee et al., 1999
	NetCTL 1.2	Weight-matrix method	5 HLA-A and 7 HLA-B	0.75	Larsen et al., 2007
	IEDB consensus	Artificial neural network	18 HLA-A and 32 HLA-B	IC_{50}< 500nm	Nielsen et al., 2007
$CD4^+$	MHC2Pred	Support vector machine (SVM) method	11 HLA-DQ and 27 HLA-DR	0.5	Bhasin and Raghava, 2004
	Propred	Quantitative matrices	51 HLA-DR	3%	Singh and Raghava, 2001
	IEDB consensus	Artificial neural network	5 HLA-DP, 6 HLA-DQ and 14 HLA-DR	IC_{50}< 500nm	Nielsen et al., 2007

3.4. B cell epitope prediction

Presence of B cell epitopes in the selected peptide fragments was checked with the help of ABCPred tool. This tool utilizes a recurring neural network method to predict linear B cell epitopes (Saha and Raghava, 2006). The window length was kept as 10. The peptides devoid of B cell epitopes were not considered further.

3.5. Elimination of peptides showing autoimmune, allergic and toxic response

3.5.1. BLASTp

Immunogenic response is not mounted against self-peptides unless failure of clonal deletion leads to selection of auto-reactive T cells. Further, similarity of the identified peptides with host self-protein may sometimes lead to an unwanted immunogenic response (autoimmunity). Hence, in order to obtain non-self-peptides, the selected peptide fragments were analysed for their homology with annotated human proteome using BLASTp (Altschul et al., 1990). A fragment was considered as self if it shared a minimum seven consecutive identical amino acids with the human peptides (Dhiman et al., 2016). Such fragments were eliminated from further studies.

3.5.2. AlgPred and ToxinPred

AlgPred and ToxinPred were used for screening of peptides having allergic and toxic nature respectively. AlgPred provides the option to follow various prediction approaches such as mapping of IgE epitopes and MEME/MAST motif to measure allergenicity of the query sequence (Saha and Raghava, 2006). ToxinPred classifies toxic and non-toxic peptides based on dataset of non-toxic or random peptides (Gupta et al., 2013).

3.6. HLA coverage analysis

This analysis was carried out to assess the promiscuous (ability to interact with multiple HLA alleles) nature of the identified peptide fragments. Various epitope prediction tools cater to different types of HLA supertypes/alleles (HLA-A, HLA-B, HLA-DP, HLA-DQ and HLA-DR types). All HLA supertypes/alleles, for which epitopes present in the selected peptide fragments were predicted by the six T cell epitope prediction tools, were listed manually. This list was further used as input for population coverage analysis.

3.7. Population coverage analysis

There is a drastic difference in the frequencies of expression of various HLA types among individuals from different continents. Because of this, population coverage analysis of the HLA alleles corresponding to the predicted immunogenic peptides becomes important. The population coverage tool by IEDB calculates the proportion of individuals responding to a given set of epitopes with known HLA restriction taking into account HLA genotypic frequency (Bui et al., 2006). Population coverage analysis tool obtains the HLA allele frequencies for different individual populations from the allele frequency net database (Gonzalez-Galarza et al., 2015). In population coverage analysis tool, the populations are organized in a pecking order based upon geographical area, country and ethnicity. The data from the individual populations in each group was combined to evaluate the allele frequencies for each merged population. Analysis can be achieved for population distributed in different continents viz., Asia (East, Northeast, South, Southeast and Southwest Asia), Europe, Africa (East, West, Central, North and South Africa) and America (North, South and Central America). Population coverage analysis of the peptides was carried out for the population distributed across these continents. Analysis was also done for the whole world.

3.8. Molecular docking

The interaction of the predicted peptides containing epitopes with various HLA class I and II molecules was further studied by means of molecular docking. High-resolution crystallographic structures of nine HLA class I and nine HLA class II molecules (Table 3.2, page 56) were retrieved from the protein data bank (Berman et al., 2000). The naturally bound peptides (native peptides) of the HLA molecule were separated using the Discovery studio visualizer v4.1 tool. The resultant HLA molecules were used as receptors to dock the *in silico* identified epitopes/peptides and the separated naturally bound peptides (positive control) using AutoDock Vina (Trott and Olson, 2010) and CABS-dock (Blaszczyk et al., 2016). The structure of epitopes present in the selected peptide fragments (for docking with HLA class I molecules) and the selected peptides themselves (for docking with HLA class II molecules) was generated using the PEP-FOLD server (Shen et al., 2014). PEP-FOLD provides a fast and user friendly approach for the *de novo* design of small peptides (9 to 36 residues). It is based on OPEP coarse grained force field for molecular simulation. The top model structures predicted for the peptides were used for docking.

Table 3.2: HLA alleles and their resolution

HLA Class I			HLA Class II		
PDBid	**HLA molecule**	**Resolution (Å)**	**PDBid**	**HLA molecule**	**Resolution (Å)**
3C9N	B*1501	1.87	4P5M	DPA1	1.7
4JQV	B*18:01	1.5	IUVQ	DQ0602	1.8
2A83	B*2705	1.4	2NNA	DQ8	2.1
4O2F	B*3901	1.9	2FSE	DR1	3.1
2BVP	B5703	1.35	1KLU	DRB1*0101	1.93
3BO8	A1	1.8	1BX2	DRB1*1501	2.6
3MRE	A2	1.1	1A6A	DR3	2.75
3RL1	A*0301	2.0	1D5M	DR4	2.0
3WL9	A24	1.66	3C5J	DR5	1.8

3.8.1. AutoDock Vina

Docking with AutoDock Vina involves a series of steps including receptor and ligand preparation, defining the search space (grid), preparing the configuration file and visualisation of the AutoDock Vina results. AutoDockTools (ADT) is the graphical user Interphase (GUI) employed by AutoDock to set up the ligand, receptor, search space and to visualise the results. The input files for the receptor and the ligand as well as the output obtained after docking are in PDBQT format which stores the atomic coordinates, partial charges and AutoDock atom types.

3.8.1.1. Receptor preparation

HLA molecules (receptor) were obtained from PDB database in the form of a coordinate data file (PDB file format). PDB structure of HLA was opened in Discovery studio visualizer v4.1 tool. Water molecules were deleted while all the ligands and heteroatoms were selectively removed. Hydrogen atoms were added to the HLA peptide chains and the resultant file was saved as “receptor.pdb”. As the receptor.pdb file was opened in ADT using macromolecule option in Grid widget, ADT detected the charges on molecules. Inherent charge on the receptor molecule was preserved. In case of absence of any inherent charges, Gasteiger charges were automatically added to each atom. ADT determined the type of atoms in the macromolecule. Output was saved as “receptor.pdbqt” file.

3.8.1.2. Ligand preparation

Structure of peptides (ligand) was generated from PEP-FOLD server in form of a coordinate data file (PDB file format). Naturally bound peptides were obtained from parent HLA by selectively removing the HLA chains and heteroatoms from the PDB structures in Discovery studio visualizer v4.1 tool. Hydrogen atoms were added at polar and non-polar positions and the resultant file was saved as the "ligand.pdb". As the ligand.pdb file was opened in ADT, it computed and added Gasteiger charges for the entire ligand (in case the charge is zero); else partial charges were used. ADT assigned an autodock type to each atom in the ligand: atoms forming hydrogen bond and aromatic carbons. Output was saved as "ligand.pdbqt" file.

3.8.1.3. Grid optimization

Grid optimization is the most significant step of docking via AutoDock Vina. It requires us to specify the space where the peptide is ideally expected to bind; the 3D space called as the binding groove. If a peptide binds outside the binding groove, it is considered as a non-binder and as futile for vaccine development. Therefore, defining the grid precisely returns true results. In order to define the location and size of the 3D area to be searched during the AutoDock Vina procedure, grid box option in grid widget of ADT was used. The thumbwheels provided are controlled by dragging the pointer in L to R direction using the conventional left mouse button. Using the thumbwheels in grid option panel, search space was defined by specifying x, y, and z center, the number of points in each dimension (x, y, z) and the spacing (in angstrom) between points. The output grid dimension file was saved as "grid.txt" and the grid option was closed.

3.8.1.4. Configuration file preparation

A text file called configuration file which contains the information regarding the name of inputs files i.e. ligand.pdbqt and receptor.pdbqt, name of the output file to be generated, search space parameters (center as well as the size in each dimension) and exhaustiveness (the time spent on the search) was prepared and saved as "conf.txt". The format of this text file is as mentioned:-

receptor = *(Name of the receptor.pdbqt file)*

ligand = *(Name of the ligand.pdbqt file)*

out = out.pdbqt

center_x = *(The value of center for x obtained during grid optimization)*

center_y = *(The value of center for y obtained during grid optimization)*

center_z = *(The value of center for z obtained during grid optimization)*

size_x = *(The number of points in x dimension obtained during grid optimization)*

size_y = *(The number of points in y dimension obtained during grid optimization)*

size_z = *(The number of points in z dimension obtained during grid optimization)*

exhaustiveness = *(The time spent on the search)*

3.8.1.5. AutoDock Vina process

All the prepared files viz., receptor.pdb, receptor.pdbqt, ligand.pdb, ligand.pdbqt, grid.txt and conf.txt were kept in one folder. The command line interface terminal was opened and the following steps were followed:-

cd.. *(Command to change the working directory)*

↓

cd The folder was dragged onto the terminal followed by Vina software

↓

--config conf.txt --log log.txt *(Command to start the docking process)*

↓

Vina-split software was dragged onto the terminal

↓

--input out.pdbqt *(Command to generate individual outputs from a multimodel file)*

Outcome was visualised by using analyze widget in ADT to confirm if the peptide was a binder/non-binder as well as to analyse the molecular interactions.

3.8.2. CABS-dock

CABS is an abbreviation of single amino acid residue letters: carbon alpha (CA), carbon beta (B) and side chain (S). It is based on CABS model which has been designed to control multimeric protein chains (Blaszczyk et al., 2016). CABS-dock server follows a four stage protocol:-

a) Flexible docking based on CABS algorithm
b) Selection of probable models from the complete list of generated models
c) Selection of representative models
d) Optimization of final models.

Unlike AutoDock Vina, CABS-dock does not require the user to define a search space. Also, the .pdbqt format is not needed to work with CABS-dock. Moreover, generation of peptide structure is not a pre-requisite. CABS-dock allows for blind protein-peptide docking where it generates random structures of the peptide and places them randomly on the surface of the sphere centered at the receptor's geometrical center. Starting from random conformations and positions of the peptide, CABS-dock performs simulation search for the binding site allowing for full flexibility of the peptide (Blaszczyk et al., 2016). It accepts the receptor file in .pdb format and peptide sequence for docking.

3.8.2.1. Receptor preparation

HLA molecules (receptor) were obtained from protein data bank (PDB) database in the form of a coordinate data file (PDB file format). PDB structure of HLA was opened in Discovery studio visualizer v4.1 tool. Water molecules were deleted while all the ligands and

heteroatoms were selectively removed. Hydrogen atoms were added to the HLA peptide chains and the resultant file was saved as “receptor.pdb”.

3.8.2.2. File submission

As input, CABS-dock requires the receptor.pdb file generated earlier and the peptide sequence. The native ligand sequence was obtained by using the “Show Sequence” option under “Sequence” tab in Discovery studio visualizer v4.1 tool. As CABS-dock is an online server, an email address can be provided (optional) to receive results once the job is finished.

The link to docking results leads to the page which helps in viewing individual models and analyze if the peptide is a binder/non-binder.

3.9. Peptide synthesis

The selected peptides were synthesised by GL Biochem (Shanghai) Ltd (China) with purity ≥ 90%. Peptides were dissolved in DMSO at a concentration of 1 mg/mL and stored in small aliquots at -20 °C.

3.10. Blood sampling from healthy volunteers

Healthy human volunteers of age ≥ 18 years having no history of hepatitis B, C and/or HIV infection were included in the study. Blood was drawn via venipuncture by the trained technicians of Rajindra Hospital, Patiala and Nitin Hospital, Patiala in blood collection EDTA quoted tubes (BD Vacutainer® Tubes). All the volunteers gave their informed consent to donate blood for the experiments and the study was approved by institutional ethical committee.

3.11. Chemicals and reagents used

All the chemicals and reagents used in this work were of cell culture grade (Table 3.3, page 61-62).

Table 3.3: List of chemicals and reagents used in the current study

S.No.	Reagents/Chemical	Company
1	ABTS (2,2'-Azinobis[3-ethylbenzothiazoline-6-sulfonicacid]-diammonium salt) substrate	ThermoFisherScientific, USA
2	Amphotericin B	Sigma Aldrich, USA
3	Bovine serum albumin (BSA)	Sigma Aldrich
4	Concanavalin A (Con A)	Sigma Aldrich
5	Dimethyl Sulphoxide (DMSO)	Merck, Germany
6	Fetal bovine Serum (FBS)	Gibco®Life Technologies, USA
7	HEPES buffer	Sigma Aldrich
8	Histopaque® -1077	Sigma Aldrich
9	Human IFN-γ Mini ABTS ELISA Development Kit	PeproTech, USA
10	MTT(3-(4,5-Dimethylthiazol-2-yl)-2,5-Diphenyltetrazolium Bromide)	Sigma Aldrich
11	Penicillin Sodium	Himedia
12	Potassium Chloride (KCl)	Himedia
13	Potassium phosphate monobasic (KH_2PO_4)	Himedia
14	Rosewell Park Memorial Institute (RPMI)-1640 medium	Sigma Aldrich
15	Sodium Bicarbonate ($NaHCO_3$)	Himedia
16	Sodium Chloride (NaCl)	Himedia

17	Sodium phosphate dibasic (Na_2HPO_4)	Himedia
18	Streptomycin	Sigma Aldrich
19	Trypan blue	Himedia
20	Tween 20	Sigma Aldrich

3.12. Peripheral blood mononuclear cells isolation

Peripheral blood mononuclear cells (PBMC) were isolated by ficoll density gradient method (Lohia and Baranwal, 2017). Blood was carefully layered onto Histopaque®-1077 in the ratio of 1:1 and centrifuged at 400g for 30 min at room temperature in a swinging bucket rotor (Thermo Scientific Biofuge Stratos). This density based centrifugation technique fractionates blood into plasma, peripheral blood mononuclear cells (PBMC) and red blood cells. After centrifugation, the upper plasma layer was discarded with a micropipette and opaque interface (buffy coat) containing PBMC was transferred into a sterile 15 mL falcon. The cells were washed twice in 8 mL isotonic phosphate buffered saline solution by centrifugation at 330g for 12 min. Finally, the cell pellet was re-suspended in 1 mL of complete media (RPMI-1640 supplemented with 10% fetal bovine serum, 100 µg/mL streptomycin, 100 IU/mL penicillin and 10 mM HEPES).

3.13. Cell counting

Cells were counted with the help of a haemocytometer by trypan blue exclusion assay (Strober, 2001). Briefly, 10 µL of cell sample was diluted with 10 µL 0.4% Trypan Blue solution and the final volume was raised to 100 µL with the help of complete media. 20-25 µL of this preparation was loaded on the haemocytometer and examined immediately under a microscope (Nikon Eclipse E100-LED) at 40X magnification. The number of unstained viable cells (non-viable cells take up dye and appear blue) were counted in all the four corner squares. Number of viable cells in the original cell suspension was calculated from the following formula:-

$$\text{Number of viable cells/mL} = \frac{\text{Number of unstained cells} * 10000 * \text{dilution factor}}{4}$$

where dilution factor = 10 and number of unstained cells = total cells present in four corner squares

3.14. PBMC proliferation assay

Peptide induced proliferation of PBMC was measured by MTT based colorimetric assay (do Livramento et al., 2013). MTT assay measures the cell viability based on reduction of the yellow colored tetrazolium salt MTT into purple formazan crystals by mitochondrial succinate dehydrogenase of metabolically active cells (Mosmann, 1983). To measure proliferation, the isolated cells were seeded in a 96 well cell culture plate at a concentration of $2*10^5$ cells per well on 0^{th} day. Three sets of cells were made viz., unstimulated cells (negative control), peptide stimulated cells (cells stimulated with peptide at a concentration of 50 µg/mL) and cells stimulated with concanavalin A (positive control) at a concentration of 10 µg/mL. All experiments were carried out in triplicates and the plate was incubated at 37°C and 5% CO_2. On 3^{rd} day, peptide and ConA restimulation was given to cells. On 5^{th} day, MTT (0.5 mg/mL) was added to the cell culture and the plate was again incubated at 37°C and 5% CO_2. After an incubation of 4 h, 120 µL of the supernatant was removed without disturbing the purple colored crystals and stored at -80°C. These formazan crystals were solubilized in 100 µL DMSO. Finally, absorbance of each well was recorded at 570 nm with 630 nm as the reference wavelength using a microplate reader (Tecan Austria). Proliferation was calculated in terms of stimulation index (SI) as following:-

$$\text{SI} = \frac{\text{Average absorbance of peptide stimulated cells}}{\text{Average absorbance of unstimulated cells}}$$

3.15. IFN-γ measurement assay

In order to detect extracellular secretion of IFN-γ by peptide stimulated T cells, the supernatant extracted (as discussed in section 3.14.) was analyzed with sandwich ELISA (Agallou et al., 2014; Ohkuri et al., 2009) in 96-well ELISA plate (Nunc MaxiSorp®) as per manufacturer's instruction. Briefly, 100 µL of capture antibody (1µg/mL in PBS) was added to each ELISA plate well and incubated overnight at room temperature. The wells were aspirated to remove liquid and the plate was washed four times with 300 µL wash buffer

(0.05% Tween-20 in PBS) per well. After the last wash, the plate was tapped in inverted position to remove residual buffer on paper towel. 300 µL of blocking buffer (1% BSA in PBS) was added to each well and the plate was incubated for 1h at room temperature. After washing the plate four times with wash buffer, 100 µL of the test sample and the IFN-γ standard were added to each well in triplicate. The plate was incubated for 2 h at room temperature. The plate was washed 4 times and 100 µL of detection antibody (1µg/mL in sample diluent buffer) was added to each ELISA plate well and incubated at room temperature for 2 h. Plate was washed four times and 100 µL of diluted avidin-HRP conjugate (1:2000) in sample diluent was added and incubated for 30 min. Plate was again washed four times and 100 µL of ABTS substrate solution was added to each well. The plate was wrapped in a foil and incubated at room temperature for 20 min for color development. Absorbance was recorded at 405 nm with wavelength correction set at 650 nm in ELISA plate reader (Tecan, Austria). IFN-γ production was expressed as fold change in cytokine release as following:-

$$\text{Fold change} = \frac{\text{Average absorbance of peptide stimulated cells}}{\text{Average absorbance of unstimulated cells}}$$

3.16. Statistical analysis of docking and *in vitro* results

GraphPad Prism was used to generate column statistics (mean, median, mode, 25th percentile, max, min, 75th percentile and standard deviation and skewness) of results obtained by molecular docking (Autodock Vina and CABS-dock). One way ANOVA analysis followed by Tukey's multiple comparison test was carried out to check if there was a significant difference between the results obtained by docking native peptides and those obtained by docking Ebola peptides.

GraphPad Prism was used to analyse the MTT and ELISA results by performing t-test for unpaired data wherein p-values < 0.05 were considered as significant. Further, correlation analysis was carried out by calculating Pearson r and two-tailed p value.

Chapter 4: Results

Objective 1

4.1. Prediction of conserved peptides containing multiple T and B cell epitopes of glycoprotein and nucleoprotein

T_c ($CD8^+$) and T_h ($CD4^+$) cells recognize antigens presented to them by HLA class I molecules (present on the surface of nucleated cells) and HLA class II molecules (present on the surface of antigen presenting cells) respectively. In order to identify the peptides containing multiple epitopes, full length protein sequences of Ebola virus (EBOV) glycoprotein (GP) and nucleoprotein (NP) were obtained from viprbrc and NCBI databases. MUSCLE tool was used for multiple sequence alignment. Aligned sequences of these proteins were used to identify conserved sequences by AVANA. These conserved regions were used as input for various T cell epitope prediction tools. They were searched for presence of T cell epitopes by using six epitope prediction tools (three for each, $CD4^+$ and $CD8^+$ T cell epitopes). Peptides containing multiple $CD8^+$ and $CD4^+$ epitopes respectively were obtained by merging overlapping epitopes. Further, common regions amongst these peptide fragments were identified so as to obtain such fragments containing both, $CD8^+$ and $CD4^+$ epitopes. B cell epitope prediction was carried out with the help of ABCPred tool. Screening of peptides for undesirable responses was done with the help of BLAST (autoimmunity), AlgPred (allergenicity) and ToxinPred (toxicity). Finally, peptides devoid of unwanted responses and containing multiple T and B cell epitopes were considered for further study.

4.1.1. Identification of GP and NP conserved regions

1092 GP and 2407 NP sequences belonging to human infecting EBOV species were downloaded from NCBI and viprbrc databases. These sequences were manually screened for redundancy resulting in 173 GP and 195 NP unique sequences (Table 4.1, page 66). To develop a vaccine useful for a comparatively longer period of time, it is essential to identify conserved regions in the viral proteins as they are least likely to undergo mutation in future. Therefore, the sequences were aligned with the help of MUSCLE tool and the conserved regions exhibiting ≥90% conservancy were identified with the help of AVANA tool. Ten

(Table 4.2, page 67) and four (Table 4.3, page 68) conserved regions were identified in GP and NP sequences respectively. These conserved regions were used as input for various epitope prediction tools.

Table 4.1: Glycoprotein and nucleoprotein sequences belonging to human infecting Ebola virus species

Protein	Time frame	Total sequences	Unique sequences	
Glycoprotein	1976 – 30th Jan 2018	1092	*Aggregate**	173
			Zaire	164
			Sudan	5
			Bundibugyo	3
			Tai Forest	1
Nucleoprotein	1976 – May 2018	2407	*Aggregate*	195
			Zaire	187
			Sudan	5
			Bundibugyo	2
			Tai Forest	1

*Aggregate is the summation of the sequences of the four EBOV virus species viz., Zaire, Sudan, Bundibugyo and Tai Forest.

Table 4.2: Conserved peptide fragments of Ebola virus glycoprotein

	Conserved peptide fragment	Location	Length
CGp1	MGVTGILQLPRDRFKRTSFFLWVIILFQRTFSIPLGVIHNSTLQVSDVDKLVCRDKLSSTNQLRSVGLNLEGN GVATDVPS	1-81	81
CGp2	TKRWGFRSGVPPKVVNYEAGEWAENCYNLEIKKPDGSECLPAAPDGIRGFPRCRYVHKVSGTGPCAGDFA FHKEGAFFLYDRLASTVIYRGTTFAEGVVAFLILPQAKKDFFSSHPLREPVNATEDPSSGYYSTTIRYQATGF GTNE	83-229	147
CGp3	EYLFEVDNLTYVQLESRFTPQFLLQLNETIY	231-261	31
CGp4	SGKRSNTTGKLIWKVNPEIDTTIGEWAFWETKKNLTRKIRSEELSFTAVSN	263-313	51
CGp5	KNISGQSPARTSSDP	316-330	15
CGp6	EDHKIMASENSSAMVQVHSQGR	337-358	22
CGp 7	AAVSHLTTLAT	360-370	11
CGp8	GPDNSTHNTPVYKLDISEATQ	390-410	21
CGp9	SETAGNNNTHHQDTGEESASSGKLGLITNTIAGVAGLITGGRRTRRE	465-511	47
CGp10	IVNAQPKCNPNLHYWTTQDEGAAIGLAWIPYFGPAAEGIYTEGLMHNQDGLICGLRQLANETTQALQLFLR ATTELRTFSILNRKAIDFLLQRWGGTCHILGPDCCIEPHDWTKNITDKIDQIIHDFVDKTLPDQGDNDNWWT GWRQWIPAGIGVTGVIIAVIALFCICKFVF	513-685	173

Table 4.3: Conserved peptide fragments of Ebola virus nucleoprotein

	Conserved peptide fragment	Location	Length
CNp1	MDSRPQKVWMTPSLTESDMDYHKILTAGLSVQQGIVRQRVIPVYQVNNLEEICQLIIQAFEAGVDFQESADSFLLMLC LHHAYQGDYKLFLESGAVKYLEGHGFRFEVKK	1-110	110
CNp2	DGVKRLEELLPAVSSGRNIKRTLAAMPEEETTEANAGQFLSFASLFLPKLVVGEKACLEKVQRQIQVHAEQGLIQYPT AWQSVGHMMVIFRLMRTNFLIKFLLIHQGMHMVAGHDANDAVISNSVAQARFSGLLIVKTVLDHILQKTERGVRLHP LARTAKVKNEVNSFKAALSSLAKHGEYAPFARLLNLSGVNNLEHGLFPQLSAIALGVATAHGSTLAGVNVGEQYQQL REAATEAEKQLQQYAESRELDHLGLDDQEKKILMNFHQKKNEISFQQTNAMVTLRKERLAKLTEAITAASLPKTSGH YDDDDDIPFPGPINDDDNPGHQDDDPTDSQDTTIPDVVVDPDDGGYGEYQSYSENGMSAPDDLVLFDLDEDDEDTKP VPNRSTKGGQQKNSQKGQHTEGRQTQSTPTQN	112-529	418
CNp3	GPRRTIHHASAPLTDNDRRNEPSGSTSPRMLTPINEEADPLDDADDETSSLPPLESDDEEQDRDGTSNRTPTVAPPAPVY RDHSEKKELPQDEQQDQDH	534-632	99
CNp4	NQDSDNTQPEHSFEEMYRHILRSQGPFDAVLYYHMMKDEPVVFSTSDGKEYTYPDSLEEEYPPWLTEKEAMNDENRF VTLDGQQFYWPVMNHRNKFMAILQHHQ	638-741	104

4.1.2. Peptides containing overlapping T cell epitopes

Six T cell epitope prediction tools were used for $CD8^+$ T cell epitopes (SYFPEITHI, NetCTL 1.2 and IEDB analysis resource) and $CD4^+$ T cell epitopes (ProPred, MHC2Pred and IEDB analysis resource) prediction. The prediction results were compared and $CD8^+$ and $CD4^+$ epitopes commonly predicted by all the tools were considered. The overlapping $CD8^+$ and $CD4^+$ T cell epitopes obtained were merged into peptide fragments.

4.1.2.1 Ebola glycoprotein

261, 186 and 469 redundant epitopes were predicted with the help of $CD8^+$ T cell epitope prediction tools viz., SYFPEITHI, NetCTL 1.2 and IEDB analysis resource respectively (data not shown). 84 epitopes were found to be commonly predicted by all the tools (Table 4.4, page 70). The overlapping epitopes were merged to obtain 15, multiple epitope containing peptide fragments (Table 4.5, page 71). Similarly, 72 epitopes were found to be commonly predicted by three $CD4^+$ T cell epitope prediction tools viz., MHC2Pred, Propred and IEDB analysis resource (Table 4.6, page 72) out of initial raw prediction data of 2381, 820 and 2717 redundant epitopes (data not shown) respectively. 13 peptide fragments containing multiple $CD4^+$ T cell epitopes were generated by merging the overlapping epitopes (Table 4.7, page 73).

Table 4.4: $CD8^+$ T cell epitopes commonly predicted for glycoprotein by the three prediction tools

$CD8^+$ T cell epitopes				
ILFQRTFSI	EWAENCYNL	YLFEVDNLT	THNTPVYKL	TELRTFSIL
RTSFFLWVI	TVIYRGTTF	TYVQLESRF	LITNTIAGV	NQDGLICGL
VIHNSTLQV	AAPDGIRGF	EYLFEVDNL	KLGLITNTI	AEGIYTEGL
QLRSVGLNL	GVPPKVVNY	SRFTPQFLL	ITGGRRTRR	NETTQALQL
VIILFQRTF	LPQAKKDFF	LFQRTFSIP	NTIAGVAGL	QDEGAAIGL
FFLWVIILF	GPCAGDFAF	LESRFTPQF	ETAGNNNTH	AQPKCNPNL
DRFKRTSFF	FPRCRYVHK	TTGKLIWKV	TIAGVAGLI	RATTELRTF
LPRDRFKRT	FAFHKEGAF	LIWKVNPEI	SASSGKLGL	FEVDNLTYV
RFKRTSFFL	KRWGFRSGV	NTTGKLIWK	EESASSGKL	
FQRTFSIPL	DRLASTVIY	TTIGEWAFW	GLMHNQDGL	
KRTSFFLWV	FHKEGAFFL	ETKKNLTRK	RTFSILNRK	
RTFSIPLGV	FAEGVVAFL	DTTIGEWAF	QIIHDFVDK	
ATEDPSSGY	KDFFSSHPL	TRKIRSEEL	ETTQALQLF	
TEDPSSGYY	AENCYNLEI	KRSNTTGKL	LANETTQAL	
HKEGAFFLY	AEGVVAFLI	SEELSFTAV	FSILNRKAI	
FLYDRLAST	LFEVDNLTY	KIRSEELSF	QRWGGTCHI	
RLASTVIYR	LLQLNETIY	AAVSHLTTL	LRTFSILNR	
FLILPQAKK	FLLQLNETI	NSTHNTPVY	NRKAIDFLL	
TFAEGVVAF	FTPQFLLQL	STHNTPVYK	IHDFVDKTL	

Table 4.5: Peptide fragments of glycoprotein containing overlapping $CD8^+$ T cell epitopes

	Peptides fragments	Length	Number of epitopes
CGp1	LPRDRFKRTSFFLWVIILFQRTFSIPLGV	29	10
CGp2	KRWGFRSGVPPKVVNY	16	2
	EWAENCYNLEI	11	2
	GPCAGDFAFHKEGAFFLYDRLASTVIYRGTTFAEGVVAFLILPQAKKDFFSSHPL	55	14
	ATEDPSSGYY	10	2
CGp3	EYLFEVDNLTYVQLESRFTPQFLLQLNETIY	31	10
CGp4	KRSNTTGKLIWKVNPEI	17	4
	DTTIGEWAFW	10	2
	ETKKNLTRKIRSEELSFTAV	20	4
CGp8	NSTHNTPVYKL	11	3
CGp9	EESASSGKLGLITNTIAGVAGLI	23	6
CGp10	AEGIYTEGLMHNQDGLICGL	20	3
	LANETTQALQLF	12	3
	RATTELRTFSILNRKAIDFLL	21	6
	QIIHDFVDKTL	11	2

Table 4.6: $CD4^+$ T cell epitopes commonly predicted for glycoprotein by the three prediction tools

$CD4^+$ T cell epitopes			
WVIILFQRT	FHKEGAFFL	FLLQLNETI	LANETTQAL
VIILFQRTF	YDRLASTVI	WAFWETKKN	WTKNITDKI
ILFQRTFSI	IYRGTTFAE	IMASENSSA	IGVTGVIIA
FQRTFSIPL	IRYQATGFG	MVQVHSQGR	LHYWTTQDE
IILFQRTFS	FFLYDRLAS	MASENSSAM	LFLRATTEL
FLWVIILFQ	FLYDRLAST	YKLDISEAT	LRTFSILNR
FFLWVIILF	FLILPQAKK	VYKLDISEA	ILNRKAIDF
LWVIILFQR	VVAFLILPQ	LGLITNTIA	LQRWGGTCH
FKRTSFFLW	VIYRGTTFA	LRQLANETT	IPYFGPAAE
LGVIHNSTL	YRGTTFAEG	FSILNRKAI	YFGPAAEGI
VIHNSTLQV	YNLEIKKPD	WRQWIPAGI	IAVIALFCI
IHNSTLQVS	IRGFPRCRY	IIAVIALFC	LQLFLRATT
LQVSDVDKL	YVHKVSGTG	IVNAQPKCN	WIPYFGPAA
LVCRDKLSS	FFSSHPLRE	VIALFCICK	LLQRWGGTC
MGVTGILQL	YYSTTIRYQ	IGLAWIPYF	IALFCICKF
LRSVGLNLE	VAFLILPQA	IIHDFVDKT	
LNLEGNGVA	YVQLESRFT	VIIAVIALF	
LEGNGVATD	FEVDNLTYV	LICGLRQLA	
FRSGVPPKV	YLFEVDNLT	FLRATTELR	

Table 4.7: Peptide fragments of glycoprotein containing overlapping $CD4^+$ T cell epitopes

	Peptide fragments	Length	Number of epitopes
CGp1	FKRTSFFLWVIILFQRTFSIPLGVIHNSTLQVSDVDKLVCRDKLSS	46	14
	LRSVGLNLEGNGVATD	16	3
CGp2	FHKEGAFFLYDRLASTVIYRGTTFAEG	27	7
	VVAFLILPQAKK	12	3
	IRGFPRCRYVHKVSGTG	17	2
	YYSTTIRYQATGFG	14	2
CGp3	YLFEVDNLTYVQLESRFT	18	3
CGp6	IMASENSSAMVQVHSQGR	18	3
CGp8	VYKLDISEAT	10	2
CGp10	IGLAWIPYFGPAAEGI	16	4
	LICGLRQLANETTQALQLFLRATTELRTFSILNRKAIDF	39	9
	LLQRWGGTCH	10	2
	WRQWIPAGIGVTGVIIAVIALFCICKF	27	7

4.1.2.2 Ebola nucleoprotein

Initially, a total of 890 redundant epitopes (data not shown) were predicted by $CD8^+$ T cell epitope prediction tools (347, 157 and 386 epitopes by SYFPEITHI, NetCTL 1.2 and IEDB analysis resource respectively). 105 epitopes were found to be commonly predicted by all three tools and hence, were considered further (Table 4.8, page 75). Overlapping epitopes were merged to obtain 18 peptide fragments containing multiple $CD8^+$ epitopes (Table 4.9, page 76).

7247 redundant epitopes (data not shown) were predicted by $CD4^+$ T cell epitope prediction tools (2544, 965 and 3738 epitopes by MHC2Pred, Propred and IEDB analysis resource respectively) out of which, 79 epitopes were found to be commonly predicted (Table 4.10, page 77). 15 peptide fragments containing multiple $CD4^+$were generated by merging the overlapping epitopes (Table 4.11, page 78).

Table 4.8: $CD8^+$ T cell epitopes commonly predicted for nucleoprotein by the three prediction tools

$CD8^+$ T cell epitopes					
FLESGAVKY	LESGAVKYL	ILMNFHQKK	EVNSFKAAL	ERGVRLHPL	AMNDENRFV
SLTESDMDY	LEEICQLII	KILMNFHQK	FPQLSAIAL	TEANAGQFL	AVLYYHMMK
SADSFLLML	SSLAKHGEY	TVLDHILQK	HPLARTAKV	AQARFSGLL	VVFSTSDGK
ESADSFLLM	VVDPDDGGY	LMRTNFLIK	APFARLLNL	LEHGLFPQL	NHRNKFMAI
KILTAGLSV	HAEQGLIQY	RSTKGGQQK	VATAHGSTL	GEYAPFARL	FEEMYRHIL
YQGDYKLFL	SLPKTSGHY	SFASLFLPK	QARFSGLLI	AESRELDHL	HRNKFMAIL
LLMLCLHHA	MSAPDDLVL	TLRKERLAK	LRKERLAKL	GEKACLEKV	SQGPFDAVL
IIQAFEAGV	FLSFASLFL	AITAASLPK	VAQARFSGL	IQVHAEQGL	LEEEYPPWL
KLFLESGAV	RLMRTNFLI	RFSGLLIVK	FRLMRTNFL	WQSVGHMMV	KEYTYPDSL
YQVNNLEEI	KLTEAITAA	QTNAMVTLR	QEKKILMNF	ISFQQTNAM	FVTLDGQQF
AYQGDYKLF	RLEELLPAV	AVSSGRNIK	MRTNFLIKF	TVAPPAPVY	
IVRQRVIPV	QLSAIALGV	ISNSVAQAR	KRLEELLPA	GTSNRTPTV	
IPVYQVNNL	GLFPQLSAI	GVRLHPLAR	QQTNAMVTL	TIHHASAPL	
YHKILTAGL	RLAKLTEAI	VSSGRNIKR	ARLLNLSGV	DETSSLPPL	
VRQRVIPVY	FASLFLPKL	LLPAVSSGR	ARFSGLLIV	TSDGKEYTY	
HAYQGDYKL	GMSAPDDLV	EYAPFARLL	HQKKNEISF	VTLDGQQFY	
FQESADSFL	SQDTTIPDV	QFLSFASLF	GHMMVIFRL	FSTSDGKEY	
QKVWMTPSL	LIHQGMHMV	NFLIKFLLI	TEAITAASL	YPDSLEEEY	
QESADSFLL	GQFLSFASL	KVKNEVNSF	QQYAESREL	HMMKDEPVV	

Table 4.9: Peptide fragments of nucleoprotein containing overlapping $CD8^+$ T cell epitopes

	Peptide fragments containing multiple $CD8^+$ T cell epitopes	Length	Epitope Number
CNp1	QKVWMTPSLTESDMDYHKILTAGLSV	26	4
	IVRQRVIPVYQVNNLEEICQLIIQAFEAGV	30	6
	FQESADSFLLMLCLHHAYQGDYKLFLESGAVKYL	34	11
CNp2	KRLEELLPAVSSGRNIKR	18	5
	TEANAGQFLSFASLFLPKL	19	6
	IQVHAEQGLIQY	12	2
	WQSVGHMMVIFRLMRTNFLIKFLLIHQGMHMV	32	8
	ISNSVAQARFSGLLIVK	17	6
	ERGVRLHPLARTAKVKNEVNSFKAAL	26	5
	SSLAKHGEYAPFARLLNLSGV	21	5
	LEHGLFPQLSAIALGVATAHGSTL	24	5
	QQYAESRELDHL	12	2
	QEKKILMNFHQKKNEISFQQTNAMVTLRKERLAKLTEAITAASLPKTSGHY	51	14
	GMSAPDDLVL	10	2
CNp3	GTSNRTPTVAPPAPVY	16	2
	SQGPFDAVLYYHMMKDEPVVFSTSDGKEYTYPDSLEEEYPPWL	43	9
CNp4	AMNDENRFVTLDGQQFY	17	3
	NHRNKFMAIL	10	2

Table 4.10: $CD4^+$ T cell epitopes commonly predicted for nucleoprotein by the three prediction tools

$CD4^+$ T cell epitopes				
YHKILTAGL	LLIHQGMHM	VRLHPLART	LAGVNVGEQ	VLYYHMMKD
IVRQRVIPV	MHMVAGHDA	ILMNFHQKK	LHPLARTAK	FMAILQHHQ
FLLMLCLHH	FKAALSSLA	VKNEVNSFK	FHQKKNEIS	LYYHMMKDE
LLMLCLHHA	LLNLSGVNN	FARLLNLSG	YQSYSENGM	
VKYLEGHGF	FPQLSAIAL	LFPQLSAIA	LIKFLLIHQ	
YQGDYKLFL	VQRQIQVHA	ISFQQTNAM	IHQGMHMVA	
YQVNNLEEI	VNSFKAALS	FQQTNAMVT	LMNFHQKKN	
VQQGIVRQR	LGVATAHGS	LIQYPTAWQ	LMRTNFLIK	
VWMTPSLTE	MVTLRKERL	IQYPTAWQS	LQKTERGVR	
VYQVNNLEE	LVVGEKACL	VIFRLMRTN	IHHASAPLT	
YKLFLESGA	MMVIFRLMR	LLIVKTVLD	FVTLDGQQF	
FLESGAVKY	IFRLMRTNF	LLPAVSSGR	YRHILRSQG	
MLCLHHAYQ	MRTNFLIKF	VGHMMVIFR	MMKDEPVVF	
IPVYQVNNL	LNLSGVNNL	LSGVNNLEH	VTLDGQQFY	
MDYHKILTA	LGLDDQEKK	IKFLLIHQG	YHMMKDEPV	
LMLCLHHAY	FLIKFLLIH	LVLFDLDED	VMNHRNKFM	
FLSFASLFL	FLLIHQGMH	VVGEKACLE	MYRHILRSQ	
MVIFRLMRT	VISNSVAQA	FLPKLVVGE	YTYPDSLEE	
FRLMRTNFL	LIVKTVLDH	YAPFARLLN	YWPVMNHRN	

Table 4.11: Peptide fragments of nucleoprotein containing overlapping $CD4^+$ T cell epitopes

	Peptide fragments containing multiple $CD4^+$ T cell epitopes	Length	Epitope Number
CNp1	MDYHKILTAGL	11	2
	VQQGIVRQRVIPVYQVNNLEEI	22	5
	FLLMLCLHHAYQGDYKLFLESGAVKYLEGHGF	32	8
CNp2	FLSFASLFLPKLVVGEKACLE	21	4
	LIQYPTAWQS	10	2
	VGHMMVIFRLMRTNFLIKFLLIHQGMHMVAGHDA	34	15
	LLIVKTVLDH	10	2
	LQKTERGVRLHPLARTAK	18	3
	VKNEVNSFKAALSSLA	16	3
	YAPFARLLNLSGVNNLEH	18	5
	LFPQLSAIALGVATAHGS	18	3
	ILMNFHQKKNEISFQQTNAMVTLRKERL	28	6
CNp4	MYRHILRSQG	10	2
	VLYYHMMKDEPVVF	14	4
	FVTLDGQQFYWPVMNHRNKFMAILQHHQ	28	5

4.1.3. Identification of peptides containing multiple $CD8^+$ and $CD4^+$ epitopes

Peptides containing $CD8^+$ T cell epitopes and peptides containing $CD4^+$ epitopes were searched for presence of common regions so as to obtain peptide sequences containing multiple $CD8^+$ as well as $CD4^+$ T cell epitopes. In this way, six GP and twelve NP sequences containing multiple $CD8^+$ and $CD4^+$ T cell epitopes were identified (Table 4.12, page 80). The total number of $CD8^+$ and $CD4^+$ epitopes present in each of these peptide sequences are mentioned in Table 4.12, page 80.

4.1.4. Presence of B cell epitopes

The identified peptides were looked for the presence of B cell epitopes via ABCPred tool. Two EBOV GP peptide fragments (VVAFLILPQAKK and YLFEVDNLTYVQLESRFT) were found to be devoid of B cell epitopes while four (FLLMLCLHHAYQGDYKLFLESGAVKYL, VKNEVNSFKAAL, ILMNFHQKKNEISFQQTNAMVTLRKERL and FVTLDGQQFY) of the identified EBOV NP fragments lacked B cell epitopes. The peptides lacking B cell epitopes were not included in further analysis.

4.1.5. Screening for undesirable responses

BLAST analysis was performed to screen for similarity to human proteome which might result in autoimmune responses. Amongst the GP peptides, YLFEVDNLTYVQLESRFT was found to exhibit similarity with human cytoskeletal linker protein epiplakin1 via BLAST analysis. VGHMMVIFRLMRTNFLIKFLLIHQGMHMV and YAPFARLLNLSGV of NP were found to be similar to human Anaphase-promoting complex/cyclosome and homeobox protein, Hox-B9 respectively.

Allergenicity and toxicity of the identified peptides were predicted with the help of AlgPred and ToxinPred. None of the predicted peptides was found to be either allergic or toxic in nature.

Hence, the remaining four GP peptides (Table 4.13, page 81) and six NP peptides (Table 4.14, page 82) containing multiple T and B cell epitopes were considered for further studies.

Table 4.12: Peptides of glycoprotein and nucleoprotein containing multiple $CD8^+$ and $CD4^+$ T cell epitopes

Glycoprotein peptides	Total number of epitopes	Nucleoprotein peptides	Total number of epitopes
FKRTSFFLWVIILFQRTFSIPL	15	FLLMLCLHHAYQGDYKLFLESGAVKYL	14
FHKEGAFFLYDRLASTVIYRGTTFAEG	13	IVRQRVIPVYQVNNLEEI	8
VVAFLILPQAKK	4	MDYHKILTAGL	3
YLFEVDNLTYVQLESRFT	7	FLSFASLFLPKL	4
LANETTQALQLF	9	ERGVRLHPLARTAK	4
RATTELRTFSILNRKAIDF	4	VKNEVNSFKAAL	3
		LFPQLSAIALGVATAHGS	5
		ILMNFHQKKNEISFQQTNAMVTLRKERL	11
		VLYYHMMKDEPVVF	5
		FVTLDGQQFY	4
		VGHMMVIFRLMRTNFLIKFLLIHQGMHMV	20
		YAPFARLLNLSGV	4

Table 4.13: Peptide fragments of glycoprotein containing T and B cell epitopes

Peptide fragments	Epitopes		
	$CD8^+$ T cell	**$CD4^+$ T cell**	**B cell**
FKRTSFFLWVIILFQRTFSIPL (**Gp1**)	KRTSFFLWV RTSFFLWVI FFLWVIILF VIILFQRTF ILFQRTFSI FQRTFSIPL	FFLWVIILF LWVIILFQR FKRTSFFLW WVIILFQRT VIILFQRTF ILFQRTFSI FQRTFSIPL IILFQRTFS FLWVIILFQ	RTSFFLWVII SFFLWVIILF LWVIILFQRT LFQRTFSIPL
FHKEGAFFLYDRLASTVIYRGTTFAEG (**Gp2**)	FHKEGAFFL HKEGAFFLY FLYDRLAST DRLASTVIY RLASTVIYR TVIYRGTTF	FHKEGAFFL YDRLASTVI IYRGTTFAE YRGTTFAEG FFLYDRLAS FLYDRLAST VIYRGTTFA	HKEGAFFLYD GAFFLYDRLA VIYRGTTFAE LASTVIYRGT ASTVIYRGTT FHKEGAFFLY STVIYRGTTF LYDRLASTVI RLASTVIYRG
LANETTQALQLF (**Gp3**)	LANETTQAL NETTQALQL ETTQALQLF	LANETTQAL	NETTQALQLF
RATTELRTFSILNRKAIDF (**Gp4**)	RATTELRTF TELRTFSIL LRTFSILNR RTFSILNRK FSILNRKAI	FSILNRKAI LRTFSILNR ILNRKAIDF	ELRTFSILNR SILNRKAIDF RATTELRTFS LRTFSILNRK RTFSILNRKA TTELRTFSIL FSILNRKAID

Table 4.14: Peptide fragments of nucleoprotein containing T and B cell epitopes

Peptide fragments	Epitopes		
	$CD8^+$ T cell	**$CD4^+$ T cell**	**B cell**
MDYHKILTAGL (**Np1**)	YHKILTAGL	MDYHKILTA YHKILTAGL	DYHKILTAGL
IVRQRVIPVYQVNNLEEI (**Np2**)	IVRQRVIPV VRQRVIPVY IPVYQVNNL YQVNNLEEI	IVRQRVIPV IPVYQVNNL VYQVNNLEE YQVNNLEEI	VYQVNNLEEI RVIPVYQVNN RQRVIPVYQV
FLSFASLFLPKL (**Np3**)	FLSFASLFL SFASLFLPK FASLFLPKL	FLSFASLFL	FLSFASLFLP LSFASLFLPK
ERGVRLHPLARTAK (**Np4**)	ERGVRLHPL GVRLHPLAR	LHPLARTAK VRLHPLART	RLHPLARTAK ERGVRLHPLA VRLHPLARTA
LFPQLSAIALGVATAHGS (**Np5**)	FPQLSAIAL QLSAIALGV	LFPQLSAIA FPQLSAIAL LGVATAHGS	AIALGVATAH LSAIALGVAT FPQLSAIALG ALGVATAHGS QLSAIALGVA
VLYYHMMKDEPVVF (**Np6**)	HMMKDEPVV	VLYYHMMKD LYYHMMKDE YHMMKDEPV MMKDEPVVF	YHMMKDEPVV VLYYHMMKDE HMMKDEPVVF

4.1.6. Conservancy analysis

In a bid to identify cross-protective peptide vaccine candidates which can confer protection against current strains as well as can be relied upon to protect against future viral strains, it is significant to conduct conservancy analysis of the identified peptide fragments. The peptides were looked for conservancy amongst all the human infecting Ebola virus species as well as in other filoviridae members viz., *Marburg virus* and *Lloviu virus*.

All the GP peptides were found to be nearly 100% conserved in *Zaire ebolavirus*, the most pathogenic EBOV species (Table 4.15, page 84). Gp3 was found to be 100% conserved amongst all EBOV species while Gp4 was found to vary by two amino acids in *Sudan ebolavirus*. Gp2 was found to be the most variable of all the identified GP peptides. Further, Gp1 could not be located in *Marburg virus* while other peptides were located in the virus with ample variations (Table 4.15, page 84). *Lloviu virus* contained only Gp2 peptide but with variations.

All the NP peptides were found to be 100% conserved in *Zaire ebolavirus* (Table 4.16, page 85). Np1 exhibited a single amino acid variation w.r.t its first amino acid amongst various EBOV species. Further, Np3 established itself as a potent candidate as it exhibited 100% conservancy in all EBOV species as well as *Lloviu virus* with variation of only one amino acid in *Marburg virus* (Table 4.16, page 85). Amongst the EBOV species, Np5 was found to vary by a single amino acid in Sudan ebolavirus. The Np5 sequence for *Sudan ebolavirus*, *Marburg virus* and *Lloviu virus* was found to be the same. Np1, Np2 and Np6 could not be located in Marburg and Lloviu viruses.

Table 4.15: Conservancy analysis of identified glycoprotein peptides in Ebola virus species and other filoviridae members

Virus species	FKRTSFFLWVIILFQRTFSIPL (Gp1)	FHKEGAFFLYDRLASTVIYRGTTFAEG (Gp2)	LANETTQALQLF (Gp3)	RATTELRTFSILNRKAIDF (Gp4)
ZAIRE	99.39%	100%	100%	100%
SUDAN	FRKSSFFVWVIILFQKAFSMPL	FHKDGAFFLYDRLASTVIYRGVNFAEG	100%	RATTELRTYTILNRKAIDF
TAI FOREST	FRKTSFFVWVIILFHKVFSIPL	FHKEGAFFLYDRLASTIIYRGTTFAEG	100%	100%
BUNDIBUGYO	FRKTSFFVWVIILFHKVFPIPL	FHKEGAFFLYDRLASTIIYRSTTFSEG	100%	100%
MARBURG	--	LHLWGAFFLYDRVASTTMYRGKVFTEG	LANQTAKSLELL	RVTTEERTFSLINRHAIDF
LLOVIU	--	FHKHGSFFLYHGMASTVIYHGVTFTEG	--	--

Table 4.16: Conservancy analysis of identified nucleoprotein peptides in Ebola virus species and other filoviridae members

Virus species	MDYHKILTAGL (Np1)	IVRQRVIPVYQVNNLEEI (Np2)	FLSFASLFLPKL (Np3)	ERGVRLHPLARTAK (Np4)	LFPQLSAIALGVATAHGS (Np5)	VLYYHMMKDEPVVF (Np6)
ZAIRE	100%	100%	100%	100%	100%	100%
*SUDAN	LDYHKILTAGL	IVRQRIIPVYVVSDLEGI/ IVRQRVIPVYVVNDLEGI	100%	DLGVRLHPLARTAK	LYPQLSAIALGVATAHGS	INYYHLMSDEPIAF
TAÏ FOREST	TDYHKILTAGL	IVRQRVIQVHQVTNLEEI	100%	EHGVRLHPLARTAK	100%	ILYYYMMTEEPIVF
BUNDIBUGYO	ADYHKILTAGL	IVRQRIIPVYQISNLEEV	100%	EHGVRLHPLARTAK	100%	ILYYHMMKEEPIIF
LLOVIU	-	-	100%	EAGVQLHPLARTSK	LYPQLSAIALGVATAHGS	-
*MARBURG	-	-	FLSFCSLFLPKL	DSGVTLHPLVRTSK/ DSGVALHPLVRTSK	LYPQLSAIALGVATAHGS	-

* Np2 and Np4 of Sudan and Marburg virus respectively have two variable sequences

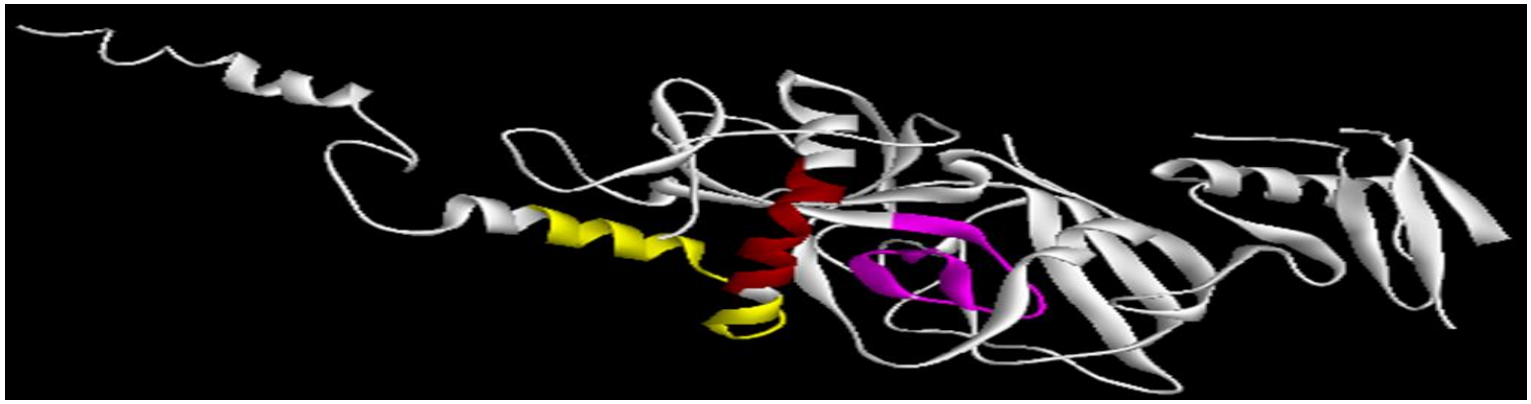
4.1.7. Peptide mapping

The identified GP and NP peptides were located on their respective three dimensional protein structures obtained from protein data bank (PDB) (Figure 4.1 a, page 87; Figure 4.2 a, page 88). All peptides were found to be parts of regions crucial for viral functioning. Further, a schematic representation helped in better understanding of the location of the peptides within EBOV GP and NP (Figure 4.1 b, page 87; Figure 4.2 b, page 88).

The GP peptides viz., Gp2, Gp3 and Gp4 were mapped onto the crystal structure of Ebola glycoprotein (PDB id: 5JQ3) while Gp1 could not be located in the structure owing to missing residues. Further, Gp1 formed a part of signal peptide at the glycoprotein N terminus which directs the nascent glycoprotein to the endoplasmic reticulum where carbohydrates are added on its surface (Feldmann et al., 1994; Lee et al., 2008). Gp2 formed a part of glycoprotein subunit GP1 while Gp3 and Gp4 were found to be located in the heptad repeat regions (HR1 and HR2) of glycoprotein subunit, GP2 (Figure 4.1 b, page 87).

NP peptides (except Np1 and Np6) were mapped onto the crystal structure of Ebola nucleoprotein (PDB id: 4YPI). Further, Np1-Np5 formed a part of N terminal 1-450 residues, a polypeptide which in itself is sufficient for viral genome replication (Watanabe et al., 2006). A part of Np1 was located in N-tail while the rest of the peptide, along with Np2-Np5 peptides, was found to be located in the core domain of viral nucleoprotein. Np6 was found to be a part of C tail (Figure 4.2 b, page 88).

a)



b)

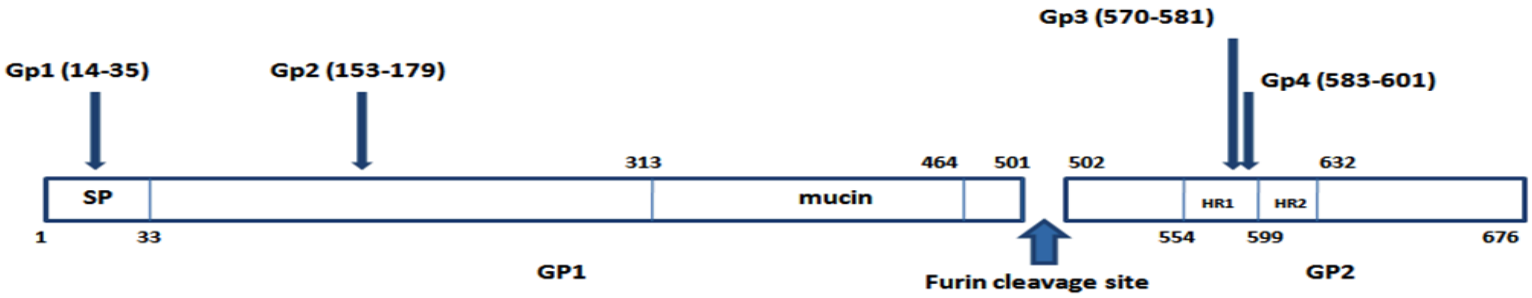
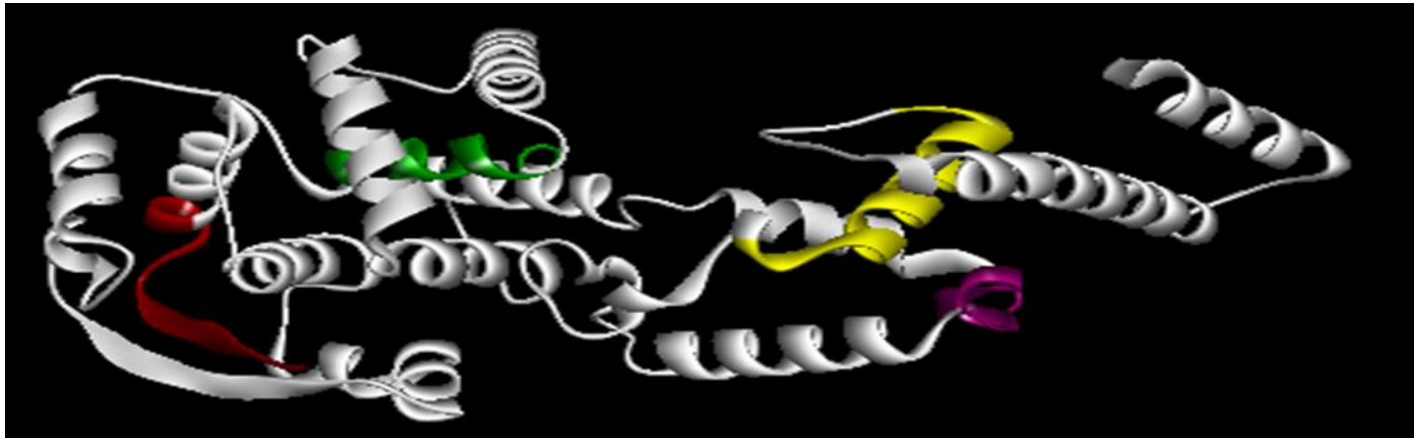


Figure 4.1: a) Mapping of identified peptide fragments onto the crystal structure of Ebola glycoprotein (PDB id: 5JQ3). Gp2, Gp3 and Gp4 are depicted in pink, red and yellow respectively. Gp1 could not be located in the structure owing to missing residues. b) Schematic presentation of peptides in different regions of Ebola glycoprotein.

a)



b)

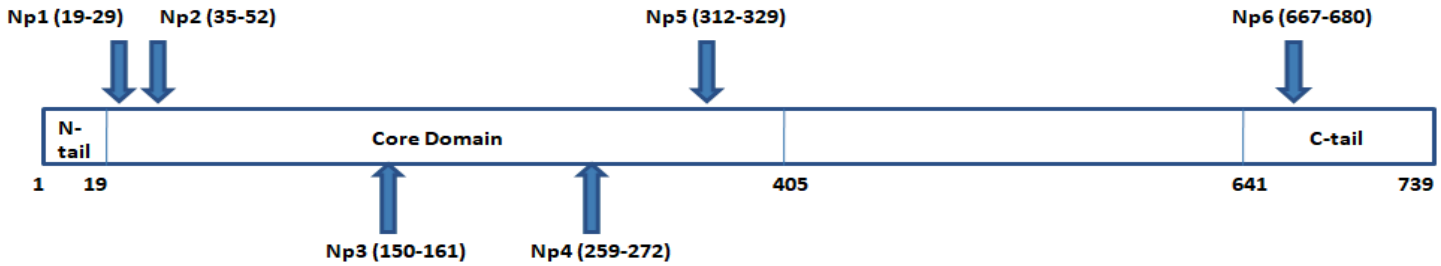


Figure 4.2: a) Mapping of identified peptide fragments onto the crystal structure of Ebola nucleoprotein (PDB id: 4YPI). Np2, Np3, Np4 and Np5 are represented in red, green, pink and yellow respectively b) Schematic presentation of peptides in different regions of Ebola nucleoprotein.

Objective 2

4.2. Peptide-HLA interaction analysis

Human leukocyte antigen (HLA) molecules are present on the surface of various cells and are a result of translation of mRNA transcribed from a cluster of approximately 200 genes present on the 6^{th} chromosome. HLA class I molecules are located on the cell surface of all nucleated cells and present epitope to T_c ($CD8^+$) cells while HLA class II molecules are found specifically on cell surface of antigen presenting cells (APC) and present epitope to T_h ($CD4^+$) cells. Therefore, the study of peptide-HLA interaction is of utmost significance in the prediction of peptide vaccine candidates with the ability to generate an effective cell mediated immune response.

In the current study, peptide-HLA interaction studies were carried out in three phases viz., HLA coverage analysis to assess the promiscuous nature of identified peptides, population coverage analysis to assess the potential of predicted peptides as global vaccine candidates and molecular docking analysis carried out via Autodock Vina and CABS-dock.

4.2.1. HLA coverage analysis

T cell epitope prediction tools help in predicting epitopes corresponding to particular HLA alleles/supertypes. A supertype is a cluster of functionally related HLA alleles that share binding specificities towards the same panel of peptides owing to similar structural features of HLAs peptide binding groove (Lund et al., 2004). The HLA alleles/supertypes predicted by the six T cell epitope prediction tools for $CD8^+$ and $CD4^+$ T cell epitopes present in the finally selected peptide fragments were listed for both, EBOV GP and NP (Table 4.17 & Table 4.18, page 91).

Amongst the GP peptides, Gp1 was predicted for the maximum number of HLA class I alleles/supertypes by the three different prediction tools while Gp2 showed coverage for maximum HLA class II alleles/supertypes (Table 4.17, page 91). Gp3 showed the least class I and class II coverage while Gp4 exhibited an average HLA coverage for both the classes (Table 4.17, page 91).

Amongst the NP peptides, Np2 displayed the best HLA coverage for both, class I and II according to all the epitope prediction tools. Np6 displayed the least HLA class I coverage while Np3 and Np5 were jointly lowest in HLA class II coverage (Table 4.18, page 91).

Further, epitopes present in Np1 and Np6 were not predicted to bind to HLA-A and HLA-B respectively by any of the HLA class I T cell epitope prediction tools (Table 4.18, page 91).

Overall, Gp1 exhibited the maximum HLA class I coverage amongst all the identified peptides where as Np2 lead the HLA class II predictions. Np2 emerged as the peptide with maximum HLA coverage when both HLA classes were considered together (Table 4.17 & Table 4.18, page 91). Interestingly, the four GP peptide fragments were predicted for a greater number of HLA class I alleles/sypertypes by the six epitope prediction tools than the six NP peptide fragments (Table 4.17 & Table 4.18, page 91).

Table 4.17: Multiple eptiope containing peptides of Ebola virus glycoprotein exhibited binding to diverse HLA supertypes/alleles

	Class I						Class II					
Tools	**SYFPEITHI**		**NetCTL 1.2**		**IEDB**		**ProPred**	**MHC2Pred**		**IEDB**		
HLA type	**A**	**B**	**A**	**B**	**A**	**B**	**DR**	**DQ**	**DR**	**DP**	**DQ**	**DR**
Number*	7	26	5	7	18	32	51	11	27	5	6	14
FKRTSFFLWVIILFQRTFSIPL (Gp1)	3	7	3	6	9	8	47	1	3	4	4	10
FHKEGAFFLYDRLASTVIYRGTTFAEG (Gp2)	6	6	5	6	7	5	47	5	9	5	3	12
LANETTQALQLF (Gp3)	2	6	1	4	1	7	3	1	-	5	-	4
RATTELRTFSILNRKAIDF (Gp4)	2	6	1	5	6	8	36	4	7	3	1	10

*Total number of HLA alleles/supertypes in each tool

Table 4.18: Multiple eptiope containing peptides of Ebola virus nucleoprotein exhibited binding to diverse HLA supertypes/alleles

	Class I						Class II					
Tools	**SYFPEITHI**		**NetCTL 1.2**		**IEDB**		**ProPred**	**MHC2Pred**		**IEDB**		
HLA type	**A**	**B**	**A**	**B**	**A**	**B**	**DR**	**DQ**	**DR**	**DP**	**DQ**	**DR**
Number*	7	26	5	7	18	32	51	11	27	5	6	14
MDYHKILTAGL (Np1)	-	4	-	2	-	1	22	5	5	-	1	10
IVRQRVIPVYQVNNLEEI (Np2)	1	5	1	5	5	6	51	9	18	3	2	14
FLSFASLFLPKL (Np3)	3	2	3	1	9	-	18	2	4	5	1	7
ERGVRLHPLARTAK (Np4)	3	5	1	1	2	3	24	-	7	-	1	7
LFPQLSAIALGVATAHGS (Np5)	2	4	1	3	3	6	15	3	6	1	2	10
VLYYHMMKDEPVVF (Np6)	1	-	1	-	2	-	18	4	7	4	2	13

*Total number of HLA alleles/supertypes in each tool

4.2.2. Population coverage analysis

Population coverage analysis IEDB tool provides an insight into the geographical ethnicity/ population probable to react to a given epitope set. All the HLA alleles predicted for respective peptide fragments by the epitope prediction tools were considered and analysis was carried out for four different continents (Africa, America, Asia and Europe) consisting of 14 geographical areas and whole world.

GP peptide fragments exhibited a 77-100% population coverage for various continents while a tremendous coverage of 96-100% was found for the whole world (Figure 4.3, page 93). Individually, Gp1, Gp2, Gp3 and Gp4 exhibited an average percentage population coverage of 95, 96, 91 and 94 respectively for all continents.

In case of EBOV NP, peptide fragments exhibited a 69-100% population coverage for various continents while a coverage of 90-100% was found for the whole world (Figure 4.4, page 93). Majority of peptides exhibited >95% population coverage for all continents except the African continent. Individually, Np1, Np2, Np3 and Np4, Np5 and Np6 exhibited an average percentage population coverage of 94.7, 96, 95.4, 85, 95 and 94.3 respectively for all continents.

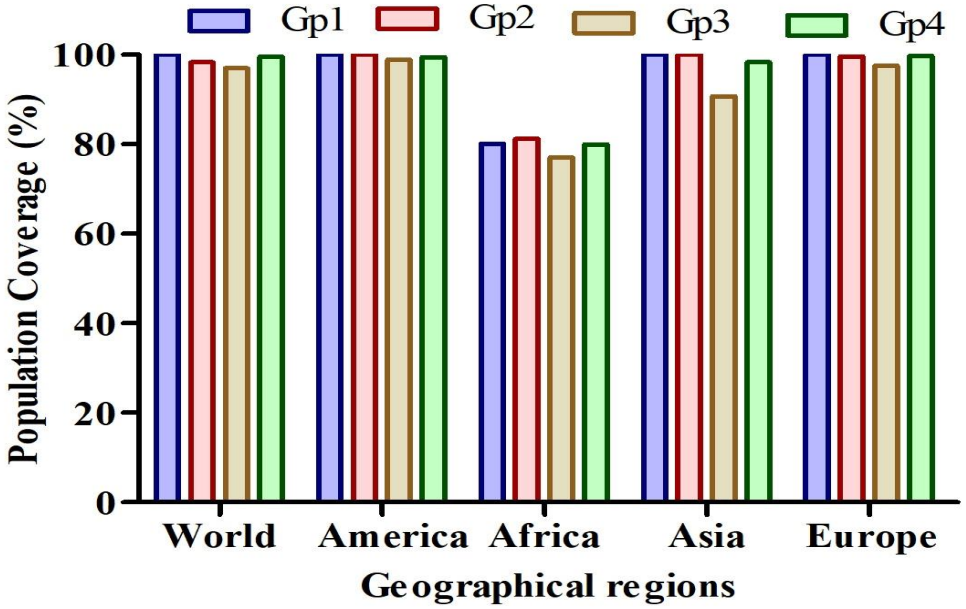


Figure 4.3: Population coverage analysis of peptides belonging to glycoprotein

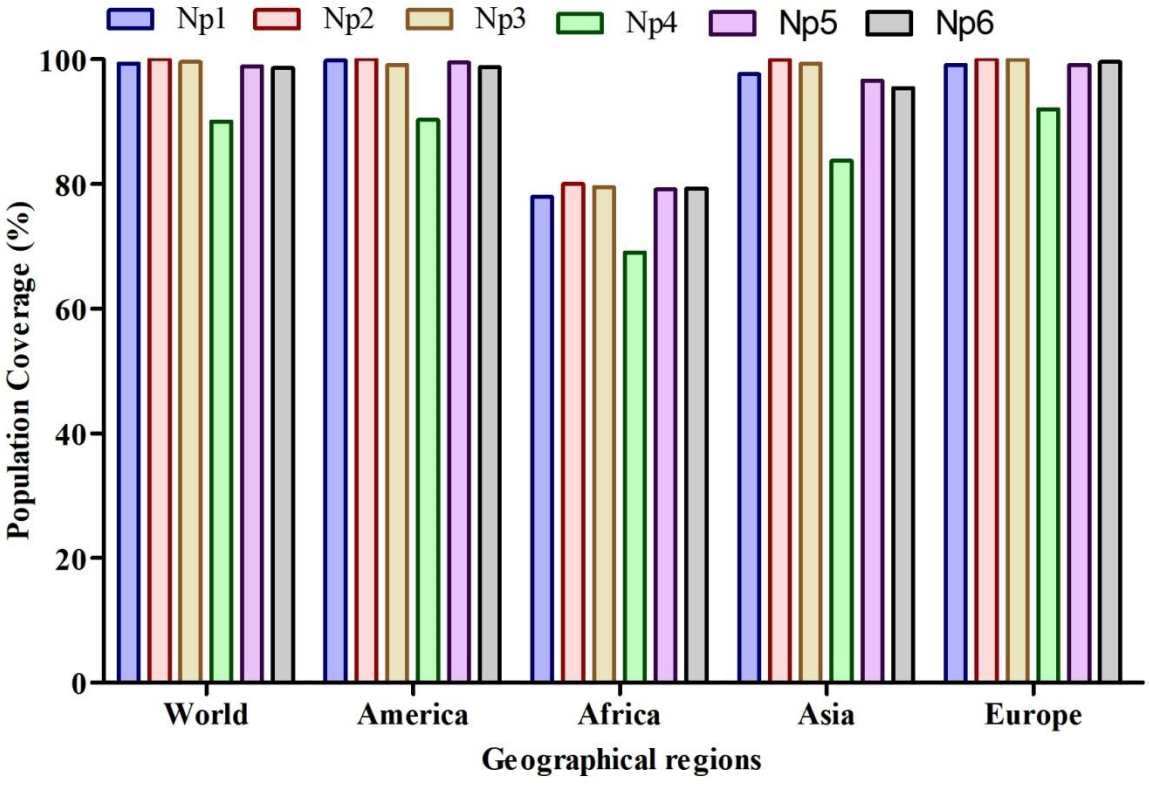


Figure 4.4: Population coverage analysis of peptides belonging to nucleoprotein

4.2.3. Molecular docking

The first and foremost step towards the generation of an efficient adaptive immune response is the binding of immunogenic peptide to the peptide binding groove of HLA molecule present on the surface of host cell. Molecular docking was employed to study the interaction of the predicted peptides containing epitopes with various HLA class I and II molecules. High resolution crystallographic structures of peptide bound HLA molecules (nine each for HLA class I and II typifying different HLA molecules) were obtained from the PDB database for docking. Autodock Vina and CABS-dock tools were employed to calculate the binding energy and root mean square deviation (RMSD) value respectively to assess the binding affinity of the selected peptides containing multiple epitopes with HLA class I and II molecules. Generally, a lower binding energy and RMSD value indicates a stable interaction with a good binding pose (Dhiman et al., 2016; Blaszczyk et al., 2016). Binding energy and RMSD values obtained by docking the natural peptides separated from the parent peptide-HLA complex with the corresponding HLA molecules served as positive control.

HLA class II accommodates longer peptides (~18–20 amino acid) in contrast to HLA class I which binds to smaller peptides (8-10 amino acid) (Andreatta and Nielsen, 2016). Accordingly, $CD8^+$ T cell epitopes present in the peptide fragments were docked with HLA class I molecules while the selected peptide fragments were docked as such with HLA class II molecules. For Autodock Vina, PEP-FOLD was used to generate structure of $CD8^+$ T cell epitopes and the selected peptide fragments. In case of CABS-dock, peptide structure generation is not necessary and hence, the peptide sequence and PDB file of the receptor were provided to the online server to find the RMSD value.

4.2.3.1. Binding energy and RMSD values obtained by docking the peptides with HLA class I molecules

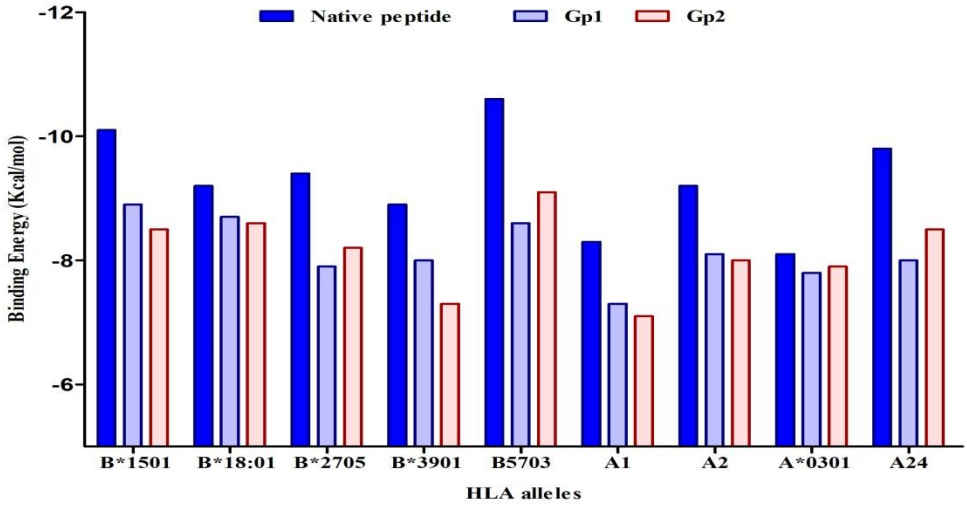
Twenty and thirteen $CD8^+$ T cell epitopes were present in the selected GP and NP peptide fragments respectively and they were docked with the HLA class I molecules. The mean of the binding energy and RMSD obtained for all epitopes belonging to a particular peptide fragment was considered and used for analysis.

4.2.3.1.1. Ebola glycoprotein

The mean binding energy (Autodock Vina) and RMSD (CABS-dock) values of identified peptide fragments for HLA class I molecules were found to be comparable to the values obtained for native peptide (Figure 4.5, page 96; Figure 4.6, page 97). Interesting results were obtained in case of CABS-dock as several peptide-HLA complex exhibited better mean RMSD value than the corresponding native peptide-HLA complex. Gp1, Gp3 and Gp4 scored a RMSD value of 2.4, 1.8 and 2.1 respectively as compared to native peptide (3.7) for B*1801 while they scored 1.9, 1.8 and 2.1 respectively as compared to native peptide (3.3) for B*3901 (Figure 4.6, page 97). Gp1 showed a better RMSD value of 1.4 than native peptide (2.0) for A1 allele while Gp4 (1.6) outperformed native peptide (2.4) for A*0301 allele.

A detailed statistical analysis of the results was carried out by considering all HLA molecules for each peptide (Table 4.19, page 98). One way ANOVA analysis followed by Tukey's multiple comparison test indicated no significant difference between the performances of selected peptides and native peptide in case of CABS-dock results. Although significant variation was observed in Autodock Vina results but the range of mean binding energies considering all HLA class I molecules for each peptide is -7.54 to -8.14 which is not much deviated from -9.28 of native peptide (Table 4.19, page 98). Skewness was also calculated where positive skewness indicates a stable interaction between a greater number of HLA molecules and selected peptides, normal distribution where skewness = 0 and negative skewness which indicates a highly stable interaction between a few HLA molecules and selected peptides. Most skewness values for HLA class I molecules were found to be either positive or close to zero for both, mean binding energy as well as mean RMSD. This indicated a relatively uniform stable interaction between GP epitopes belonging to all the selected peptides with majority of HLA class I molecules (Table 4.19, page 98).

a)



b)

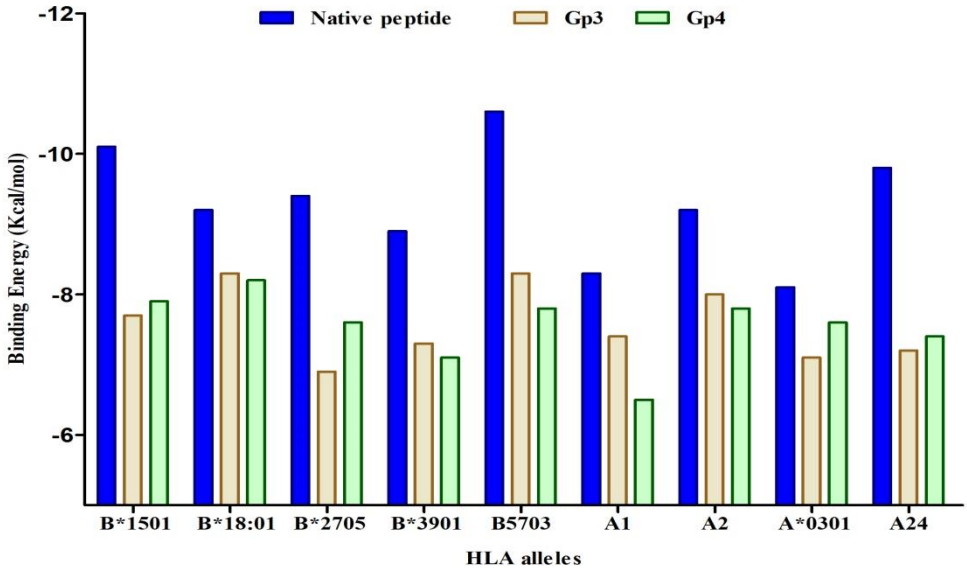
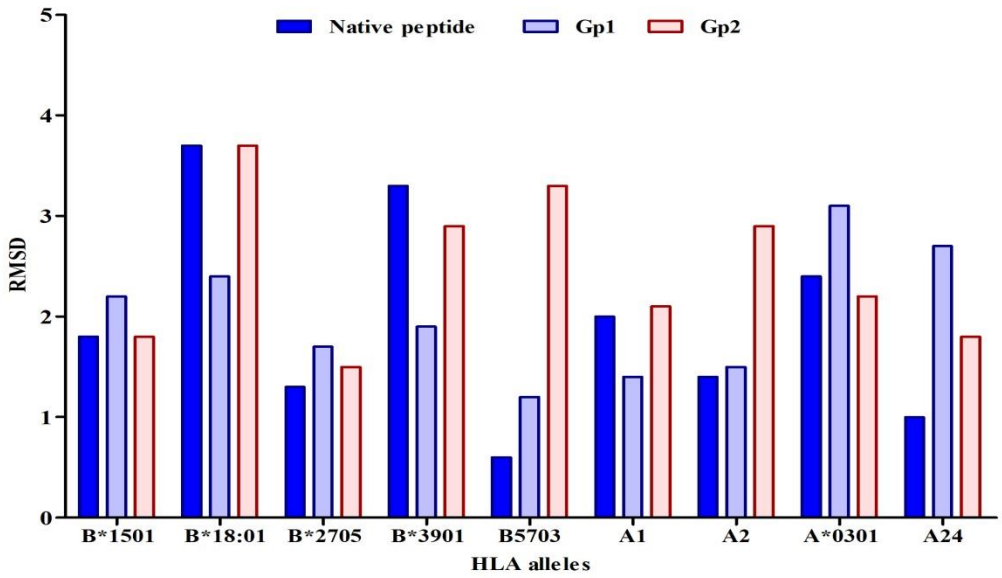


Figure 4.5: Mean binding energies for the identified glycoprotein peptide fragments against HLA class I molecules. a) Gp1 and Gp2 b) Gp3 and Gp4

a)



b)

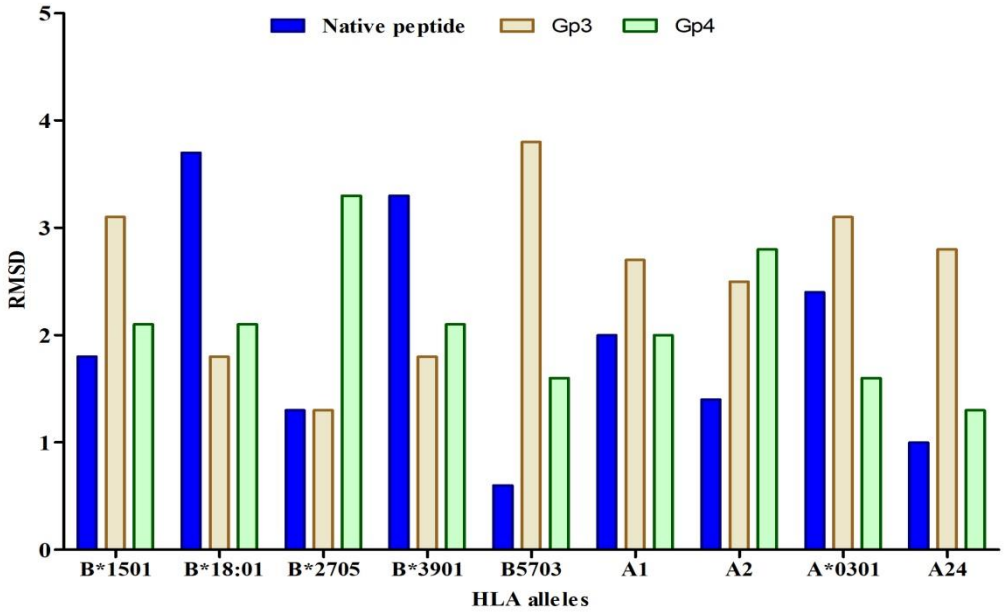


Figure 4.6: Mean RMSD values for the identified glycoprotein peptide fragments against HLA class I molecules. a) Gp1 and Gp2 b) Gp3 and Gp4

Table 4.19: Comparative analysis of mean binding energies (Autodock Vina) and RMSD (CABS-dock) values of glycoprotein peptides with native peptide for HLA class I molecules

	Minimum	25% Percentile	Median	75% Percentile	Maximum	Mean	SD	Skewness	Significant $p < 0.05$
Binding energy (Kcal/mol)									
Native peptide	-10.6	-9.95	-9.2	-8.6	-8.1	-9.28	0.80	-0.06	
Gp1	-8.9	-8.65	-8	-7.85	-7.3	-8.14	0.50	-0.04	*
Gp2	-9.1	-8.55	-8.2	-7.6	-7.1	-8.13	0.63	0.36	*
Gp3	-8.3	-8.15	-7.4	-7.15	-6.9	-7.57	0.52	-0.37	*
Gp4	-8.2	-7.85	-7.6	-7.25	-6.5	-7.54	0.50	1.11	*
Root mean square deviation (RMSD)									
Native peptide	0.6	1.15	1.8	2.85	3.7	1.94	1.03	0.62	
Gp1	1.2	1.45	1.9	2.55	3.1	2.01	0.63	0.46	ns
Gp2	1.5	1.8	2.2	3.1	3.7	2.46	0.75	0.38	ns
Gp3	1.3	1.8	2.7	3.1	3.8	2.54	0.78	-0.13	ns
Gp4	1.3	1.6	2.1	2.45	3.3	2.1	0.62	0.88	ns

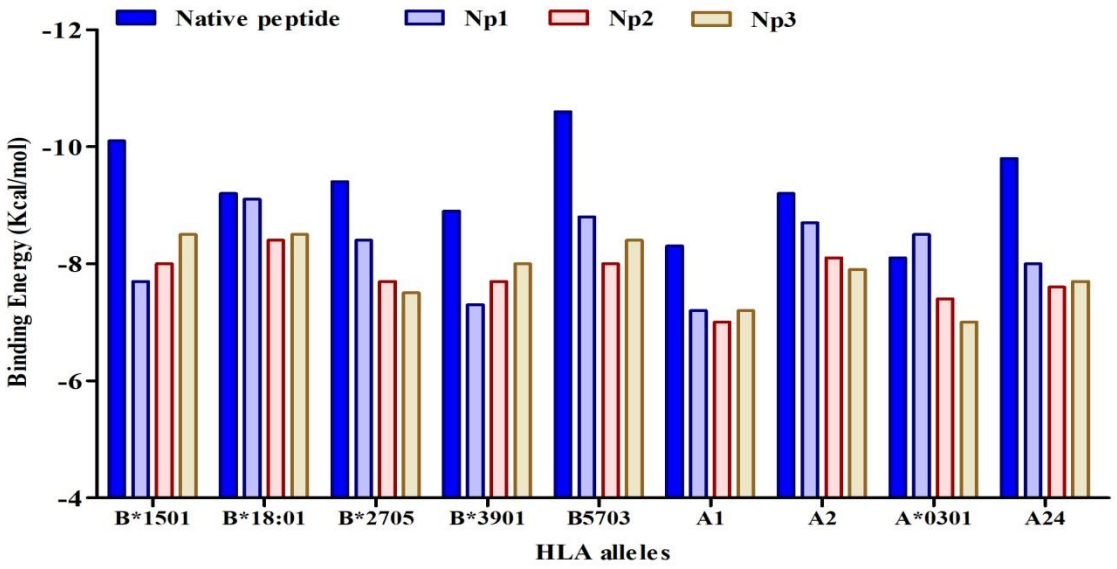
The mean of the binding energy and RMSD obtained for all epitopes belonging to a particular peptide fragment was considered as the binding energy and RMSD respectively for that peptide and these values have been depicted and analysed via skewness and one way ANOVA analysis followed by Tukey's multiple comparison test. Significant differences are indicated by * ($p<0.05$) where as ns stands for non-significant.

4.2.3.1.2. Ebola nucleoprotein

The mean binding energy and RMSD of each peptide for HLA class I molecules was plotted (Figure 4.7, page 100; Figure 4.8, page 101). Encouragingly, Np1 (-8.5 Kcal/mol) displayed better mean binding energy than native peptide (-8.1 Kcal/mol) for HLA A*0301 (Figure 4.7, page 100). All peptides viz., Np1 (1.2), Np2 (0.9), Np3 (2.2), Np4 (1.0), Np5 (2.7) and Np6 (2.1) displayed a better mean RMSD value for B*3901 allele than native peptide (3.3). Np1 was found to have a mean RMSD value of 0.6, superior to native peptide value (1.8) for B*1501 allele while both, Np1 (1.1) and Np6 (0.2) displayed better results than native peptide (1.4) for A2 allele. All peptides except Np2 exhibited better mean RMSD value for A1 molecule than native peptide (Figure 4.8, page 101). Peptide-HLA interactions either exhibiting an RMSD >5 or with peptide binding outside the HLA binding groove are considered as poor (Blaszczyk et al., 2016; Lohia and Baranwal, 2015) and hence, the RMSD values for Np4-A*0301, Np6-A*0301 and Np5-A24 peptide-HLA complexes were not plotted (Figure 4.8, page 101).

One way ANOVA analysis followed by Tukey's multiple comparison test indicated no significant difference in the mean RMSD values of peptides and native peptides for CABS-dock results (Table 4.20, page 102) and even though significant variations were found for Autodock Vina results, three peptides viz., Np1, Np5 and Np4 displayed a mean binding energy of ≥ -8 Kcal/mol for six (B*1801, B*2705, B*5703, A2, A*0301 and A24), six (B*1501, B*1801, B*2705, B*3901, B*5703 and A2) and five (B*1501, B*2705, B*3901, B*5703 and A24) HLA molecules respectively and the range of mean binding energies considering all HLA class I molecules for each peptide was found to be -7.32 to -8.18 which is not much deviated from -9.28 of native peptide typifying desirable interactions between majority of peptide-HLA complex (Table 4.20, page 102). Moreover, positive or normal skewness of all peptides (except Np5 and Np6) for HLA class I molecules was found for Autodock Vina results which further indicates stable interactions of the epitopes with majority of class I molecules. In case of CABS-dock results, positive or close to zero skewness was found for three (Np1, Np3 and Np5) peptides indicating favourable interactions with a high percentage of HLA class I molecules while Np2, Np4 and Np6 were found to interact in a highly stable and superior manner for some of the molecules (Table 4.20, page 102).

a)



b)

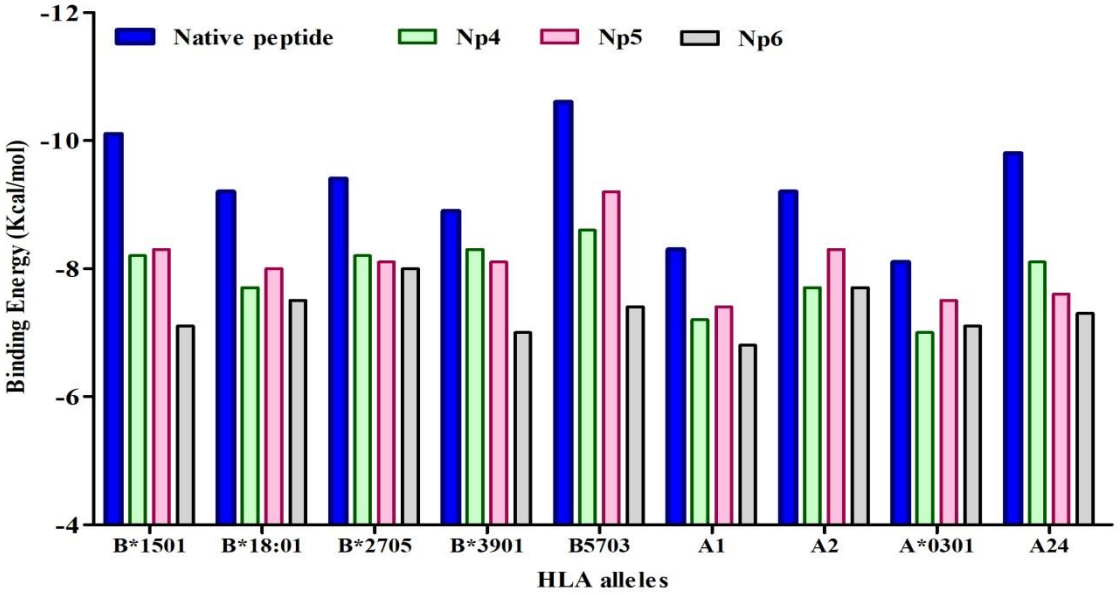
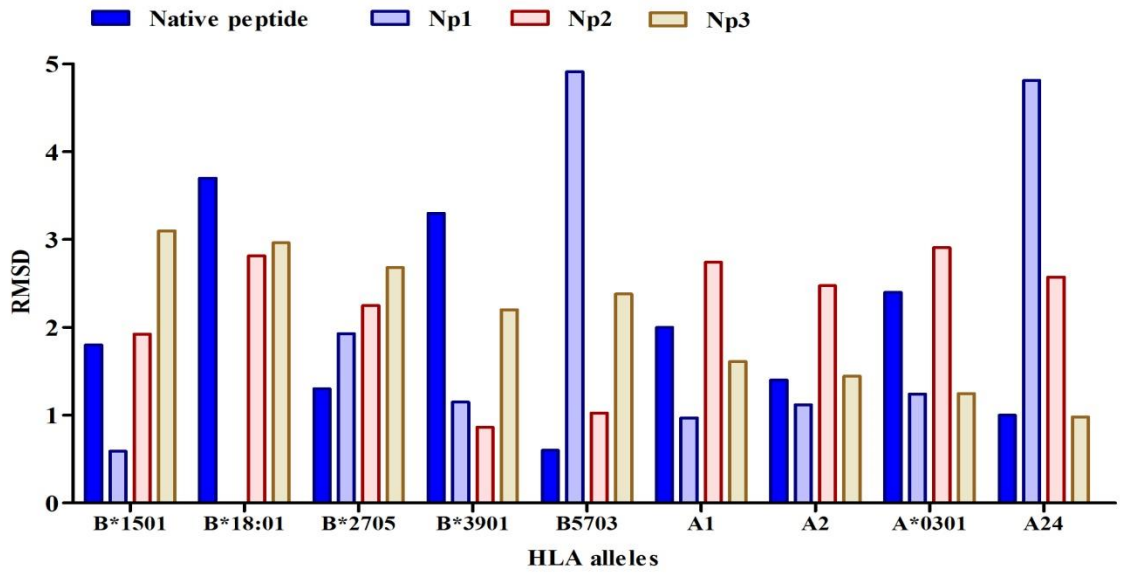


Figure 4.7: Mean binding energies for the identified nucleoprotein peptide fragments against HLA class I molecules. a) Np1, Np2 and Np3 b) Np4, Np5 and Np6

a)



b)

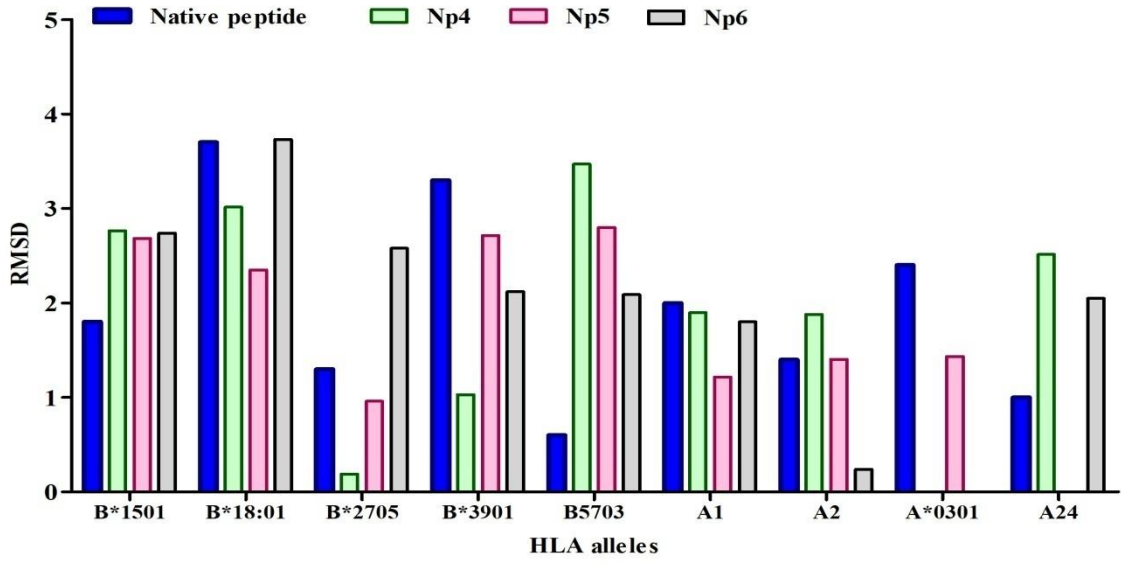


Figure 4.8: Mean RMSD values for the identified nucleoprotein peptide fragments against HLA class I molecules. a) Np1, Np2 and Np3 b) Np4, Np5 and Np6

Table 4.20: Comparative analysis of mean binding energies (Autodock Vina) and RMSD (CABS-dock) values of nucleoprotein peptides with native peptide for HLA class I molecules

	Minimum	25% Percentile	Median	75% Percentile	Maximum	Mean	SD	Skewness	Significant $p < 0.05$
Binding energy (Kcal/mol)									
Native peptide	-10.6	-9.95	-9.2	-8.6	-8.1	-9.28	0.80	-0.06	
Np1	-9.1	-8.75	-8.4	-7.5	-7.2	-8.18	0.67	0.33	*
Np2	-8.4	-8.05	-7.7	-7.5	-7	-7.76	0.41	0.42	*
Np3	-8.5	-8.45	-7.9	-7.35	-7	-7.85	0.55	0.24	*
Np4	-8.6	-8.25	-8.1	-7.45	-7	-7.88	0.53	0.58	*
Np5	-9.2	-8.3	-8.1	-7.55	-7.4	-8.05	0.54	-0.96	*
Np6	-8	-7.6	-7.3	-7.05	-6.8	-7.32	0.37	-0.53	*
Root mean square deviation (RMSD)									
Native peptide	0.60	1.15	1.80	2.85	3.70	1.94	1.03	0.62	
Np1	0.60	1.02	1.20	4.07	4.90	2.08	1.74	1.27	ns
Np2	0.90	1.45	2.50	2.75	2.90	2.16	0.75	-1.01	ns
Np3	1.0	1.30	2.20	2.85	3.10	2.06	0.79	-0.01	ns
Np4	0.20	1.40	1.90	2.90	3.50	2.06	1.02	-0.52	ns
Np5	1.0	1.25	1.90	2.70	2.80	1.95	0.76	-0.02	ns
Np6	0.20	1.87	2.10	2.67	3.70	2.16	0.98	-0.73	ns

The mean of the binding energy and RMSD obtained for all epitopes belonging to a particular peptide fragment was considered as the binding energy and RMSD respectively for that peptide and these values have been depicted and analysed via skewness and one way ANOVA analysis followed by Tukey's multiple comparison test. Significant differences are indicated by * ($p<0.05$) where as ns stands for non-significant.

4.2.3.2. Binding energy and RMSD values obtained by docking the peptides with HLA class II molecules

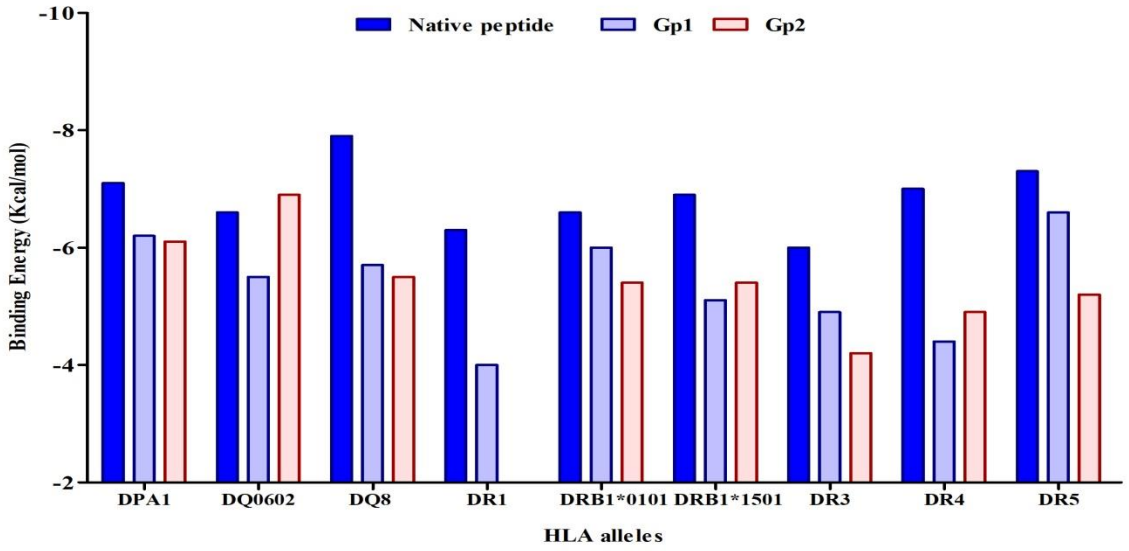
The structure of the four GP and six NP peptide fragments generated via PEP-FOLD was docked as such with the selected HLA class II molecules to obtain binding energy via Autodock Vina. The peptide sequence was provided to the online server to obtain RMSD value via CABS-dock.

4.2.3.2.1. Ebola glycoprotein

The binding energy and RMSD values for all peptides were plotted against HLA class II molecules (Figure 4.9, page 104; Figure 4.10, page 105). Gp3 was found to outperform native peptides for multiple alleles viz., DR3, DRB1*1501, DR1 and DQ8 while Gp4 exhibited better results for DRB1*0101 in terms of binding energy (Figure 4.9, page 104). Gp1 exhibited a superior RMSD value of 2.2, 0.2, 0.3, 1.4, 0.8, 2.0 and 2.2 than native peptide for DP1 (2.4), DQ0602 (1.5), DQ8 (3.0), DRA1 (2.2), DRB1 (1.8), DR2 (3.4) and DR4 (3.2) respectively (Figure 4.10, page 105). Gp2 exhibited positive binding energy for DR1 and showed poor CABS-dock results for 5 (DQ8, DRB1, DR2, DR4 and DR5) HLA class II alleles, proving to be either a non-binder or exhibiting an RMSD value >5. Therefore, these results were not plotted (Figure 4.9, page 104; Figure 4.10, page 105). An appreciably better RMSD value was found for Gp3 and Gp4 in case of 5 (DP1, DQ8, DRA1, DRB1 and DR2) and 3 (DP1, DQ0602 and DR5) HLA alleles respectively.

Gp3 displayed the best results in all statistical categories for HLA class II molecules. Also, Gp1 and Gp3 displayed a better mean RMSD value than native peptide when all the HLA class II molecules were considered (Table 4.21, page 106). One way ANOVA analysis followed by Tukey's multiple comparison test indicated a significant variation for Autodock Vina amongst the values of peptides when compared to native peptides (except for Gp3) while no significant variation was found in case of CABS-dock results (Table 4.21, page 106). Gp1 and Gp4 showed positive skewness for Autodock Vina results indicating highly stable interaction with majority of HLA class II molecules. Gp2 skewness value was found to be comparable to native peptide value. In case of CABS-dock, skewness values indicated that Gp3 displayed favourable interaction with majority of class II molecules while the other peptides displayed lesser desirable interactions with some of the molecules (Table 4.21, page 106).

a)



b)

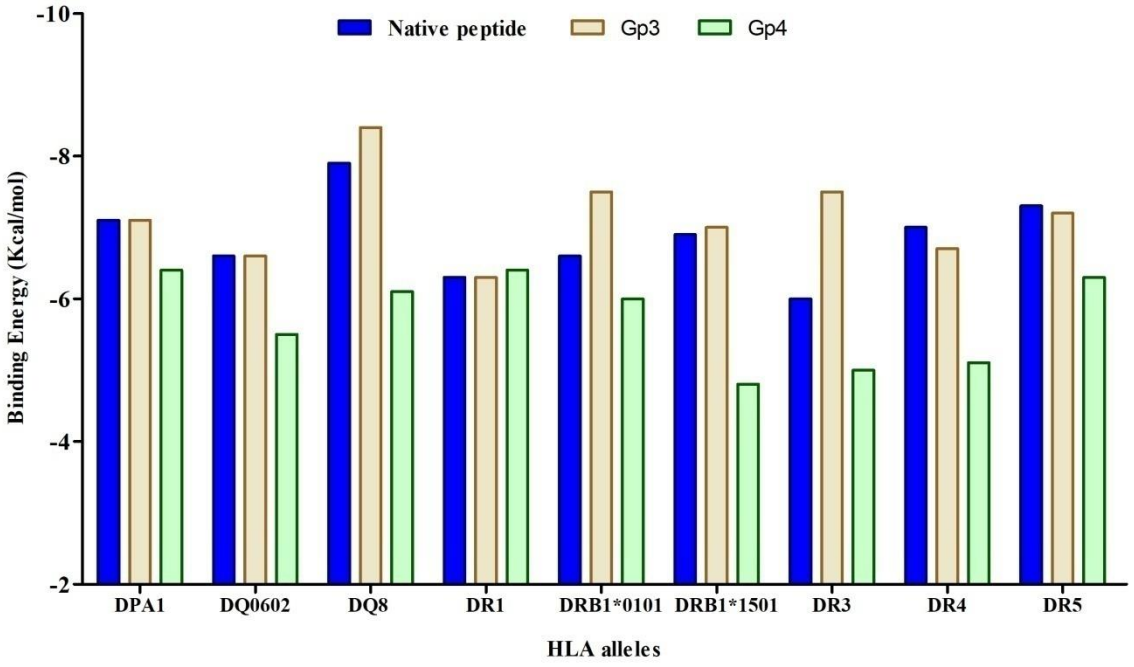
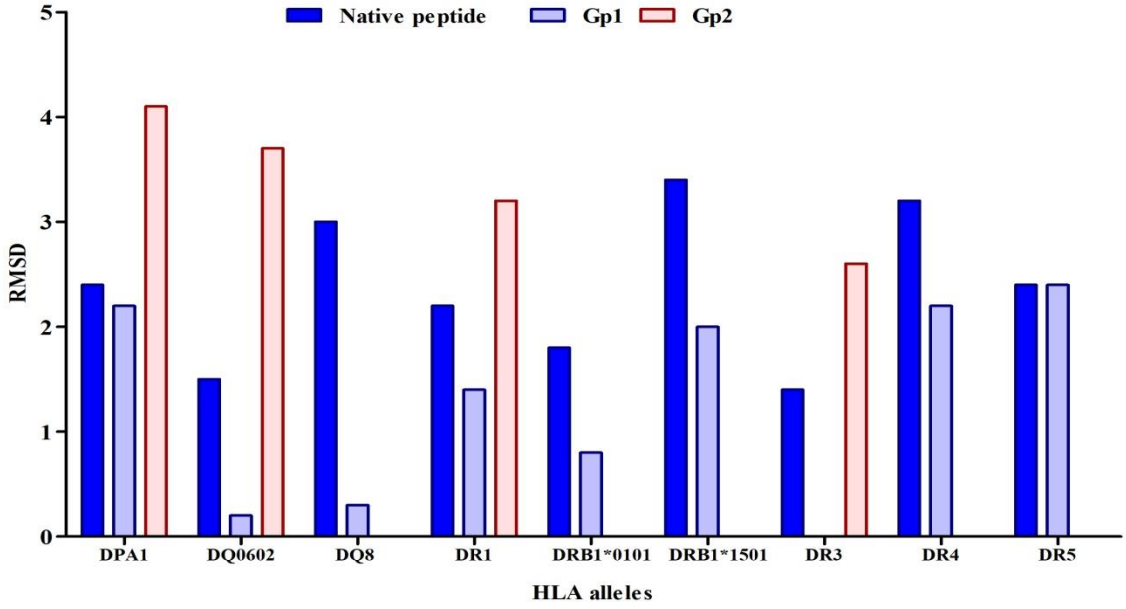


Figure 4.9: Binding energies for the identified glycoprotein peptide fragments against HLA class II molecules. a) Gp1 and Gp2 b) Gp3 and Gp4

a)



b)

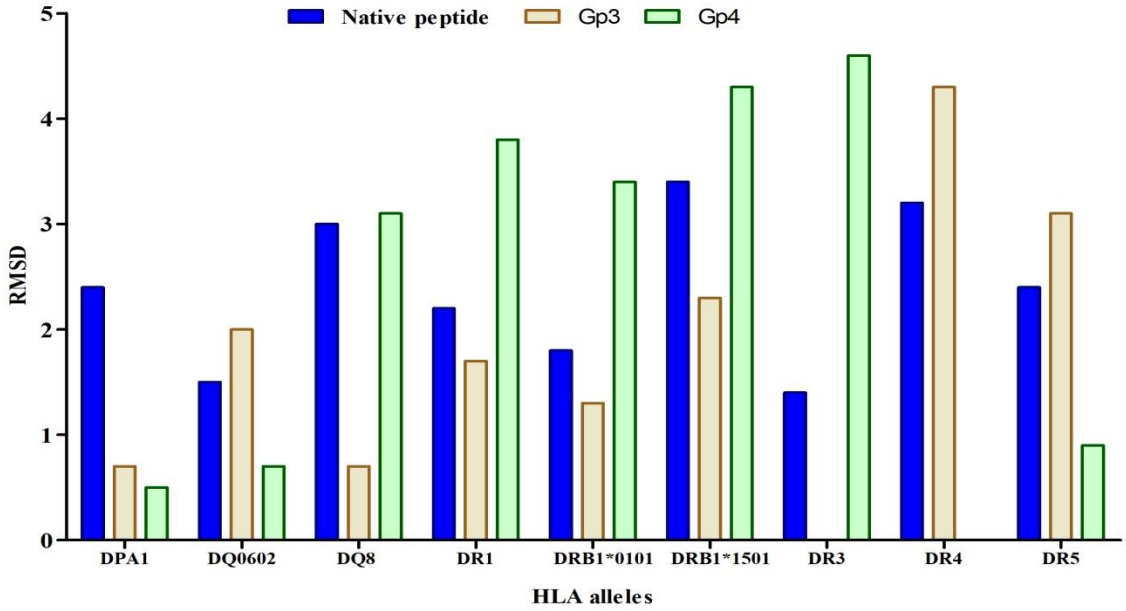


Figure 4.10: RMSD values for the identified glycoprotein peptide fragments against HLA class II molecules. a) Gp1 and Gp2 b) Gp3 and Gp4

Table 4.21: Comparative analysis of mean binding energies (Autodock Vina) and RMSD (CABS-dock) values of glycoprotein peptides with native peptide for HLA class II molecules

	Minimum	25% Percentile	Median	75% Percentile	Maximum	Mean	SD	Skewness	Significant p < 0.05
Binding energy (Kcal/mol)									
Native peptide	-7.9	-7.2	-6.9	-6.45	-6	-6.85	0.56	-0.37	
Gp1	-6.6	-6.1	-5.5	-4.65	-4	-5.37	0.85	0.26	*
Gp2	-6.9	-5.95	-5.4	-4.97	-4.2	-5.45	0.79	-0.44	*
Gp3	-8.4	-7.5	-7.1	-6.65	-6.3	-7.14	0.61	-0.82	ns
Gp4	-6.4	-6.35	-6	-5.05	-4.8	-5.73	0.64	0.38	*
Root mean square deviation (RMSD)									
Native peptide	1.4	1.65	2.4	3.1	3.4	2.36	0.72	0.08	
Gp1	0.2	0.42	1.7	2.2	2.4	1.43	0.89	-0.43	ns
Gp2	2.6	2.75	3.45	4	4.1	3.4	0.64	-0.36	ns
Gp3	0.7	0.85	1.85	2.9	4.3	2.01	1.22	0.86	ns
Gp4	0.5	0.75	3.25	4.17	4.6	2.66	1.69	-0.37	ns

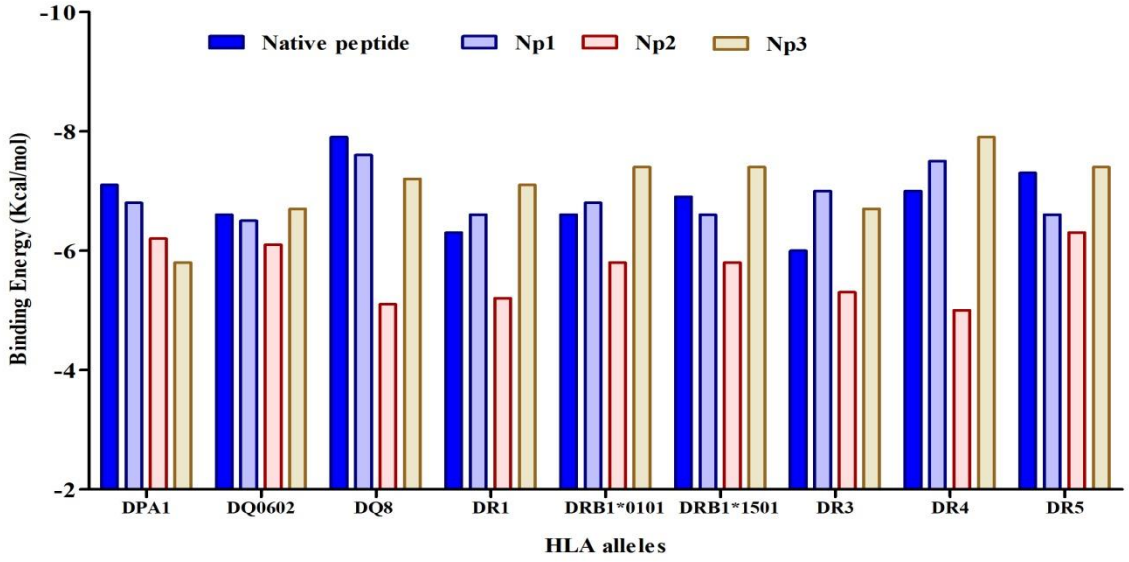
The mean of the binding energy and RMSD values obtained for all peptide fragments have been depicted and analysed via skewness and one way ANOVA analysis followed by Tukey's multiple comparison test. Significant differences are indicated by * ($p<0.05$) where as ns stands for non-significant.

4.2.3.2.2. Ebola nucleoprotein

Binding energy and RMSD values of each peptide for HLA class II molecules were plotted (Figure 4.11, page 108; Figure 4.12, page 109). Highly encouraging results were obtained as Np1, Np3, Np4 and Np5 outperformed native peptide for 4 (DR3, DR4, DR1 and DRB1*1501), 7 (DQ0602, DR1, DRB1*1501, DRB1*0101, DR3, DR4 and DR5), 4 (DR3, DRB1*0101, DR1 and DPA1) and 3 (DQ0602, DR1 and DPA1) HLA alleles respectively in case of Autodock Vina (Figure 4.11, page 108). All peptides, viz., Np1 (2.7), Np2 (1.9), Np3 (1.4), Np4 (3.0), Np5 (0.9) and Np6 (2.8) displayed a better RMSD value for DR4 allele than native peptide (3.2). Np5 (1.3) and Np6 (1.1) outperformed native peptide for DQ0602 (1.5) and DRB1 (1.8) molecules respectively. Np4 and Np5 performed better in case of DR5 molecule while Np1 and Np4 outperformed in case of DR3 molecule (Figure 4.12, page 109). CABS-dock results for Np1-DRB1*0101, Np2-DR5, Np3-DR5, Np6-DQ0602, Np4-DQ8 and Np6-DQ8 peptide-HLA complexes were not plotted as either the peptides were found to be non-binders or exhibited RMSD>5 (Figure 4.12, page 109).

One way ANOVA analysis followed by Tukey's multiple comparison test indicated no significant difference for all peptides (except for Np2 in case of Autodock Vina) (Table 4.22, page 110). Statistical analysis indicated that all peptides (except Np1) displayed a skewness value which was either positive, close to zero or comparable to native peptide value for Autodock Vina results which meant the selected peptides interacted desirably with class II molecules. Positive skewness or normal distribution for all peptides in case of CABS-dock results depicted favourable binding of selected peptides with majority of HLA molecules (Table 4.22, page 110). Significantly, the mean binding energy of Np1 and Np3 and mean RMSD value of Np2, Np3 and Np4 were found to be better than the respective mean values of native peptide for all HLA class II molecules (Table 4.22, page 110).

a)



b)

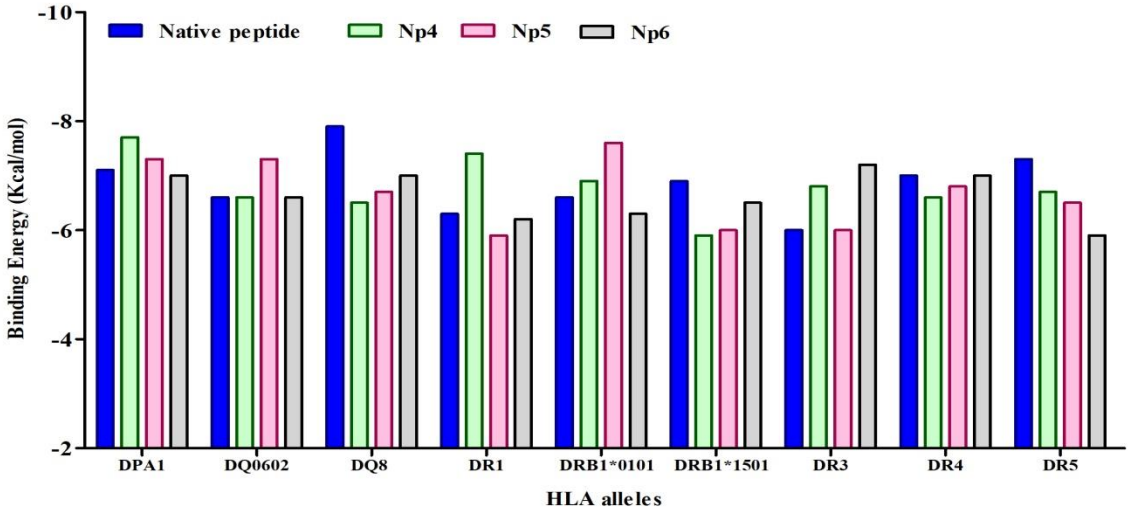
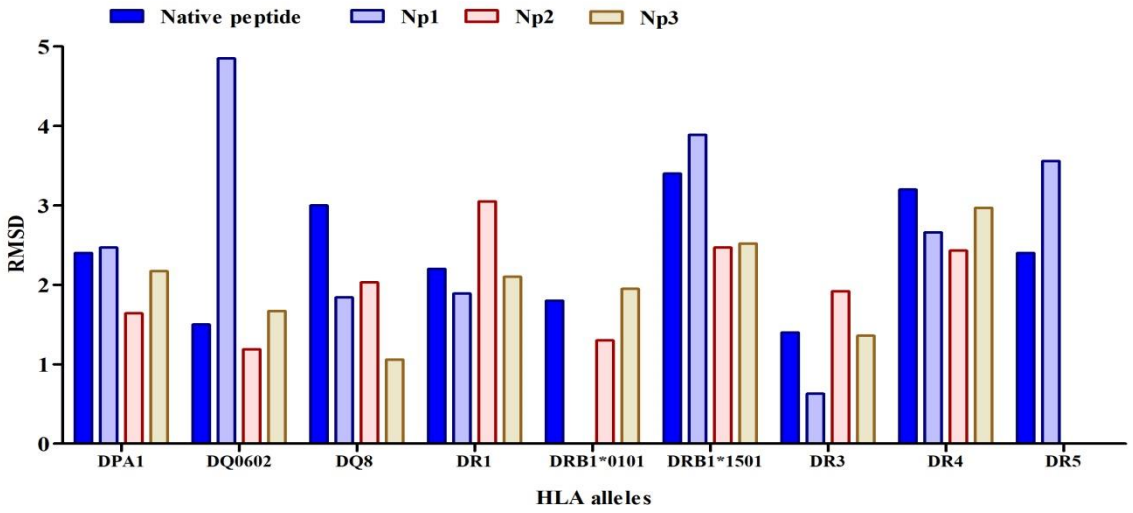


Figure 4.11: Binding energies for the identified nucleoprotein peptide fragments against HLA class II molecules. a) Np1, Np2 and Np3 b) Np4, Np5 and Np6

a)



b)

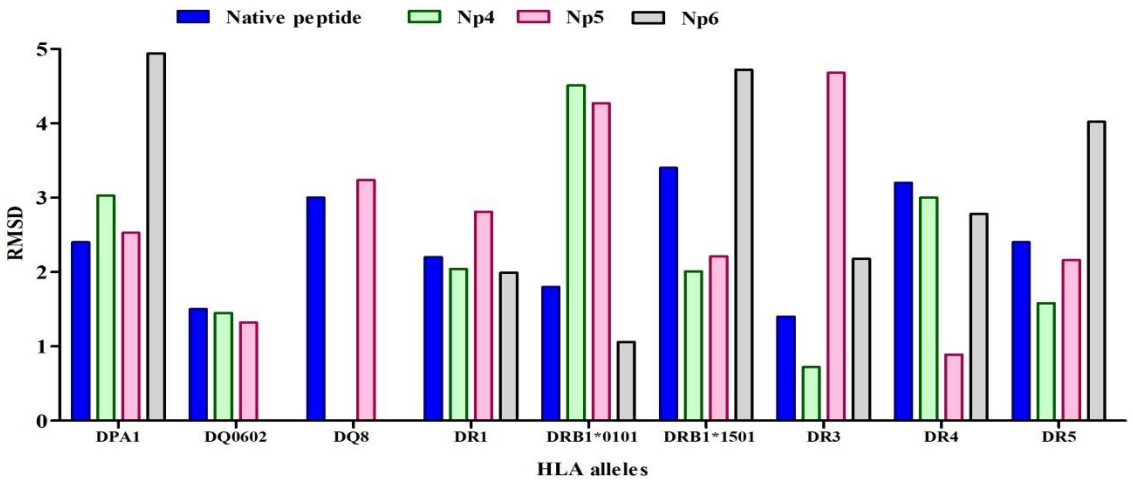


Figure 4.12: RMSD values for the identified nucleoprotein peptide fragments against HLA class II molecules. a) Np1, Np2 and Np3 b) Np4, Np5 and Np6

Table 4.22: Comparative analysis of mean binding energies (Autodock Vina) and RMSD (CABS-dock) values of nucleoprotein peptides with native peptide for HLA class II molecules

	Minimum	25% Percentile	Median	75% Percentile	Maximum	Mean	SD	Skewness	Significant $p < 0.05$
Binding energy (Kcal/mol)									
Native peptide	-7.9	-7.2	-6.9	-6.45	-6	-6.85	0.56	-0.37	
Np1	-7.6	-7.25	-6.8	-6.6	-6.5	-6.88	0.40	-1.10	ns
Np2	-6.3	-6.15	-5.8	-5.15	-5	-5.64	0.50	0.00	*
Np3	-7.9	-7.4	-7.2	-6.7	-5.8	-7.06	0.60	1.04	ns
Np4	-7.7	-7.15	-6.7	-6.55	-5.9	-6.78	0.52	-0.27	ns
Np5	-7.6	-7.3	-6.7	-6	-5.9	-6.67	0.63	-0.10	ns
Np6	-7.2	-7	-6.6	-6.25	-5.9	-6.63	0.44	0.33	ns
Root mean square deviation (RMSD)									
Native peptide	1.40	1.65	2.40	3.10	3.40	2.36	0.72	0.08	
Np1	0.60	1.82	2.60	3.82	4.90	2.73	1.36	0.08	ns
Np2	1.20	1.37	1.95	2.47	3.10	2.00	0.64	0.42	ns
Np3	1.10	1.47	2.05	2.42	3.00	2.00	0.60	0.13	ns
Np4	0.70	1.52	2.00	3.00	4.50	2.28	1.17	0.78	ns
Np5	0.90	1.75	2.50	3.75	4.70	2.67	1.25	0.34	ns
Np6	1.10	2.00	2.80	4.70	4.90	3.10	1.45	0.04	ns

The mean of the binding energy and RMSD values obtained for all peptide fragments have been depicted and analysed via skewness and one way ANOVA analysis followed by Tukey's multiple comparison test. Significant differences are indicated by * ($p<0.05$) where as ns stands for non-significant.

4.2.3.3. Peptides displaying best interactions and non-binders for various HLA alleles

Representative poses of HLA class I and II dockings with $CD8^+$ T cell epitope and $CD4^+$ epitope enriched peptide respectively, which resulted in highest binding energies are shown in Figure 4.13, page 112 (glycoprotein) and Figure 4.14, page 113 (nucleoprotein).

A peptide was considered as a non-binder if:-

a) The peptide-HLA complex exhibited a positive binding energy during Autodock Vina analysis
b) The peptide-HLA complex exhibited a RMSD > 5 during CABS-dock analysis
c) The peptide was found to bind outside the HLA binding groove during Autodock Vina and/or CABS-dock analysis

A list of all peptides and the number of HLA molecules for which they were found to be non-binder has been presented in table 4.23, page 114. Peptides (Gp2, Np1, Np4, and Np6) were observed to be non-binders for more than one HLA alleles and were not considered for the evaluation of immunogenic response in the *in vitro* system in this study.

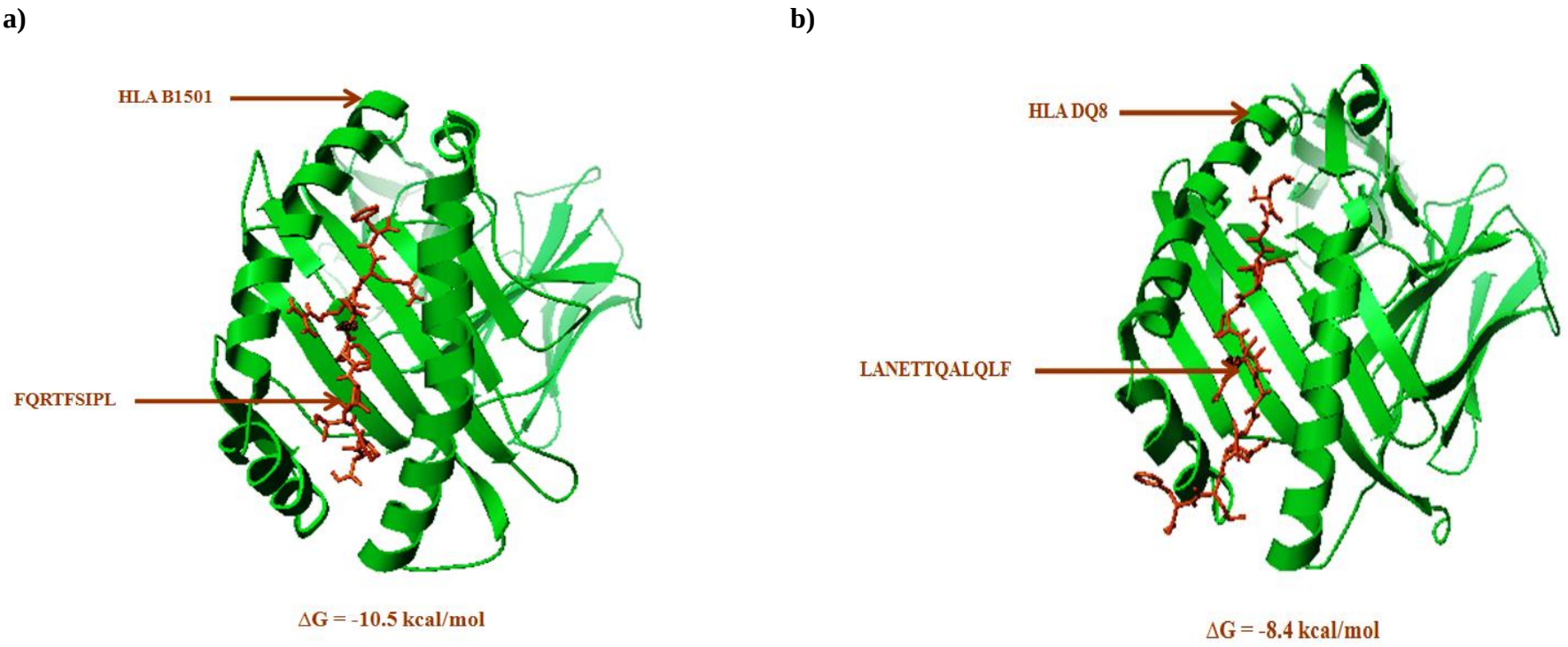


Figure 4.13: Autodock Vina docking poses showing highest binding energy. a) $CD8^+$ T cell epitope (FQRTFSIPL) of glycoprotein with HLA class I molecule (HLA B1501) b) $CD4^+$ T cell peptide (LANETTQALQLF) of glycoprotein with HLA class II molecule (HLA DQ8)

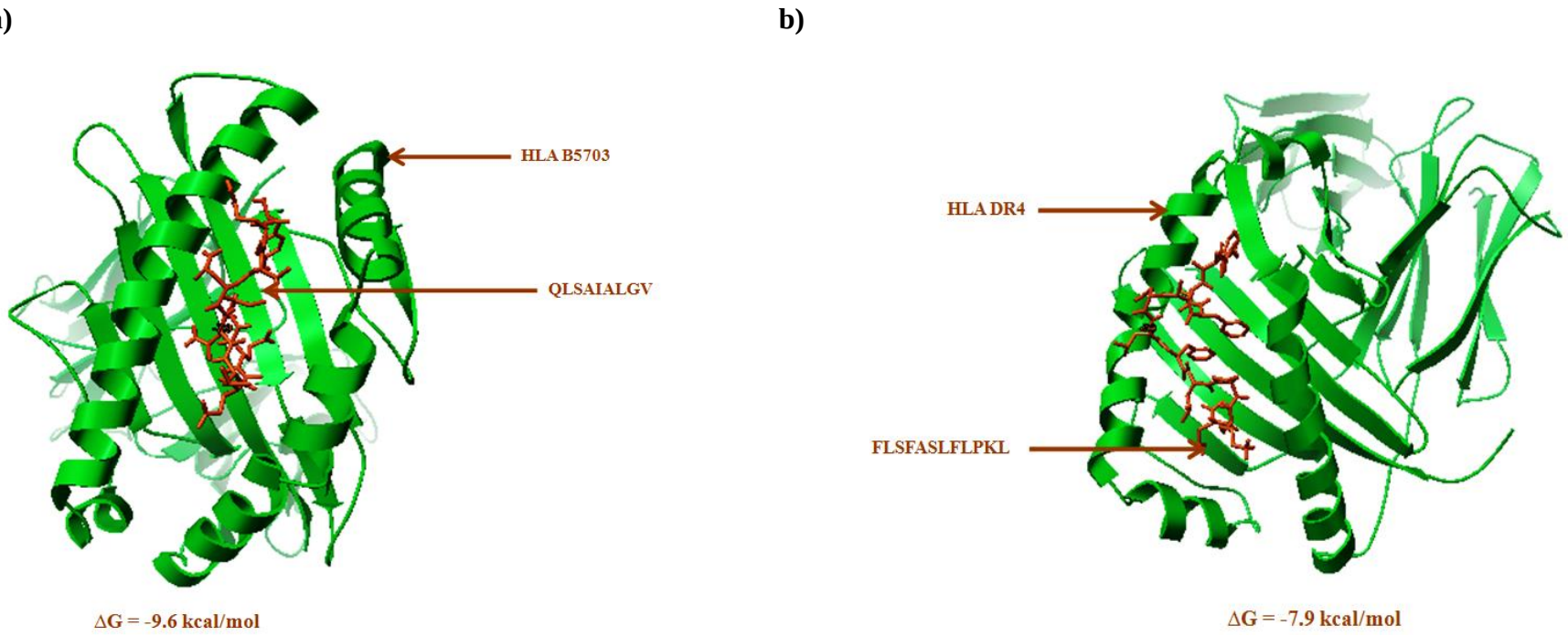


Figure 4.14: Autodock Vina docking poses showing highest binding energy. a) $CD8^+$ T cell epitope (QLSAIALGV) of nucleoprotein with HLA class I molecule (HLA B5703) b) $CD4^+$ T cell peptide (FLSFASLFLPKL) of nucleoprotein with HLA class II molecule (HLA DR4)

Table 4.23: Number of HLA molecules for which the glycoprotein and nucleoprotein peptide fragments acted as non-binders during molecular docking analysis

Peptide	Peptide sequence	Number of non-binders
Gp1	FKRTSFFLWVIILFQRTFSIPL	1
Gp2	FHKEGAFFLYDRLASTVIYRGTTFAEG	6
Gp3	LANETTQALQLF	1
Gp4	RATTELRTFSILNRKAIDF	1
Np1	MDYHKILTAGL	2
Np2	IVRQRVIPVYQVNNLEEI	-
Np3	FLSFASLFLPKL	1
Np4	ERGVRLHPLARTAK	2
Np5	LFPQLSAIALGVATAHGS	1
Np6	VLYYHMMKDEPVVF	3

Objective 3

4.3. Estimation of immunogenic response of *in silico* screened peptides in peripheral blood mononuclear cells

The peptides showing best results during *in silico* analysis were selected for assessing the immunogenic potential based on peptide induced cell proliferation and interferon-gamma (IFN-γ) release in peripheral blood mononuclear cells (PBMC). The selection was based on three parameters viz., number of HLA alleles covered, peptide affinity with HLA molecules based on molecular docking analysis and best population coverage among different continents. Based on these criteria, Gp1, Gp3 and Gp4 for glycoprotein and Np2, Np3 and Np5 for nucleoprotein were selected and commercially synthesized with a purity level >90%.

4.3.1. Peptide induced cell proliferation in PBMC

In order to measure the peptide induced cell proliferation, PBMC isolated from the blood samples of healthy volunteers were given repeated peptide stimulus and MTT assay was carried out to assess the proliferation of peptide stimulated PBMC. Concanavalin A (con A) was used as positive control whereas unstimulated cells served as negative control.

On 0^{th} day, $2*10^5$ cells were seeded per well and stimulated with the synthesized peptide. Restimulation was done on 3^{rd} day while MTT assay was carried out on 5^{th} day. All the experiments were carried out in triplicates.

The PBMC proliferation resulting due to each selected peptide was measured in PBMC isolated from ten different healthy blood samples. Proliferation was represented as stimulation index which is the ratio of average absorbance of the peptide stimulated cells and average absorbance of unstimulated cells. Samples displaying a stimulation index >1 were considered as positive responders. Amongst the GP peptides, Gp1, Gp3 and Gp4 showed a positive response in eight, six and nine healthy individuals respectively (Figure 4.15, page 116). Further, analysis with the help of unpaired t-test revealed that results for seven, six and eight individuals for Gp1, Gp3 and Gp4 respectively were statistically significant. In case of EBOV NP peptides, Np3 was found to have a positive response in all ten healthy individuals (nine statistically significant) while nine and six positive responders (six and five statistically significant) were found for Np2 and Np5 respectively (Figure 4.16, page 117).

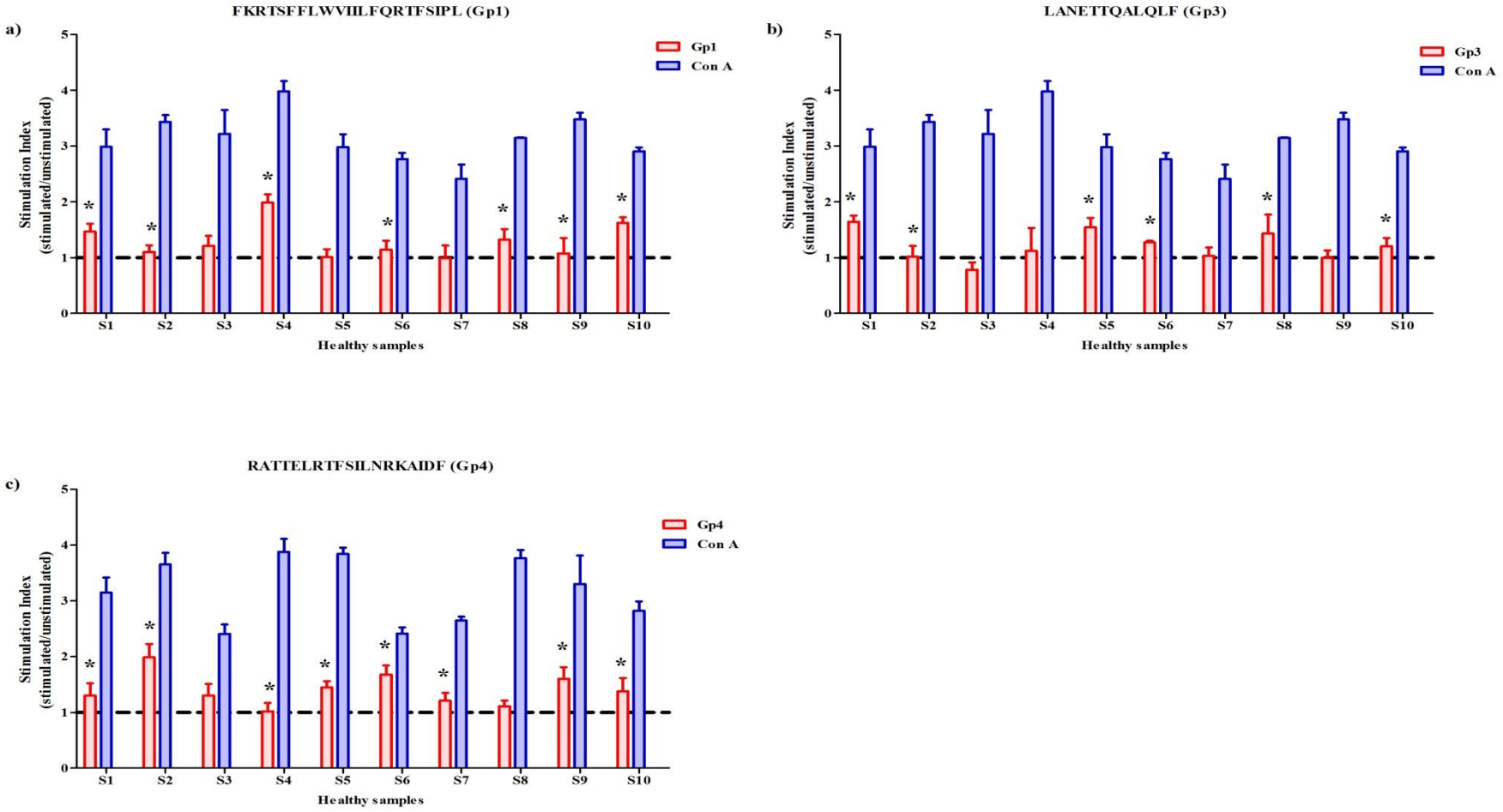


Figure 4.15: Cell proliferation by glycoprotein peptides stimulated peripheral blood mononuclear cells (PBMC) of 10 healthy blood samples. Concanavalin A (Con A) was used as positive control. Significant differences between unstimulated and peptide stimulated cells are indicated by * ($p<0.05$). S1-S10 represent the 10 different healthy blood samples.

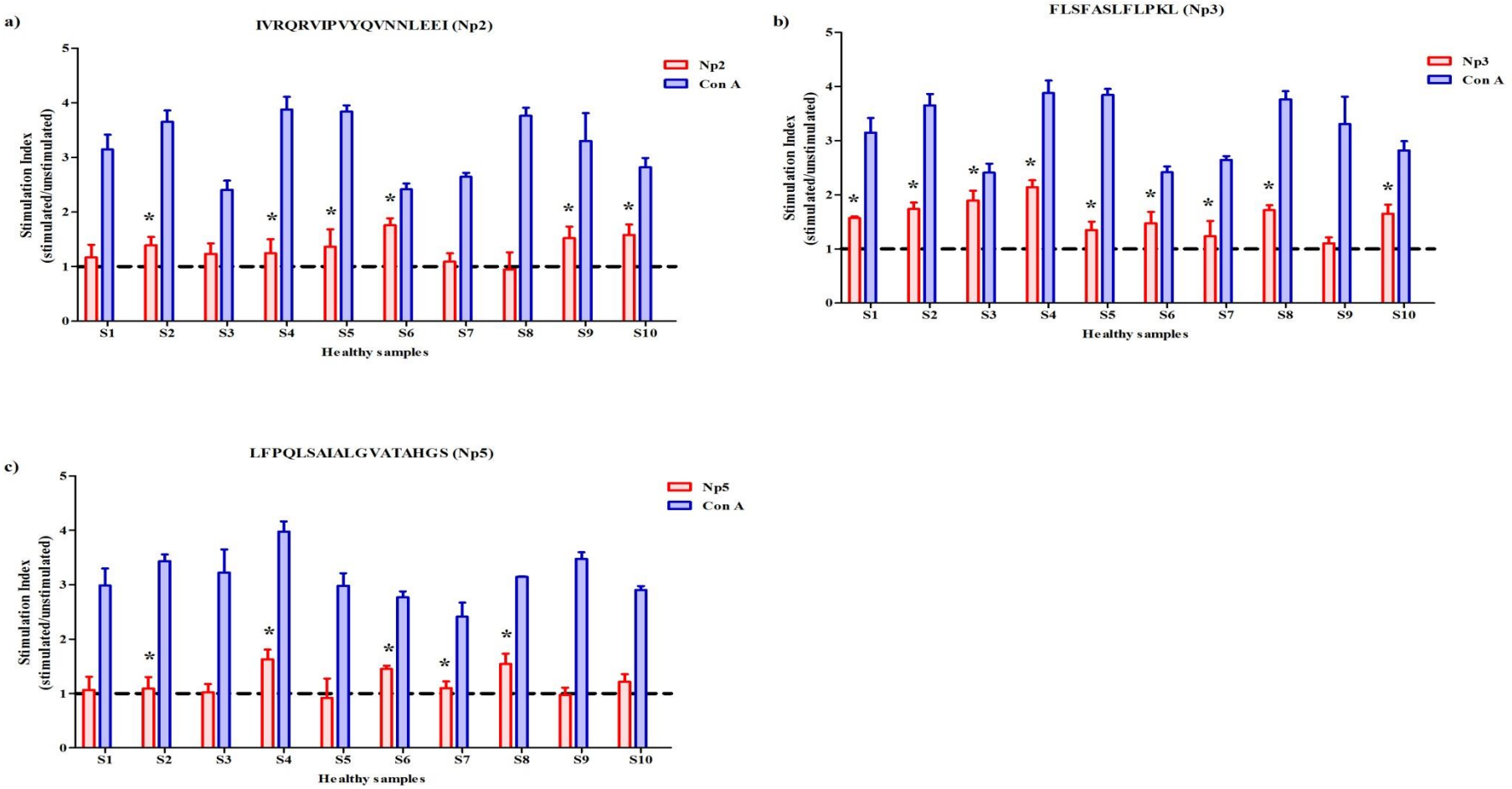


Figure 4.16: Cell proliferation by nucleoprotein peptides stimulated peripheral blood mononuclear cells (PBMC) of 10 healthy blood samples. Concanavalin A (Con A) was used as positive control. Significant differences between unstimulated and peptide stimulated cells are indicated by * ($p<0.05$). S1-S10 represent the 10 different healthy blood samples.

4.3.2. Peptide induced IFN-γ release in PBMC

In the present study, sandwich ELISA was carried to measure the IFN-γ production by peptide stimulated PBMC. IFN-γ release was checked for each peptide in ten different healthy blood samples and the experiment was carried out in triplicates. PBMC were stimulated with the selected peptides on day 0 and 3 and on 5^{th} day, culture supernatant was collected and tested for the presence of IFN-γ.

Increase in IFN-γ secretion in response to con A and peptide stimulus was expressed as fold increase which is the ratio of average absorbance of peptide stimulated cells and unstimulated cells. Amongst the GP peptides, Gp4 exhibited the best results as significant positive response was found for all ten healthy blood samples (Figure 4.17, page 119). Other GP peptides viz., Gp1 and Gp3 showed positive response in eight and seven samples (seven statistically significant results for each) respectively (Figure 4.17, page 119). In case of NP peptides, Np2 and Np3 showed good response in eight and nine samples respectively though seven statistically significant results were found for each of these peptides. Np5 displayed the least desirable results as significant IFN-γ release was detected only in two samples (Figure 4.18, page 120).

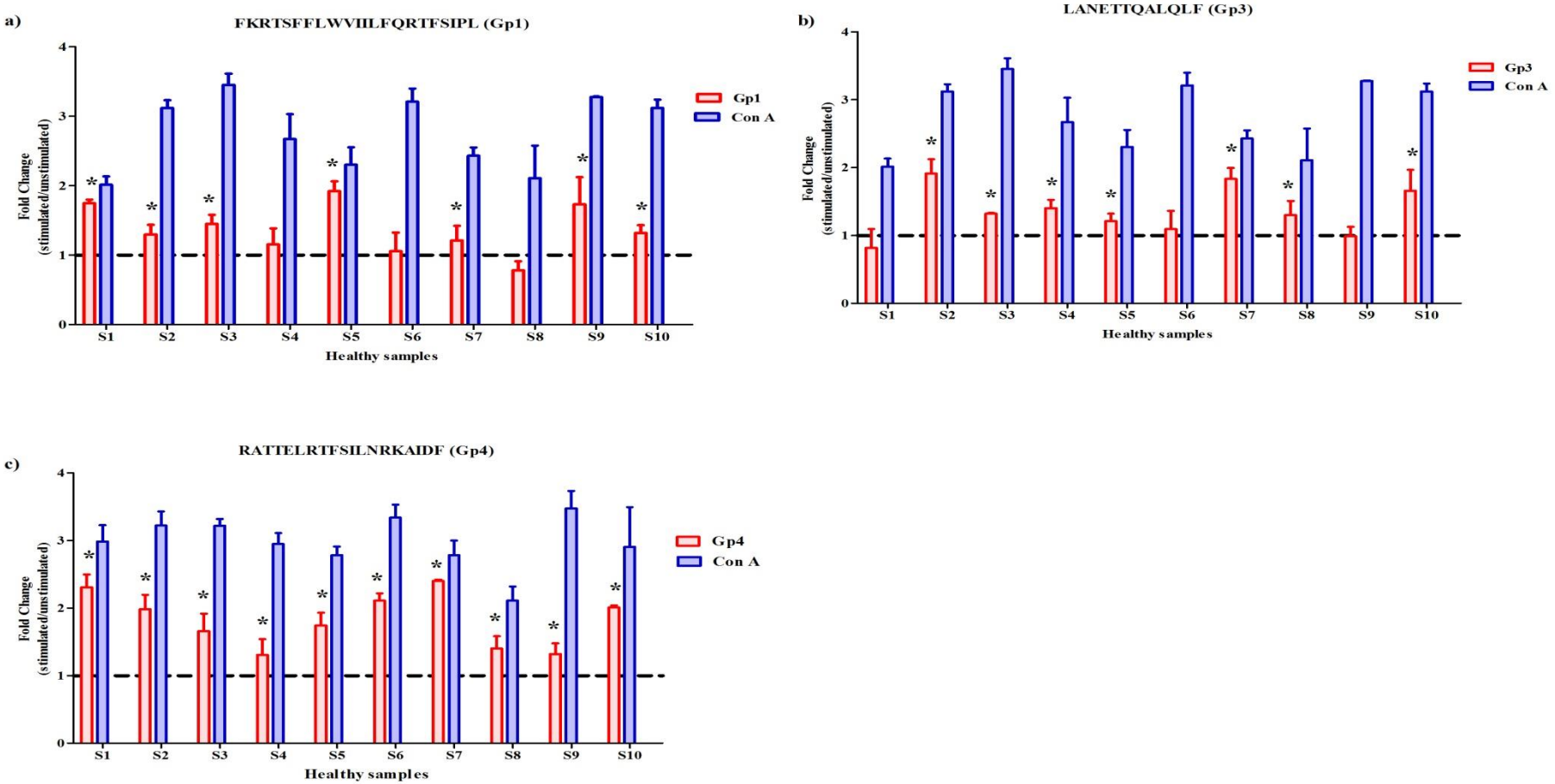


Figure 4.17: Interferon-gamma (IFN-γ) secretion by glycoprotein peptides stimulated peripheral blood mononuclear cells (PBMC) of 10 healthy blood samples. Concanavalin A (Con A) was used as positive control. Significant differences between unstimulated and peptide stimulated cells are indicated by * ($p<0.05$). S1-S10 represent the 10 different healthy blood samples.

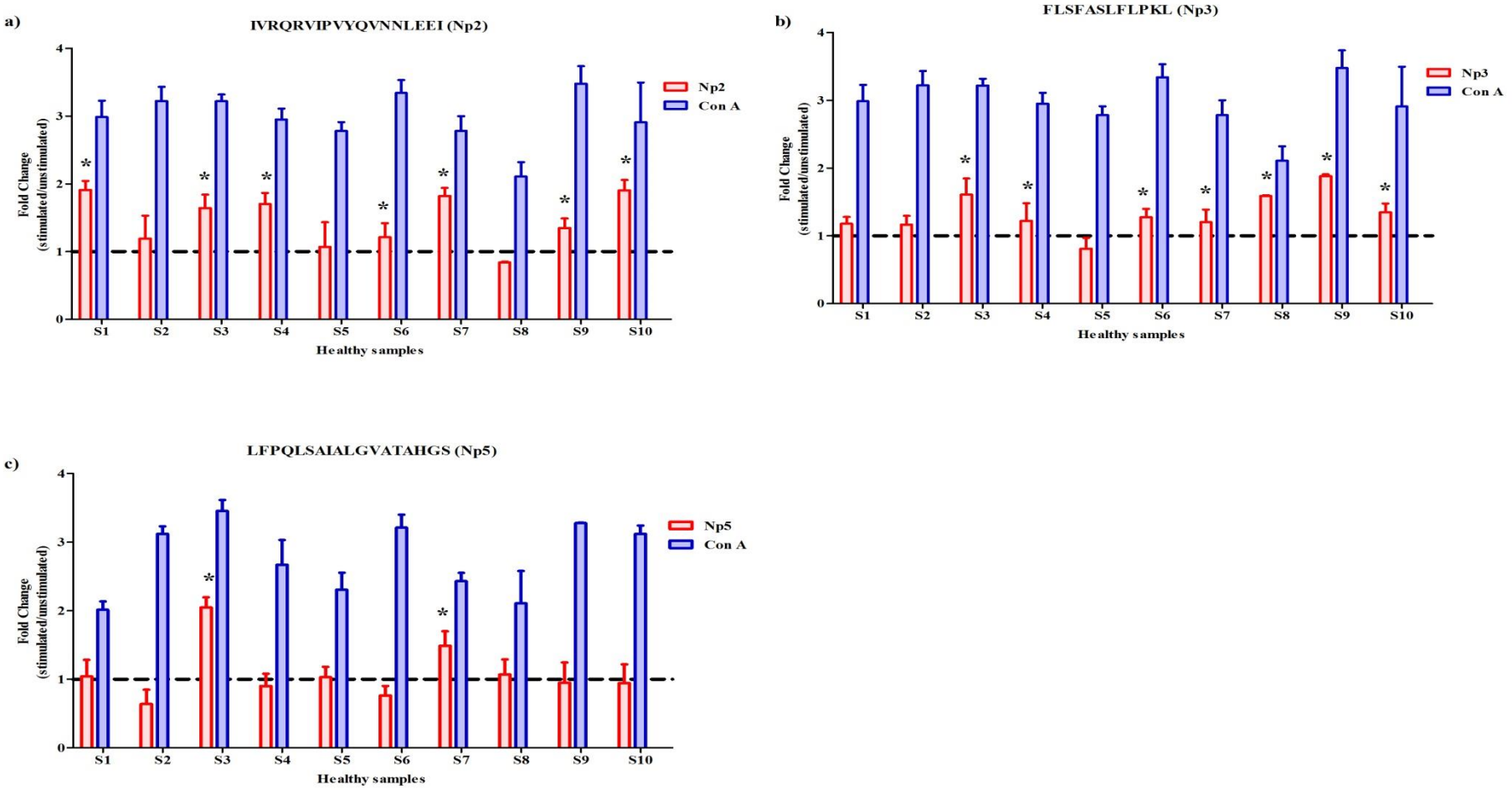


Figure 4.18: Interferon-gamma (IFN-γ) secretion by nucleoprotein peptides stimulated peripheral blood mononuclear cells (PBMC) of 10 healthy blood samples. Concanavalin A (Con A) was used as positive control. Significant differences between unstimulated and peptide stimulated cells are indicated by * ($p<0.05$). S1-S10 represent the 10 different healthy blood samples.

4.3.3. Comparative analysis of the cell proliferation and IFN-γ secretion measurement results

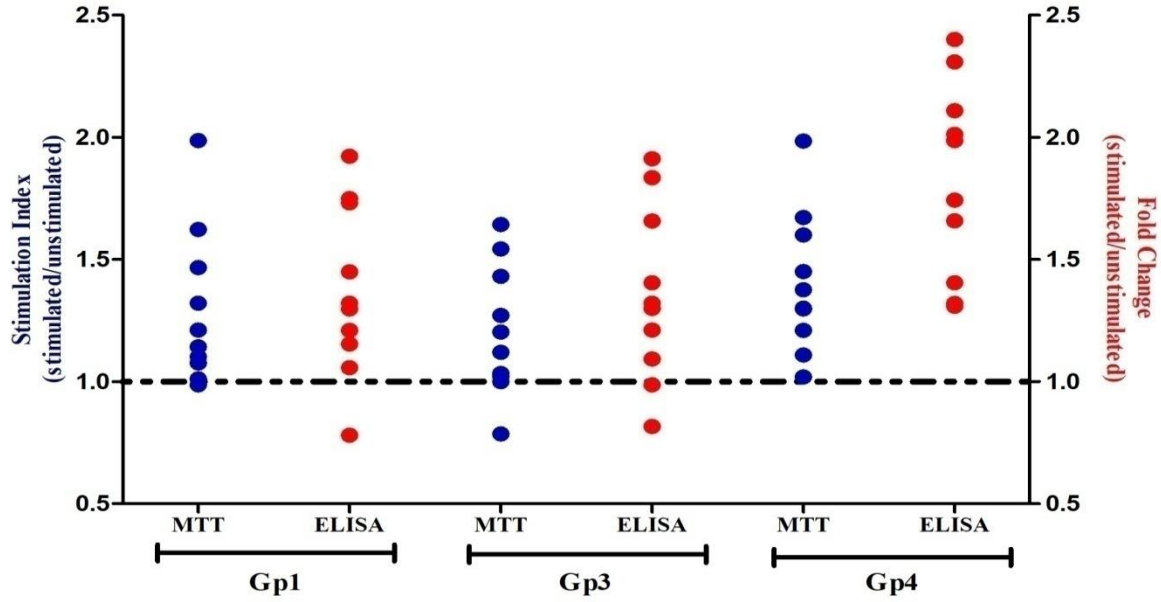
The average stimulation index (cell proliferation) and fold change (IFN-γ) of each of six peptides for all 10 samples was found to be >1 indicating the immunogenic potential of these peptides (Table 4.24, page 121). Amongst GP peptides, Gp4 returned the best results for both, MTT (cell proliferation) and ELISA (IFN-γ) as statistically significant positive responders are ≥ 8 while other the two peptides (Gp1 and Gp3) exhibited good results as well (Table 4.24, page 121; Figure 4.19 a), page 122). In case of NP peptides, overall results in terms of both, cell proliferation and IFN-γ release for two peptides (Np2 and Np3) were promising while one peptide (Np5) responded poorly especially for IFN-γ release (Table 4.24, page 121; Figure 4.19 b), page 122). Correlation analysis between MTT and ELISA results revealed a significant correlation with r value ranging from 0.624 to 0.718 for all six peptides (Table 4.25, page 123; Table 4.26, page 124).

Table 4.24: Positive responders of the selected six EBOV peptides along with average stimulation index and fold change for 10 healthy blood samples

Peptide	Sequence	Positive responders		Average (mean ± SD)	
		MTT	ELISA	MTT (stimulation index)	ELISA (fold Change)
Gp1	FKRTSFFLWVIILFQRTFSIPL	8 (7*)	8 (7*)	1.29 ± 0.31	1.36 ± 0.33
Gp3	LANETTQALQLF	6 (6*)	7 (7*)	1.2 ± 0.26	1.35 ± 0.33
Gp4	RATTELRTFSILNRKAIDF	9 (8*)	10 (10*)	1.4 ± 0.28	1.82 ± 0.37
Np2	IVRQRVIPVYQVNNLEEI	9 (6*)	8 (7*)	1.32 ± 0.24	1.46 ± 0.36
Np3	FLSFASLFLPKL	10 (9*)	9 (7*)	1.58 ± 0.31	1.32 ± 0.28
Np5	LFPQLSAIALGVATAHGS	6 (5*)	2 (2*)	1.2 ± 0.25	1.08 ± 0.38

*Significant at $p<0.05$

a)



b)

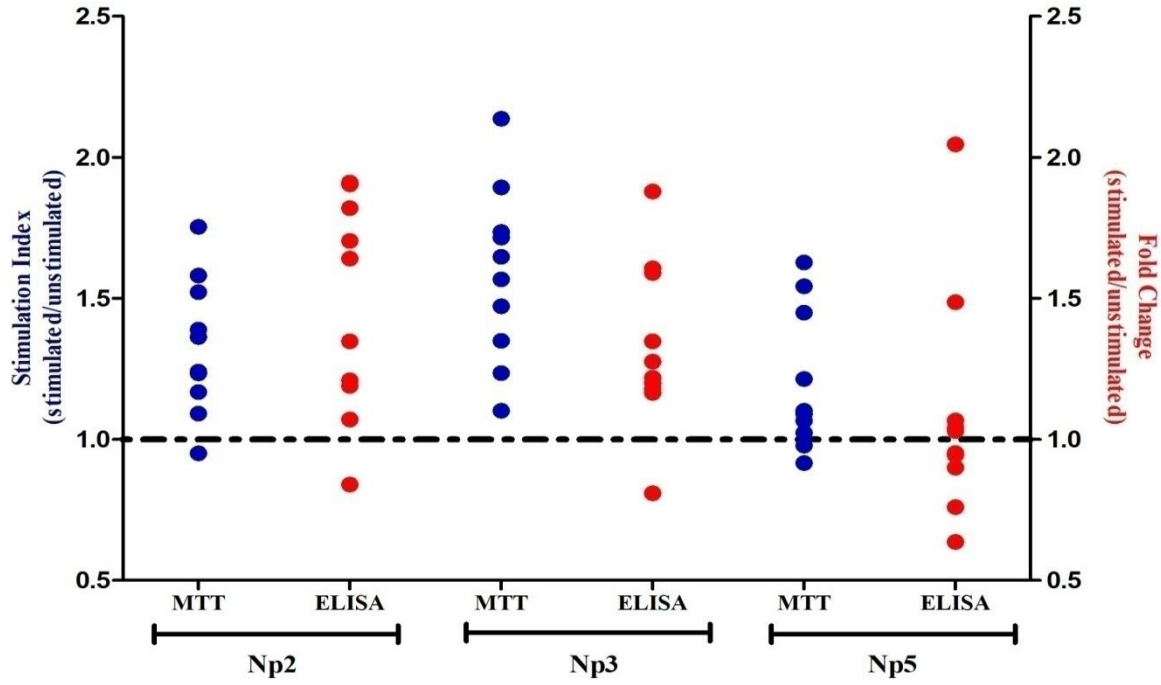


Figure 4.19: A comparative representation of the cell proliferation (MTT) and IFN-γ secretion (ELISA) results of 10 healthy blood samples for a) glycoprotein peptides and b) nucleoprotein peptides

Table 4.25: Correlation analysis of MTT and ELISA data obtained for glycoprotein peptides

Healthy sample	Gp1				Gp3				Gp4			
	MTT	ELISA	Correlation		MTT	ELISA	Correlation		MTT	ELISA	Correlation	
			Pearson r	P value (two-tailed)			Pearson r	P value (two-tailed)			Pearson r	P value (two-tailed)
S1	1.467	1.749	0.643	0.032	1.643	0.817	0.624	0.040	1.301	2.309	0.633	0.036
S2	1.102	1.298			1.021	1.912			1.985	1.987		
S3	1.211	1.450			0.786	1.321			1.298	1.659		
S4	1.987	1.154			1.120	1.405			1.019	1.309		
S5	1.011	1.923			1.543	1.211			1.451	1.743		
S6	1.143	1.057			1.271	1.093			1.672	2.109		
S7	0.987	1.210			1.034	1.835			1.210	2.401		
S8	1.321	0.781			1.431	1.300			1.109	1.405		
S9	1.076	1.732			1.000	0.987			1.600	1.320		
S10	1.623	1.321			1.203	1.657			1.376	2.012		

Table 4.26: Correlation analysis of MTT and ELISA data obtained for nucleoprotein peptides

Healthy sample	Np2				Np3				Np5			
	MTT	ELISA	Correlation		MTT	ELISA	Correlation		MTT	ELISA	Correlation	
			Pearson r	P value (two-tailed)			Pearson r	P value (two-tailed)			Pearson r	P value (two-tailed)
S1	1.168	1.911	0.713	0.013	1.568	1.179	0.718	0.012	1.066	1.041	0.691	0.018
S2	1.390	1.190			1.736	1.165			1.091	0.637		
S3	1.234	1.641			1.894	1.607			1.024	2.046		
S4	1.241	1.704			2.137	1.219			1.628	0.900		
S5	1.363	1.071			1.350	0.809			0.916	1.030		
S6	1.754	1.210			1.472	1.276			1.450	0.760		
S7	1.092	1.820			1.235	1.200			1.102	1.487		
S8	0.951	0.840			1.716	1.590			1.543	1.067		
S9	1.522	1.348			1.102	1.879			0.978	0.952		
S10	1.581	1.905			1.648	1.348			1.214	0.945		

Chapter 5: Discussion

Ebola virus (EBOV) presents a global challenge as it threatens bio-security and health of individuals worldwide. Though the outbreaks are sporadic, yet the intensity and lethality which they affect an area leads to a high mortality and morbidity rate (White and Schornberg, 2012). The rapid spread of virus in the affected area via person to person transmission further impedes the treatment and post-infection healthcare. Therefore, EBOV has been classified as a category A pathogen by National Institute of Allergy and Infectious Diseases (NIAID) and recognised as an urgent medical challenge by the Centre for Disease Control and Prevention (CDC) (Madara et al., 2015). It has been categorised as a biothreat pathogen as well as a U.S. Select Agent and Tier 1 Pathogen (Fed. Sel. Agent Program, 2018). Also, EBOV is a Biosafety Level 4 (BSL4) pathogen which restricts the work on combating the virus to a handful of scientific community with appropriate facilities (Schneider-futschik et al., 2018).

Contrary to the popular misconception that EBOV is found only in African continent, the virus has been discovered in various other parts of the world as well. *Reston ebolavirus* was found in monkeys in the US (Jahrling et al., 1990) and Asia (Miranda and Miranda, 2011). Further, this species has also been known to affect pigs in the Philippines (Barrette et al., 2009). Human infections caused by *Sudan ebolavirus* and *Zaire ebolavirus* have been registered in England (Emond et al., 1977) and Russia (Borisevich et al., 2006) respectively. Zaire has affected people in Italy (WHO, 2015) and Spain (WHO, 2014) as well. This potential global threat posed by EBOV and the lack of an effective preventive measure (Brannan et al., 2019) calls for an urgent need to develop globally effective diagnostic and preventive measures against EBOV.

Various studies have been conducted on different EBOV proteins in a bid to develop an effective vaccine. EBOV glycoprotein (GP) has been the most targeted candidate as it is responsible for viral attachment and entry into the host cells. As the only protein on the viral surface, it is known to be the sole target for classical neutralizing antibodies (Feldmann et al., 2018). Interestingly, EBOV nucleoprotein (NP) also emerged as a potent vaccine candidate as shown by different recent studies (Gerritz et al., 2011; Kao et al., 2010). Majority of patients have displayed a positive response verified by their cytokine levels upon exposure to NP (Sakabe et al., 2018). The abundance of NP in the virus and strong antigenicity displayed by the C-terminal half of EBOV NP further strengthen its candidature (Niikura et al., 2001; Saijo

et al., 2001). Also, NP is the most conserved protein amongst all the Filoviruses (Changula et al., 2013) and thus, is a lucrative target for development of a multi-potent vaccine. Therefore, EBOV GP and NP were selected in the current study to find plausible vaccine candidates.

Traditionally, live vaccines have been employed to confer protection against various pathogens. Though highly immunogenic, live vaccines come with their bag of problems; stringent storage conditions and stability issues amongst them. Also, there is always a risk of the attenuated virus reverting back and regaining its ability to infect which might prove disastrous for the recipient of the vaccine (Olson et al., 2001). Killed or inactivated vaccines solve the stability and safety issues but are relatively weakly immunogenic and thus, require regular booster doses to induce the desired immune response. Handling of highly virulent strains is possible only in few workspaces around the world equipped with state-of-the-art facilities. Moreover, traditional vaccines might result in undesired responses such as induction of an autoimmune or allergic response (Jarzab et al., 2013). Hence, the trend is shifting towards vaccine development with the help of a minimal microbial component capable of inducing a robust immune response and conferring long lasting immunity (Skwarczynski and Toth, 2016). The present study is reporting three highly conserved multi-epitopic peptides each for Ebola virus glycoprotein and nucleoprotein which have the potential to be a part of a globally effective EBOV peptide vaccine.

Objective 1

5.1. Immunoinformatics based prediction of peptides containing epitopes of Ebola virus

Uncovering potential epitopes is a pivotal step in the design of a peptide based vaccine. Computational immunology or immunoinformatics is a branch of science that addresses the organisation of massive, raw immunological data with the help of computational approaches into an understandable form, enabling the researcher to generate meaningful interpretations (Korber et al., 2006). Epitope prediction with the help of immunoinformatics approach is efficient as it narrows down the list of potential candidates thus, saving a lot of hassle encountered during experimental approaches. Moreover, the recent surge in the development of immunological databases and prediction tools based on different parameters enables the unveiling of potential vaccine candidates and development of peptide based vaccine in lesser time (Jackwood et al., 2008) as opposed to traditional vaccine development which might take upto 15 years.

Various immmunoinformatics approaches have been used successfully in many studies for identifying vaccine candidates against several pathogens viz. *Human coronavirus* (Oany et al., 2008), *Saint Louis encephalitis virus* (Hasan et al., 2013), *Nairovirus* (Oany et al., 2015), *Chikungunya virus* (Hasan et al., 2015) etc. Epitopes of *Leishmania infantum*, *Mycobacterium tuberculosis*, *Mycobacterium bovis*, *Schistosoma mansoni* and epitopes responsible for peanut-allergy have also been identified *in silico* and successfully proven to be immunogenic (Agallou et al., 2014; Chen et al., 2012; Mustafa, 2011; Oliveira et al., 2016; Pascal et al., 2013).

Creating vaccines is particularly difficult for RNA viruses as they can quickly mutate their different exposed proteins (Twiddy et al., 2003). The 2014 West African EBOV outbreak (the worst in history) highlighted the ability of the virus to mutate during an epidemic (Carroll et al., 2015; Diehl et al., 2016). Hence, to account for current species as well as future viral mutations, highly conserved GP and NP regions of all unique GP and NP sequences available in NCBI and viprbrc databases belonging to human infecting EBOV species (*Zaire ebolavirus*, *Sudan ebolavirus*, *Taï Forest ebolavirus* and *Bundibugyo ebolavirus*) were selected.

The predicted epitopes must be able to induce a robust, long-lasting immune response. Thus, to predict epitopes which surpass the criteria set based on different parameters and algorithms, a consensus approach was applied in the current study. Previous reports have used consensus approach to predict *Influenza virus* epitopes in matrix protein (Lohia and Baranwal, 2015) which were further validated *in vitro* for immune response (Lohia and Baranwal, 2017). The consensus prediction approach was also applied in various other studies (Moutaftsi et al., 2006; Trost et al., 2007; Chen et al., 2012; Ichihashi et al., 2011). Therefore, three T cell epitope prediction tools each for HLA class I ($CD8^+$) and HLA class II ($CD4^+$) were utilised in the present study. These tools follow different algorithms and predict epitopes based on different parameters such as proteasomal C-terminal cleavage, TAP (transporter associated with antigen processing) transport efficiency, stability of the peptide–HLA complex and peptide susceptibility to proteolysis in lysosome.

HLA class II binding groove is open ended and accommodates larger and flexible peptides as compared to closed HLA class I binding groove (Andreatta and Nielsen, 2016). Hence, the prediction tools of HLA class II are not as accurate as those of HLA class I. X-ray studies have shown that the binding groove of HLA class II is occupied by large peptides in which the core peptide is of nine amino acids (also called peptide registers) and rest of amino acids

extend to different sides of groove (Dimitrov et al., 2010). There is existence of multiple peptide registers and as a result, long peptides adopt different conformations which contribute to the poor prediction accuracy of class II binding epitopes. However, class II predicted epitopes have shown encouraging results in *in vitro* or *in vivo* system (Kovjazin et al., 2013; Burgosa et al., 2010; Brintnell et al., 2007). In the current study, peptides having both HLA I ($CD8^+$) and II ($CD4^+$) restricted epitopes were selected for GP and NP. In earlier studies, large peptides containing HLA I and II restricted epitopes of *Influenza virus* were found to exhibit peptide induced cell proliferation and IFN-γ release in peripheral blood mononuclear cells (Lohia and Baranwal, 2018).

Computational studies for elucidation of epitopes in Ebola virus have been conducted earlier as well but in contrast to present study, Sundar et al used one prediction tool (Sundar et al., 2007) while two prediction tools were used by Dutta et al (Dutta et al., 2015). An approach similar to the current study was employed by Dikhit et al where three different prediction algorithms were used to define only $CD8^+$ T cell epitopes but not $CD4^+$ T cell epitopes (Dikhit et al., 2015). In all these studies, the work was carried out only for identifying $CD8^+$ T cell epitopes. Further, in contrast to previous studies where protein sequences were taken from single EBOV strain, present study has considered the sequences from 1976 to 2018 of all EBOV strains infecting humans.

It is beneficial to have vaccine candidates with the ability to induce both, cell mediated (T cell) as well as humoral (B cell) immunity. Therefore, presence of B cell epitopes along with T cell epitopes augurs well for a vaccine candidate. Most of the B cell epitopes are discontinuous and nearly 10% epitopes are linear (Wang et al., 2011). *In silico* prediction tools for B cell epitopes (mainly linear B cell epitopes) (Ponomarenko and Van Regenmortel, 2009) have been developed but are not as reliable as T cell epitope prediction tools. First efforts to predict B cell epitopes were made based on amino acid properties and three dimensional protein structures in the early 1980's (Hopp and Woods, 1981). In one recent study on malarial antigen, some epitopes predicted by ABCpred were found to match with experimentally identified epitopes in immune rabbit antiserum (Bergmann-Leitner et al., 2013). Hence, in the present study, B cell epitope prediction was carried out with the help of ABCPred tool which works on artificial neural network algorithm (Saha and Raghava, 2006). It is known to perform more efficiently as compared to tools based on Karplus and Schulz's flexibility scale (Karplus and Schulz, 1985) and Parker's hydrophobicity scale (Parker et al., 1986). The study revealed that four out of six GP and eight out of twelve NP peptide

fragments contained B cell epitopes and hence, these peptides were considered for further analysis.

Any component of a vaccine may contribute to adverse effects after vaccine administration. Measles-Mumps-Rubella (MMR) vaccination has been reported to induce anaphylactic reactions in some individuals (Patja et al., 2001). Varicella and HPV vaccination might induce a drop in blood pressure, bronchospasm and/or shock (Fritsche et al., 2010). Tetanus vaccination might lead to an Arthus-type reaction at the site of vaccination (Facktor et al., 1973). Therefore, the capacity of peptides identified in the current study to cause an allergic and toxic response was predicted with the help of AlgPred and ToxinPred tools respectively. The results established that all peptides under consideration were incapable of inducing any allergic or toxic response.

Development of autoimmune diseases such as Gullian Barr syndrome (influenza and oral polio vaccine), systemic lupus anaemia, mysthemia gravis, polyarthritis and thrombocytopenia (hepatitis B), diabetes mellitus (mumps), neural complications (rabies) and multiple sclerosis (swine flu) after vaccination has been reported in several studies (Cohen and Shoenfeld, 1996; Meyboom et al., 1995; Topaloglu et al., 1992). Hence, it becomes critical to eliminate the vaccine candidates which resemble human proteome to avoid autoimmunity. BLAST analysis was applied in the current study to remove potential autoimmune response elucidators. One EBOV GP peptide fragment, YLFEVDNLTYVQLESRFT (also found to be devoid of B cell epitopes) resembled the human cytoskeletal linker protein epiplakin1 while two EBOV NP peptide fragments viz., VGHMMVIFRLMRTNFLIKFLLIHQGMHMV and YAPFARLLNLSGV exhibited similarity to human Anaphase-promoting complex/cyclosome and homeobox protein, Hox-B9 respectively. Hence, these peptides were eliminated from further consideration.

After screening for undesirable responses, the selected four GP and six NP conserved peptide fragments containing multiple T and B cell epitopes were checked for their uniqueness in the IEDB database. None of the identified GP and NP peptide fragments appear to be reported exactly in any of the previous studies. Amongst the EBOV GP peptide fragments, a small sequence (LYDRLASTV) of Gp2 was reported to be capable of stimulating splenocytes in mice to secrete IFN-γ (Wu et al., 2012). Gp3 was found to be part of a large peptide (TEGLMHNQDGLICGLRQLANETTQALQLFLRATTELRT) which was reported to react with antisera obtained after vaccinating the Ebola DNA vaccine (Ripoll et al., 2017). TELRTFSI peptide which is a part of Gp4 has shown cytotoxic T lymphocytes response in

mice model (Rao et al., 1999). Amongst the EBOV NP peptide fragments, YQVNNLEEI, FLSFASLFL and FPQLSAIAL are the partial fragments of Np2, Np3 and Np5 peptides respectively that have been reported to induce $CD8^+$ T cell response in earlier studies (Sundar et al., 2007; Ruibal et al., 2016; Theaker et al., 2016; Simmons et al., 2004). In another study, mice injected with EBOV NP vaccine responded by producing protective CTLs against VYQVNNLEEIC (consisting of a part of Gp2) (Wilson and Hart, 2001). A large peptide sequence, HILRSQGPFDAVLYYHMMKDEPVVFSTSDGKEYTYP (consisting of Np6), was reported to induce $CD8^+$ T cell response in human survivors (Sakabe et al., 2018). Therefore, these studies support the potential of selected peptides to induce immunogenic response.

In a bid to identify peptides that can be a part of a cross-protective, multi-potent vaccine, conservancy analysis was carried out for other *Filoviridae* family members viz., *Marburg virus* and *Lloviu virus* along with human infecting EBOV species. Like in *Human Immunodeficiency virus*, *Influenza A virus* and *Avian leukosis and sarcoma virus*, EBOV GP undergoes proteolytic cleavage which determines the potential of virus to cause infection (McCune JM et al., 1988; Dong et al., 1992). As per various previous studies, EBOV GP consists of a highly conserved region which is part of endoproteolytic cleavage site of *Avian leukosis and sarcoma virus* (Wool-Lewis and Bates, 1999). A few domains of EBOV GP exhibit amino acid similarity to those of oncogenic retroviruses (Volchkov et al., 1992). 33-50 residues of GP (A part of Gp1 lies in this polypeptide) are responsible for differential infectivity potentials of human pathogenic and non-pathogenic EBOV species (Fujihira et al., 2018). 36–351 residues of NP (Np2-Np5 lie in this polypeptide) are highly conserved amongst all EBOV strains, *Marburg virus* as well as *Lloviu virus* (Kirchdoerfer et al., 2015). Other Mononegavirales members such as *Human respiratory syncytial virus* (Tawar et al., 2009), *Nipah virus* (Yabukarski et al., 2014) and *Human parainfluenza virus* (Alayyoubi et al., 2015) have shown structural similarity to EBOV. It was observed that all GP and NP peptide fragments (except Gp1 – 99.39%) identified in the present study were 100% conserved in *Zaire ebolavirus* (the most pathogenic EBOV species). Gp3 and Np3 were 100% conserved in all EBOV species. Np3 was found to be 100% conserved in *Lloviu virus*. These results supported the idea of development of cross-protective peptide vaccine.

Additionally, these peptides were found to be parts of crucial domains of their respective proteins. Gp1 formed a part of signal peptide at the glycoprotein N terminus which directs the nascent glycoprotein to the endoplasmic reticulum where carbohydrates are added on its

surface (Feldmann et al., 1994; Lee et al., 2008). A thick layer of oligosaccharide coats majority of the GP trimer (Lee et al., 2008) enabling formation of a shield that protects it from humoral immune responses. It is similar to glycan shields of HIV-1 gp120 (Kwong et al., 1998; Wyatt et al., 1998) and *Epstein-Barr virus* gp350 (Szakonyi et al., 2006) thus, indicating a common immune evasion method. Gp3 and Gp4 were found to be located in the heptad repeat regions (HR1 and HR2) of GP2. HR1 can be segmented into HR1A, HR1B, HR1C and HR1D. HR1A and HR1B (residues 554-575) form an alpha helix. The bend between HR1A and HR1B contains an atypical stutter similar to that in *Human parainfluenza virus* (Yin et al., 2006), which may act as a conformational switch (Weissenhorn et al., 1998). Cys511 and Cys556 form a highly conserved disulfide bond similar to the two cysteine residues bordering the internal fusion loop in *Avian sarcoma leukosis virus* (Delos et al., 2008). HR1C (residues 576-582) forms an extended coil linking to HR1D (residues 583-598) (Lee et al., 2008). Np1-Np5 formed a part of N terminal 1-450 residues, a polypeptide which in itself is sufficient for viral genome replication (Watanabe et al., 2006). Np2–Np5 formed a part of N-terminal 36–351 residues which are needed for NP oligomerization and RNA binding (Dong et al., 2015). Most of Np1 was found to form an oligomerization arm of NP (Kirchdoerfer et al., 2015). A part of Np1 laid in N-terminal 1–24 residues which enhance ssRNA binding as well as control NP intermolecular interactions (Su et al., 2018; Dong et al., 2015). One residue of Np3 (160th) is amongst the four important nucleoprotein residues (160, 171, 174 and 248) responsible for RNA encapsidation (Albertini et al., 2006) and their deletion impairs EBOV replication (Noda et al., 2010). Residues of Np4 and Np5 (264, 268 and 316) are involved in the formation of a highly conserved hydrophobic pocket significant for RNA formation (Dong et al., 2015).

Objective 2

5.2. *In silico* analysis of peptide MHC interaction

Immune response is triggered when immunogenic peptides (epitopes) are presented by human leukocyte antigen (HLA) molecules to T_c and T_h cells. These HLA molecules are present on the surface of various body cells and are a translational result of mRNA transcribed from HLA alleles present on chromosome 6. HLA alleles are the most polymorphic loci among the different individuals of the world. Till July 2019, 23,907 HLA alleles (17,191 class I and 6,716 class II) have been reported in IPD/IMGT database (EMBL-

EBI, 2019) and this number is increasing periodically (Thude et al., 2005; Wang et al., 2019). Therefore, in a bid to develop a globally effective peptide vaccine, the vaccine candidate must exhibit interaction with multiple HLA alleles belonging to different populations of the world (Ling and Whitton, 1997). Hence, in the present study, peptide-HLA interaction studies were carried out by employing HLA coverage analysis (to assess the promiscuous nature of the peptides), population coverage analysis (to assess the potential of peptides as global vaccine candidates) and molecular docking (Autodock Vina and CABS-dock) analysis.

The selected peptides in this study contain multiple epitopes predicted by six different prediction tools. The multiple epitopic peptides are advantageous as they cover a wider array of HLA alleles. HLA coverage analysis from prediction tools revealed that most of the peptides have potential to interact with diverse HLA (HLA-A, HLA-B, HLA-DP, HLA-DQ and HLA-DR) alleles. All GP peptides (except Gp3) exhibited good coverage for both, HLA class I and II. Np2, Np3, Np4 and Np5 were found to contain epitopes predicted for all HLA types present in the prediction tools.

Population coverage analysis provides an additional perspective towards the expected immune response of peptides in different geographical populations. Population coverage analysis has been performed in earlier studies during identification of vaccine candidates against Hepatitis C, influenza and malaria (Dewi et al., 2019; Lohia and Baranwal, 2015; Pritam et al., 2019). In a recent study during development of a subunit vaccine against tuberculosis, the existence of race-specific functional binding motifs of HLA alleles was contemplated after the application of IEDB tool (Perez-Martinez et al., 2017). Therefore, population coverage analysis was carried out in the current study with the help of IEDB population coverage analysis tool. The population coverage of both, GP and NP peptides was found to be in the range of 69-100% while in most of the continents, it was observed to be ≥ 90% indicating the potential of identified peptide fragments to be a part of globally effective EBOV vaccine.

In order to be presented to T cells by HLA molecules, the peptides (epitopes) must bind with the HLA molecule inside the binding groove in a stable manner and in the right configuration. This forms the basis for generation of a robust T cell mediated immune response. Molecular docking is a reliable technique to study peptide-MHC binding efficacy and is deemed as valuable in computer aided drug design (Patronov et al., 2011). In the past, it has been applied during development of vaccine candidates against cancer, schistosomiasis, leishmaniasis and tuberculosis (Mahaddalkar et al., 2015; Oliveira et al., 2016; Agallou et al.,

2014; Singh et al., 2014). In a recent study conducted on H5N1 nucleoprotein, epitopes were identified by molecular docking and further tested for their immunogenicity in *in vivo* system (Hou et al., 2012).

Molecular docking was carried out by employing two tools viz., Autodock Vina and CABS-dock and the results were analysed in terms of binding energy and root mean square deviation (RMSD) respectively. PEPFOLD was employed to predict structure of the identified peptides. Most of the B-cell epitopes required to induce the desired humoral immunity levels need to maintain their native confirmation found in the protein. In contrast to humoral immunity, T-cell mediated immunity does not require the conformational presentation of T-cell epitopes. Therefore, processing longer peptides into shorter epitopes is crucial for generation of cellular immunity but not the conformation of an epitope (Skwarczynski and Toth, 2016). Hence, even though PEPFOLD prediction and folding is presently assumed for neutral pH while the pH of lyso-endosomal compartment ranges between 5-6, the peptide structure predictions made in the current study can be expected to elicit immune response. However, the peptide vaccine candidates can be customised by either flanking them with specific sequences or by stapling where in artificial chemical bonds are induced between distinct chains of amino acids to obtain the required secondary structure. Autodock Vina has been employed to study peptide-MHC interaction in various studies on Australian box jellyfish, *Toscana virus, Nipah virus* and *Campylobacter jejuni* (Alam and Ashraf, 2013; Jain et al., 2015; Sakib et al., 2014; Yasmin et al., 2016). It provides us with binding energy (Kcal/mol) which is the amount of energy released upon binding of peptide in the HLA binding groove. According to general consensus, higher the value of negative binding energy, more stable is the peptide-HLA complex. Recent reports on development of a synthetic peptide vaccine against *Nairovirus* (Tipu, 2016) and *Influenza virus* (Sankar et al., 2018) and a subunit vaccine against *Salmonella typhi* (Samykannu et al., 2019) have indicated the ability of CABS-dock tool in studying peptide-HLA interaction and binding efficacy. CABS-

dock provides us with RMSD values which indicate docking pose quality. The root-mean-square deviation (RMSD) is the measure of the average distance between the atoms of superimposed proteins. Typically, RMSD is used as a quantitative measure of similarity between two or more protein structures (Blaszczyk et al., 2016). In the current study, the interaction of native peptide with a specific HLA molecule was used as a reference to estimate the quality of the binding pose of the predicted epitope with the same HLA molecule. The lower the RMSD value, the better are the key binding interactions which indicate a high quality docking pose. RMSD values < 3 are considered high quality predictions while $3 \leq RMSD \geq 5.5$ are considered moderate quality predictions (Blaszczyk et al., 2016). In majority of cases for both, GP and NP peptides, the binding energy and RMSD values of peptides-HLA complexes were found to be in the range of corresponding native peptide-HLA complex values. Most of the peptides exhibited a RMSD < 3 and some peptide-HLA results presented better binding interactions than native peptides. Amongst the GP peptides, Gp1, Gp3 and Gp4 showed better RMSD values than native peptides for several HLA alleles. Peptides either exhibiting a RMSD >5 or binding outside the HLA binding groove were considered as non-binders. Gp2 represented a weaker interaction amongst GP peptides as it was found to be a non-binder for six HLA alleles and therefore, was not considered for *in vitro* analysis. In case of NP peptides, several peptide-HLA complexes exhibited better binding interaction as compared to corresponding native peptides but Np1 and Np4 did not show proper interaction with two HLA alleles while Np6 didnt interact with three HLA alleles. Moreover, Np2, Np3 and Np5 were shown to cover a large number of HLA alleles. Hence, these peptides were analysed further for evaluating their immune response in peripheral blood mononuclear cells.

Objective 3

5.3. Estimation of immunogenic response of *in silico* screened peptides in the *in vitro* system

The identification of potential peptide vaccine candidates based on numerous computational approaches has several advantages but the actual immune response needs to be validated in different experimental systems. In the current study, the efficacy of six selected peptides (Gp1, Gp3, Gp4, Np2, Np3 and Np5) was evaluated by measuring cell proliferation and IFN-γ production by repetitively peptide stimulated peripheral blood mononuclear cells obtained from healthy volunteers.

On continuous encounter with an antigen, first response of the lymphocytes is to undergo proliferation followed by differentiation to generate specialised cells which specifically identify and exterminate the pathogen. Thus, lymphocyte proliferation indicates the initiation of immune response towards immunogenic peptides. MTT [3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide] assay has proved to be a successful, safe and sensitive assay for measuring cell proliferation (Weichert et al., 1991). Earlier, MTT assay has been performed to ascertain the antigenic properties of hepatitis B surface antigen (do Livramento et al., 2013) and to measure the *Human papillomavirus* vaccine candidate induced PBMC proliferation (Goncalves et al., 2015; Liao et al., 2012). In the present study, all six peptides resulted in a significant peptide induced proliferation in at least 50% (5 out of 10) of healthy samples while Np3, Gp4 and Gp1 generated a significant response in ≥ 7 out of ten samples which shows the potential of these peptides to induce peptide specific immune cell proliferation.

IFN-γ is a soluble cytokine that plays a significant role in both, innate and adaptive immune response during viral infection. It is released by NK cells during innate immune response while T_c and T_h1 (a subclass of T_h cells) cells release IFN-γ during adaptive immune response (Kang et al., 2018). Though NK cells express IFN-γ mRNA in a constitutive manner, T cells require optimum proliferation and differentiation conditions to express the genes leading to IFN-γ production. IFN-γ helps to increase phagocytosis via macrophage activation, increase the proteolysis of viral proteins by increasing the activity of proteasomes as well as increase the number of HLA molecules expressed on the cell surface which results in better presentation of epitopes to T cells (Rhein et al., 2015). Therefore, the release of IFN-γ is vital for generation of a robust immune response. Several studies have been conducted to estimate

the IFN-γ production by antigen stimulated and proliferated T cells (Atsmon et al., 2014; Hou et al., 2012; Stoloff and Caparros-Wanderley, 2007). In a report, effect of propolis, an immunomodulatory bee product, on the leishmaniasis affected human PBMC proliferation followed by IFN-γ release measurement was studied (Amarante et al., 2012). Recently, influenza peptide vaccine candidates identified via *in silico* approach were validated for human PBMC proliferation with the help of MTT assay followed by IFN-γ release measurement with the help of ELISA (Lohia and Baranwal, 2018). In another study, cytokine measurement was carried out with the help of ELISA in PBMC stimulated with lys-gingipain synthetic peptides belonging to *Porphyromonas gingivalis* after 48 hours (Santos-Lima et al., 2019). In the present work, all peptides (except Np5) displayed a significant response in inducing peptide specific IFN-γ release in ≥ 70% samples (7 out of 10 healthy samples). Gp4 emerged as the most potent candidate as it was able to induce IFN-γ secretion in all the blood samples. Significant positive correlation was observed between peptide specific cell proliferation and IFN-γ production which jointly represent that the potential of these peptides to elicit immune response.

Though advantageous over traditional vaccine development approaches in many aspects, peptide based vaccine approach is riddled with its own challenges and limitations. They result in a weak immunogenic response indicating a need of adjuvant (agents with the clear ability to enhance the immune response against the antigen of interest) or delivery system (technology to administer or transport vaccine components). Also, peptides are susceptible to undergo enzymatic degradation as opposed to fully folded whole proteins. Immunogenicity of peptides can be improved in multiple ways such as addition of adjuvants and/or usage of a mixture of synthetic peptides containing the poly-epitopes expressed in various expression systems (Rosendahl Huber et al., 2014). Adjuvants act as immunostimulants in vaccine formulations to enhance the potency of immune response generated. MF59 and Montanide ISA 51VG are some of the examples of adjuvants being used for the T cell based peptide vaccines which are currently under clinical trials (Soema et al., 2015). Other than alum, adjuvants such as squalene-based emulsions AF03 and AS03 as well as monophosphoryl lipid A containing AS04 have been licensed for human application and present possible options for the proposed vaccine candidates (Skwarczynski and Toth, 2016). Conserved T cell epitopes from various internal proteins (internal as well as surface proteins) or multiple copies of single immunodominant epitope in tandem have also been used to enhance the immunogenicity of peptide immunogens in *in vivo* systems (Rosendahl Huber et al., 2015;

Subbarao and Matsuoka, 2013; Wiersma et al., 2015). Different platforms for vaccine delivery that can be used to enhance the efficiency of same epitopes include soluble peptides with adjuvant, subunit or domain epitopes fused to a carrier protein, nanoparticles and virus-like particles (VLPs) (Zhang et al., 2014). Further, prime-boost strategy of immunisation has also been used as one of the procedures to enhance memory T cell count (Woodland, 2004). Therefore, even though peptide vaccine may not fully replace the current trend of protein-based vaccine development and other approaches, however, fully synthetic peptide-based vaccines are the potential future of vaccination as indicated by the exciting developments in this field.

To conclude, six highly conserved Ebola virus peptide vaccine candidates containing multiple T ($CD8^+$ and $CD4^+$) and B cell epitopes are being reported in this study. These peptides are devoid of any undesired responses such as autoimmune, allergic or toxic response and are capable of interacting with a wide array of HLA molecules as indicated by HLA coverage and molecular docking analysis. Also, these peptides cover a large population base as indicated by population coverage analysis. Further, most of the peptides were found to induce cell proliferation and IFN-γ production in majority of PBMC samples. Therefore, these peptides may be considered as potent candidates for development of a globally effective EBOV peptide vaccine subject to further validation in experimental animal models.

Summary

The unprecedented rise in Ebola virus epidemics poses a major health concern worldwide among humans. Owing to multiple host immune evasion methods employed by the virus and the limitations of traditional vaccine development approaches, finding a globally effective and reliable counter measure against Ebola virus remains a challenge. Development of a peptide based vaccine consisting of highly conserved immunogenic regions (epitopes) is one of the feasible options. The conventional approach of defining the epitope experimentally is an exhaustive and expensive process. Immunoinformatics or computational immunology has played an instrumental role in screening the epitopes which reduces the number of peptides needed for experimental validation for vaccine design.

All available full length protein sequences viz., 1092 and 2407 sequennces of glycoprotein (GP) and nucleoprotein (NP) respectively were retrieved from NCBI and viprbrc databases and manually screened for redundancy to obtain 173 and 195 unique GP and NP sequences respectively. Sequences were aligned with the help of MUSCLE tool and regions exhibiting ≥ 90% conservancy were identified via AVANA. Ten and four conserved peptides were identified for GP and NP respectively which were used as input for T cell epitope prediction by using a consensus approach wherein six prediction tools were utilized. 84 GP and 105 NP epitopes were found to be commonly predicted by the $CD8^+$ T cell (HLA class I) epitope prediction tools (SYFPEITHI, NetCTL 1.2 and IEDB consensus) while 72 GP and 79 NP epitopes were predicted by $CD4^+$ T-cell (HLA class II) epitope prediction tools (MHC2Pred, Propred and IEDB consensus). After merging the overlapping epitopes, fifteen GP and eighteen NP peptide fragments containing multiple $CD8^+$ T-cell epitopes while thirteen GP and fifteen NP fragments containing multiple $CD4^+$ T-cell epitopes were obtained. Regions commonly present in these peptide fragments were looked for so as to obtain six GP and twelve NP peptide fragments containing both types of epitopes. The identified peptides were searched for presence of B cell epitopes (ABCPred) as well as screened for the undesirable capability to generate autoimmune (BLAST), allergic (AlgPred) and toxic (ToxinPred) response. Four GP and six NP peptides containing multiple T and B cell epitopes and devoid of undesired responses were identified. All the peptides were found to be nearly 100% conserved in *Zaire ebolavirus*, the species responsible for maximum number of outbreaks while Gp3 and Np3 exhibited 100% conservancy in all human infecting Ebola species. In a

bid to find a cross-reactive vaccine candidate, the identified peptides were searched in other *Filoviridae* members (*Lloviu virus* and *Margburg virus*) wherein they exhibited minor variations or could not be located. Moreover, these peptides were found to be a part of GP and NP domains crucial for viral functioning.

Immunogenic response against antigen varies in different populations owing to high HLA polymorphism. Upon HLA coverage analysis of the identified peptides, majority of the peptides were predicted to bind with diverse HLA (HLA-A, HLA-B, HLA-DP, HLA-DQ and HLA-DR) types. Population coverage analysis which emphasizes the global coverage of the peptides revealed that all the peptides exhibited 69-100% coverage for four (Asia, America, Africa and European) continents with an average population coverage of 90% (GP peptides) and ≥85% (NP peptides). Autodock Vina (Binding energy) and CABS-dock (RMSD) tools were employed to dock the identified peptides with high resolution HLA class I and II molecules (nine each). In most of the cases, the binding energy and RMSD values of peptide-HLA complex were found to be near the values of corresponding native peptide-HLA complex indicating the strong potential of peptides to be presented by HLA molecules. Moreover, in some cases, the identified peptides performed better than native peptide indicating a better binding potential for various HLA molecules. However, Gp2 and Np6 were found to be non-binders for six and three HLA molecules respectitvely while Np1 and Np4 acted as non-binder two HLA alleles and therefore, these peptides were not considered for *in vitro* analysis

Further, six peptides (Gp1, Gp3, Gp4, Np2, Np3 and Np5) were commercially synthesized and assessed for their immunogenic response via *in vitro* experimentation involving estimation of lymphocyte proliferation (MTT assay) and IFN-γ production (ELISA) by peptide stimulated peripheral blood mononuclear cells obtained from ten different healthy blood samples. All peptides (except Np5) resulted in cell proliferation and IFN-γ secretion in ≥ 60% samples (six out of ten samples). Gp4 (RATTELRTFSILNRKAIDF) and Np3 (FLSFASLFLPKL) emerged as the best peptides during *in vitro* analysis as they exhibited a significantly positive IFN-γ production response for 100% and 90% samples respectively.

Thus, these GP and NP peptides are proposed to be validated further in animal models to inculcate them in design of a synthetic peptide vaccine against EBOV and related species.

References

1. Abu-Haraz A.H., Abd-eirahman K.A., Ibrahim M.S., Hussein W.H., Mohammed M.S. et al. 2017. Multi Epitope Peptide Vaccine Prediction against Sudan Ebola Virus Using Immuno-Informatics Approaches. *Adv Tech Biol Med*, 5(1), 203.
2. Adu-Gyamfi E., Digman M.A., Gratton E. and Stahelin R.V. 2012. Investigation of Ebola VP40 assembly and oligomerization in live cells using number and brightness analysis. *Biophys J*, 102(11), 2517–2525.
3. Adu-Gyamfi E., Johnson K.A., Fraser M.E., Scott J.L., Soni S.P. et al. 2015. Host Cell Plasma Membrane Phosphatidylserine Regulates the Assembly and Budding of Ebola Virus. *J Virol*, 89(94), 9440–9453.
4. Afley P., Dohre S.K., Prasad G.B. and Kumar S. 2015. Prediction of T cell epitopes of Brucella abortus and evaluation of their protective role in mice. *Appl Microbiol Biotechnol*, 99(18), 7625–7637.
5. Agallou M., Athanasiou E., Koutsoni O., Dotsika E. and Karagouni E. 2014. Experimental validation of multi-epitope peptides including promising MHC Class I- and II-restricted epitopes of four known Leishmania infantum proteins. *Front Immunol*, 5, 268.
6. Agnandji S.T., Huttner A., Zinser M.E., Njuguna P., Dahlke C. et al. 2015. Phase 1 trials of rVSV Ebola vaccine in Africa and Europe. *New Engl J Med*, 374, 1647–1660.
7. Akinfeyeva L.A., Aksyonova O.I., Vasilyevich I.V., Ginko Z.I., Zarkov K.A. et al. 2005. A case of Ebola hemorrhagic fever. *Infektsionnye Bolezni (Moscow)*, 3(1), 85–88.
8. Alam M.J. and Ashraf K.U.M. 2013. Prediction of an epitope-based computational vaccine strategy for gaining concurrent immunization against the venom proteins of Australian box Jellyfish. *Toxicol Int*, 20(3), 235-253.
9. Alayyoubi M., Leser G.P., Kors C.A. and Lamb R.A. 2015. Structure of the paramyxovirus parainfluenza virus 5 nucleoprotein-RNA complex. *Proc Natl Acad Sci USA*, 112(14), E1792–E1799.
10. Albarino C.G., Shoemaker T., Khristova M.L., Wamala J.F., Muyembe J.J. et al. 2013. Genomic analysis of filoviruses associated with four viral hemorrhagic fever outbreaks in Uganda and the Democratic Republic of the Congo in 2012. *Virology*, 442(2), 97-100.

11. Albertini A.A., Wernimont A.K., Muziol T., Ravelli R.B., Clapier C.R. et al. 2006. Crystal structure of the rabies virus nucleoprotein-RNA complex. *Science*, 313(5785), 360–363.
12. Ali M.T. and Islam M.O. 2015. A Highly Conserved GEQYQQLR Epitope Has Been Identified in the Nucleoprotein of Ebola Virus by Using an In Silico Approach. *Adv Bioinformatics*, 2015, 278197.
13. Ali S., Asad M.H.H.B., Maity S., Zada W., Rizvanov A.A. et al. 2019. Fluoro-benzimidazole derivatives to cure Alzheimer's disease: In-silico studies, synthesis, structure-activity relationship and in vivo evaluation for β secretase enzyme inhibition. *Bioorg Chem*, 88, 102936.
14. Almeida R.R., Rosa D.S., Ribeiro S.P., Santana V.C., Kallas E.G. et al. 2012. Broad and cross-clade CD4+ T-cell responses elicited by a DNA vaccine encoding highly conserved and promiscuous HIV-1 M-group consensus peptides. *PLoS ONE*, 7, e45267.
15. Altschul S.F., Gish W., Miller W., Myers E.W. and Lipman D.J. 1990. Basic local alignment search tool. *J Mol Biol*, 215 (3), 403–410.
16. Amarante M.K., Watanabe M.A., Conchon-Costa I., Fiori L.L., Oda J.M. et al. 2012. The effect of propolis on CCL5 and IFN-γ expression by peripheral blood mononuclear cells from leishmaniasis patients. *J Pharm Pharmacol*, 64(1), 154-160.
17. Andreatta M. and Nielsen M. 2016. Gapped sequence alignment using artificial neural networks: Application to the MHC class I system. *Bioinformatics*, 32(4), 511–517.
18. Anna C. 2018. Health Workers Trying to Contain Congo's Ebola Outbreak Are Being Attacked Weekly. Time, Time.
19. Ariza A., Tanner S.J., Walter C.T., Dent K.C., Shepherd D.A. et al. 2013. Nucleocapsid protein structures from orthobunyaviruses reveal insight into ribonucleoprotein architecture and RNA polymerization. *Nucleic Acids Res*, 41(11), 5912–5926.
20. Aruga A., Takeshita N., Kotera Y., Okuyama R., Matsushita N. et al. 2014. Phase I clinical trial of multiple-peptide vaccination for patients with advanced biliary tract cancer. *J Transl Med*, 12, 61.
21. Atsmon J., Caraco Y., Ziv-Sefer S., Shaikevich D., Abramov E. et al. 2014. Priming by a novel universal influenza vaccine (Multimeric-001)—a gateway for improving immune response in the elderly population. *Vaccine*, 32, 5816–5823.
22. Ayub G. and Waheed Y. 2016. Sequence analysis of the L protein of the Ebola 2014 outbreak: Insight into conserved regions and mutations. *Mol Med Rep*, 13(6), 4821-4826.

23. Azarian T., Presti A.L., Giovanetti M., Cella E., Rife B. et al. 2015. Impact of spatial dispersion, evolution, and selection on Ebola Zaire Virus epidemic waves. *Sci Rep,* 5, 10170.

24. Backert L. and Kohlbacher O. 2015. Immunoinformatics and epitope prediction in the age of genomic medicine. *Genome Med*, 7(1), 119.

25. Baker L.E., Ellena J.F., Handing K.B., Derewenda U., Utepbergenov D. et al. 2016. Molecular architecture of the nucleoprotein C-terminal domain from the Ebola and Marburg viruses. *Acta Crystallogr D Struct Biol*, 72, 49–58.

26. Bale S., Julien J.P., Bornholdt Z.A., Krois A.S., Wilson I.A. et al. 2013. Ebolavirus VP35 coats the backbone of double-stranded RNA for interferon antagonism. *J Virol,* 87, 10385–10388.

27. Balzarini J. 2007. Targeting the glycans of glycoproteins: a novel paradigm for antiviral therapy. *Nat Rev Microbiol*, 5(8), 583–597.

28. Banadyga L., Hoenen T., Ambroggio X., Dunham E., Groseth A. et al. 2017. Ebola virus VP24 interacts with NP to facilitate nucleocapsid assembly and genome packaging. *Sci Rep,* 7(1), 7698.

29. Bar S., Takada A., Kawaoka Y. and Alizon M. 2006. Detection of cell-cell fusion mediated by Ebola virus glycoproteins. *J Virol,* 80(6), 2815–2822.

30. Barkhouse D.A., Garcia S.A., Bongiorno E.K., Lebrun A., Faber M. et al. 2015. Expression of interferon gamma by a recombinant rabies virus strongly attenuates the pathogenicity of the virus via induction of type I interferon. *J Virol*, 89(1), 312–322.

31. Baron R.C., McCormick J.B. and Zubeir O.A. 1983. Ebola virus disease in southern Sudan: hospital dissemination and intrafamilial spread. *Bulletin of the World Health Organization*, 61(6), 997-1003.

32. Barrette R.W., Metwally S.A., Rowland J.M., Xu L., Zaki S.R. et al. 2009. Discovery of swine as a host for the Reston ebolavirus. *Science,* 325(5937), 204–206.

33. Barton C., Kouokam J.C., Lasnik A.B., Foreman O., Cambon A. et al. 2014. Activity of and effect of subcutaneous treatment with the broad-spectrum antiviral lectin griffithsin in two laboratory rodent models. *Antimicrob Agents Chemother*, 58(1), 120–127.

34. Basler C.F. 2015. Innate immune evasion by filoviruses. *Virology*, 479-480, 122-130.

35. Basler C.F. and Amarasinghe G.K. 2009. Evasion of interferon responses by Ebola and Marburg viruses. *J Interferon Cytokine Res*, 29(9), 511–520.

36. Bavari S., Bosio C.M., Wiegand E., Ruthel G., Will A.B. et al. 2002. Lipid raft microdomains: a gateway for compartmentalized trafficking of Ebola and Marburg viruses. *J Exp Med*, 195(5), 593–602.

37. Becker S., Rinne C., Hofs€ass U., Klenk H.D. and M€uhlberger E. 1998. Interactions of Marburg virus nucleocapsid proteins. *Virology*, 249(2), 406–417.

38. Bell B.P., Damon I.K., Jernigan D.B., Kenyon T.A., Nichol S.T. et al. 2016. Overview, Control Strategies, and Lessons Learned in the CDC Response to the 2014–2016 Ebola Epidemic. *Morbidity and Mortality Weekly Report*, 65(3), 4-11.

39. Ben-Asouli Y., Banai Y., Pel-Or Y., Shir A. and Kaempfer R. 2002. Human interferon-gamma mRNA autoregulates its translation through a pseudoknot that activates the interferon-inducible protein kinase PKR. *Cell*, 108(2), 221–232.

40. Bengtsson K.L., Song H., Stertman L., Liu Y., Flyer D.C. et al. 2016. Matrix-M adjuvant enhances antibody, cellular and protective immune responses of a Zaire Ebola/Makona virus glycoprotein (GP) nanoparticle vaccine in mice. *Vaccine*, 34(16), 1927-1935.

41. Bennett M. and Gilroy D.W. 2016. Lipid Mediators in Inflammation. *Microbiol Spectr*, 4(6), MCHD-0035-2016.

42. Bergmann-Leitner E.S., Chaudhury S., Steers N.J., Sabato M., Delvecchio V. et al. 2013. Computational and experimental validation of B and T-cell epitopes of the *in vivo* immune response to a novel malarial antigen. *PLoS One*, 8(8), e71610 .

43. Berlanda Scorza F., Tsvetnitsky V. and Donnelly J.J. 2016. Universal influenza vaccines: shifting to better vaccines. *Vaccine*, 34(26), 2926-2933.

44. Berman H.M., Westbrook J., Feng Z., Gilliland G., Bhat T.N. et al. 2000. The protein data bank. *Nucleic Acids Res*, 28(1), 235–242.

45. Bermejo M., Rodriguez-Teijeiro J.D., Illera G., Barroso A., Vila C. et al. 2006. Ebola outbreak killed 5000 gorillas. *Science*, 314(5805), 1564.

46. Bharat T.A., Noda T., Riches J.D., Kraehling V., Kolesnikova L. et al. 2012. Structural dissection of Ebola virus and its assembly determinants using cryo-electron tomography. *Proc Natl Acad Sci USA*, 109(11), 4275–4280.

47. Bhasin M. and Raghava G.P. 2004. SVM based method for predicting HLA-DRB1 *0401 binding peptides in an antigen sequence. *Bioinformatics*, 20(3), 421–423.

48. Bhasin M. and Raghava G.P. 2007. A hybrid approach for predicting promiscuous MHC class I restricted T cell epitopes. *J Biosci*, 32(1), 31–42.

49. Bhattacharjee B., Talambedu U., Sadegh S., Goyal A.K., Pande V. et al. 2011. Computer aided screening of inhibitors to 5-α reductase type 2 for prostate. *Bioinformation,* 6(7), 262–265.

50. Binning J.M., Wang T., Luthra P., Shabman R.S., Borek D.M. et al. 2013. Development of rna aptamers targeting Ebola virus vp35. *Biochemistry*, 52(47), 8406–8419.

51. Blaney J.E., Wirblich C., Papaneri A.B., Johnson R.F., Myers C.J. et al. 2011. Inactivated or live-attenuated bivalent vaccines that confer protection against rabies and Ebola viruses. *J Virol,* 85(20), 10605–10616.

52. Blaney J.E., Marzi A., Willet M., Papaneri A.B., Wirblich C. et al. 2013. Antibody quality and protection from lethal Ebola virus challenge in nonhuman primates immunized with rabies virus based bivalent vaccine. *PLoS Pathog*, 9(5), e1003389.

53. Blaszczyk M., Kurcinski M., Kouza M., Wieteska L., Debinski A. et al. 2016. Modeling of protein-peptide interactions using the CABS-dock web server for binding site search and flexible docking. *Methods,* 93, 72–83.

54. Borisevich I.V., Markin V.A., Firsova I.V., Evseev A.A., Khamitov R.A. et al. 2006. Hemorrhagic (Marburg, Ebola, Lassa, and Bolivian) fevers: epidemiology, clinical pictures, and treatment. *Vopr Virusol,* 51(5), 8–16.

55. Bornholdt Z.A., Noda T., Abelson D.M., Halfmann P., Wood M.R. et al. 2013. Structural rearrangement of Ebola virus VP40 begets multiple functions in the virus life cycle. *Cell*, 154(4), 763–774.

56. Bounds C.E., Terry F.E., Moise L., Hannaman D., Martin W.D. et al. 2017. An immunoinformatics-derived DNA vaccine encoding human class II T cell epitopes of Ebola virus, Sudan virus, and Venezuelan equine encephalitis virus is immunogenic in HLA transgenic mice. *Hum Vaccin Immunother*, 13(12), 2824–2836.

57. Brannan J.M., He S., Howell K.A., Prugar L.l., Zhu W. et al. 2019. Post-exposure immunotherapy for two ebolaviruses and Marburg virus in nonhuman primates. *Nat Commun,* 10(1), 105.

58. Bray M. and Geisbert T.W. 2005. Ebola virus: the role of macrophages and dendritic cells in the pathogenesis of Ebola hemorrhagic fever. *Int J Biochem Cell Biol*, 37(8), 1560–1566.

59. Brintnell W., Bell D.A., Hill J.A., Jevnikar A.M., Sette A. et al. 2007. The influence of MHC class II molecules containing the rheumatoid arthritis shared epitope on the immune response to aggrecan G1 and its peptides. *Scand J Immunol*, 65(5), 444–452.

60. Brubaker S.W., Bonham K.S., Zanoni I. and Kagan J.C. 2015. Innate Immune Pattern Recognition: A Cell Biological Perspective. *Annu Rev Immunol,* 33, 257-290.

61. Bruhn J.F., Kirchdoerfer R.N., Urata S.M., Li S., Tickle I.J. et al. 2017. Crystal Structure of the Marburg Virus VP35 Oligomerization Domain. *J Virol,* 91(2), e01085-e01116.

62. Bui H.H., Sidney J., Dinh K., Southwood S., Newman M.J. 2006. Predicting population coverage of T-cell epitope-based diagnostics and vaccines. *BMC Bioinform,* 7(1), 153.

63. Bui S. and Steiner R.A. 2016. New insight into cofactor-free oxygenation from combined experimental and computational approaches. *Curr Opin Struct Biol*, 41, 109–118.

64. Bukreyev A., Volchkov V., Blinov V., Dryga S. and Netesov S. 1995. The complete nucleotide sequence of the Popp (1967) strain of Marburg virus: a comparison with the Musoke (1980) strain. *Arch Virol,* 140(9), 1589-1600.

65. Bukreyev A., Yang L., Zaki S.R., Shieh W.J., Rollin P.E. et al. 2006. A single intranasal inoculation with a paramyxovirus-vectored vaccine protects guinea pigs against a lethal-dose Ebola virus challenge. *J Virol,* 80(5), 2267–2279.

66. Bukreyev A., Marzi A., Feldmann F., Zhang L., Yang L. et al. 2009. Chimeric human parainfluenza virus bearing the Ebola virus glycoprotein as the sole surface protein is immunogenic and highly protective against Ebola virus challenge. *Virology,* 383(2), 348–361.

67. Bukreyev A.A., Dinapoli J.M., Yang L., Murphy B.R. and Collins P.L. 2010. Mucosal parainfluenza virus-vectored vaccine against Ebola virus replicates in the respiratory tract of vector-immune monkeys and is immunogenic. *Virology,* 399(2), 290–298.

68. Sánchez-Burgos G., Ramos-Castañeda J., Cedillo-Rivera R. and Dumonteil E. 2010. Immunogenicity of novel Dengue virus epitopes identified by bioinformatic analysis. *Virus Res,* 153 (1), 113-20.

69. Cantoni D. and Rossman J.S. 2018. Ebolaviruses: New roles for old proteins. *PLoS Negl Trop Dis,* 12(5), e0006349.

70. Cardenas W.B., Loo Y.M., Gale M.Jr., Hartman A.L., Kimberlin C.R. et al. 2006. Ebola virus VP35 protein binds double-stranded RNA and inhibits alpha/beta interferon production induced by RIG-I signaling. *J Virol,* 80(11), 5168–5178.

71. Carette J.E., Raaben M., Wong A.C., Herbert A.S., Obernosterer G. et al. 2011. Ebola virus entry requires the cholesterol transporter Niemann–Pick C1. *Nature*, 477(7364), 340–343.

72. Carroll M.W., Matthews D.A., Hiscox J.A., Elmore M.J., Pollakis G. et al. 2015. Temporal and spatial analysis of the 2014–2015 Ebola virus outbreak in West Africa. *Nature*, 524(7563), 97–101.

73. CDC, 2019. Ebola Virus Disease Distribution Map: Cases of Ebola Virus Disease in Africa Since 1976. https://www.cdc.gov/vhf/ebola/history/distribution-map.html

74. CDC, 2019. Years of Ebola Virus Disease Outbreaks. https://www.cdc.gov/vhf/ebola/history/chronology.html

75. Chan S.Y., Empig C.J., Welte F.J., Speck R.F., Schmaljohn A. et al. 2001. Folate receptor-alpha is a cofactor for cellular entry by Marburg and Ebola viruses. *Cell*, 106(1), 117–126.

76. Changula K., Yoshida R., Noyori O., Marzi A., Miyamoto H. et al. 2013. Mapping of conserved and species-specific antibody epitopes on the Ebola virus nucleoprotein. *Virus Res*, 176(1-2), 83-90.

77. Chen W., Sun P., Lu Y., Guo W.W., Huang Y. et al. 2011. MimoPro: a more efficient web-based tool for epitope prediction using phage display libraries. *BMC Bioinformatics*, 12(1), 199.

78. Chen F., Zhai M.X., Zhu Y.H., Qi Y.M., Zhai W.J. et al. 2012. In vitro and in vivo identification of a novel cytotoxic T lymphocyte epitope from Rv3425 of Mycobacterium tuberculosis. *Microbiol Immunol*, 56(8), 548-553.

79. ClinicalTrials, 2017. A Study to Evaluate A Range of Dose Levels of Ad26.ZEBOV and MVA-BN-Filo in Healthy Adult Participants. https://clinicaltrials.gov/ct2/show/NCT02543567

80. ClinicalTrials, 2017. Clinical Trial of Ebola Vaccines cAd3-EBO, cAd3-EBOZ and MVA-EbolaZ in Healthy Adults in Uganda. https://clinicaltrials.gov/ct2/show/NCT02354404

81. ClinicalTrials, 2017. Clinical Trial to Evaluate the Efficacy and Safety of GV1001 in Alzheimer Patients. https://clinicaltrials.gov/ct2/show/NCT03184467

82. ClinicalTrials, 2019. Open-Label Study of INO-4212 With or Without INO-9012, Administered IM or ID Followed by Electroporation in Healthy Volunteers. https://clinicaltrials.gov/ct2/show/NCT02464670

83. Cohen A.D. and Shoenfeld Y. 1996. Vaccine-induced autoimmunity. *J Autoimmun*, 9(6), 699–703.

84. Collar A.L., Clarke E.C., Anaya E., Merrill D., Yarborough S. et al. 2017. Comparison of N- and O-linked glycosylation patterns of ebolavirus glycoproteins. *Virology*, 502, 39–47.

85. Crary S.M., Towner J.S., Honig J.E., Shoemaker T.R. and Nichol S.T. 2003. Analysis of the role of predicted RNA secondary structures in Ebola virus replication. *Virology* 306(2), 210-218.

86. Croyle M.A., Patel A., Tran K.N., Gray M., Zhang Y. et al. 2008. Nasal delivery of an adenovirus-based vaccine bypasses pre-existing immunity to the vaccine carrier and improves the immune response in mice. *PLoS One*, 3(10), e3548.

87. de Wit E., Marzi A., Bushmaker T., Brining D., Scott D. et al. 2015. Safety of recombinant VSV-Ebola virus vaccine vector in pigs. *Emerg Infect Dis,* 21(4), 702–704.

88. Delos S.E., Brecher M.B., Chen Z., Melder D.C., Federspiel M.J. et al. 2008. Cysteines flanking the internal fusion peptide are required for the avian sarcoma/leukosis virus glycoprotein to mediate the lipid mixing stage of fusion with high efficiency. *J Virol*, 82(6), 3131–3134.

89. Dessen A., Volchkov V., Dolnik O., Klenk H.D. and Weissenhorn W. 2000. Crystal structure of the matrix protein VP40 from Ebola virus. *EMBO J*, 19(16), 4228–4236.

90. Dewi S.K., Ali S. and Prasasty V.D. 2019. Broad Spectrum Peptide Vaccine Design Against Hepatitis C Virus. *Curr Comput Aided Drug Des*, 15(2), 120-135.

91. Dhama K., Karthik K., Khandia R., Chakraborty S., Munjal A. et al. 2018. Advances in Designing and Developing Vaccines, Drugs, and Therapies to Counter Ebola Virus. *Front Immunol*, 9, 1803.

92. Dhiman G., Lohia N., Jain S. and Baranwal M. 2016. Metadherin peptides containing CD4 + and CD8 + T cell epitopes as a therapeutic vaccine candidate against cancer. *Microbiol Immunol*, 60(9), 646–652.

93. Diehl W.E., Lin A.E., Grubaugh N.D., Carvalho L.M., Kim K. et al. 2016. Ebola virus glycoprotein with increased infectivity dominated the 2013–2016 epidemic. *Cell*, 167(4), 1088–1098.

94. Dikhit M.R., Kumar S., Vijaymahantesh., Sahoo B.R., Mansuri R. et al. 2015. Computational elucidation of potential antigenic CTL epitopes in Ebola virus. *Infect Genet Evol*, 36, 369–375.

95. Dikhit M.R., Kumar A., Das S., Dehury B., Rout A.K. et al. 2017. Identification of Potential MHC Class-II-restricted epitopes derived from Leishmania donovani antigens by reverse vaccinology and evaluation of Their CD4 + T-cell responsiveness against visceral leishmaniasis. *Front Immunol*, 8, 1763.

96. Dimitrov I., Garnev P., Flower D.R. and Doytchinova I. 2010. MHC class II binding prediction-A little help from a friend. *J Biomed Biotechnol*, 2010, 705821.

97. Dixon M.G. and Schafer I.J. 2014. Ebola viral disease outbreak – West Africa, 2014. *MMWR Morb Mortal Wkly Rep*, 63(25), 548–551.

98. Do Livramento A., Sampaio J., Schultz J., Batista K.Z.S., Treitinger A. et al. 2013. In vitro lymphocyte stimulation by recombinant hepatitis B surface antigen: A tool to detect the persistence of cellular immunity after vaccination. *J Virol Methods,* 193(2), 572-578.

99. Dolzhikova I.V., Zubkova O.V., Tukhvatulin A.I., Dzharullaeva A.S., Tukhvatulina N.M. et al. 2017. Safety and immunogenicity of GamEvac-Combi, a heterologous VSV- and Ad5-vectored Ebola vaccine: an open phase I/II trial in healthy adults in Russia. *Hum Vaccin Immunother*, 13(3), 613–620.

100. Domi A., Feldmann F., Basu R., McCurley N., Shifflett K. et al. 2018. A single dose of modified vaccinia Ankara expressing Ebola virus like particles protects nonhuman primates from lethal Ebola virus challenge. *Sci Rep*, 8(1), 864.

101. Dong J.Y., Dubay J.W., Perez L.G. and Hunter E. 1992. Mutations within the proteolytic cleavage site of the Rous sarcoma virus glycoprotein define a requirement for dibasic residues for intracellular cleavage. *J Virol*, 66(2), 865–874.

102. Dong S., Yang P., Li G., Liu B., Wang W. et al. 2015. Insight into the Ebola virus nucleocapsid assembly mechanism: crystal structure of Ebola virus nucleoprotein core domain at 1.8 Å resolution. *Protein Cell*, 6(5), 351–362.

103. Donnes P. and Elofsson A. 2001. Prediction of MHC class I binding peptides, using SVMHC. *BMC Bioinformatics*, 3, 25.

104. Doytchinova I.A., Guan P. and Flower D.R. 2006. EpiJen: a server for multi-step T cell epitope prediction. *BMC Bioinformatics*, 7, 131.

105. Durbin A.P., Skiadopoulos M.H., McAuliffe J.M., Riggs J.M., Surman S.R. et al. 2000. Human parainfluenza virus type 3 (PIV3) expressing the hemagglutinin protein of measles virus provides a potential method for immunization against measles virus and PIV3 in early infancy. *J Virol,* 74(15), 6821–6831.

106. Dutta D.K., Rhodes K. and Wood S.C. 2015. In silico prediction of Ebola Zaire GP (1,2) immuno-dominant epitopes for the Balb/c mouse. *BMC Immunol,* 16, 59.

107. Duvvuri V.R., Duvvuri B., Jamnik V., Gubbay J.B., Wu J. et al. 2013. T Cell memory to evolutionarily conserved and shared hemagglutinin epitopes of H1N1 viruses: A pilot scale study. *BMC Infect Dis*, 13, 204 .

108. Dye J.M., Herbert A.S., Kuehne A.I., Barth J.F., Muhammad M.A. et al. 2012. Postexposure antibody prophylaxis protects nonhuman primates from filovirus disease. *Proc Natl Acad Sci USA*, 109(13), 5034–5039.

109. Dziubanska P.J., Derewenda U., Ellena J.F., Engel D.A. and Derewenda Z.S. 2014. The structure of the C-terminal domain of the Zaire ebolavirus nucleoprotein. *Acta Crystallogr D Biol Crystallogr*, 70(9), 2420–2429.

110. 2014. Ebola: protection of health workers on the front line. *Lancet*, 384, 470. *(No authors listed)*

111. Edgar R.C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*, 32(5), 1792–1797.

112. Ekins S. and Coffee M. 2015. FDA approved drugs as potential Ebola treatments. *Version 2. F1000Res*, 4, 48.

113. Elisha A. and Adegboro B. 2014. Ebola virus diseases. *Afr J Cln Exper Microbiol*, 15(3), 178.

114. Elliott L.H., Kiley M.P. and McCormick J.B. 1985. Descriptive analysis of Ebola virus proteins. *Virology*, 147(1), 169–176.

115. Emanuel J., Marzi A. and Feldmann H. 2018. Filoviruses: Ecology, Molecular Biology, and Evolution. *Adv Virus Res*, 100, 189-221.

116. Emond R.T., Evans B., Bowen E.T. and Lloyd G. 1977. A case of Ebola virus infection. *Br Med J*, 2(6086), 541-544.

117. Facktor M.A., Bernstein R.A. and Fireman P. 1973. Hypersensitivity to tetanus toxoid. *J Allergy Clin Immunol*, 52(1), 1-12.

118. Falzarano D., Krokhin O., Van Domselaar G., Wolf K., Seebach J. et al. 2007. Ebola sGP—the first viral glycoprotein shown to be C-mannosylated. *Virology*, 368(1), 83–90.

119. Fan H., Du X., Zhang J., Zheng H., Lu X. et al. 2017. Selective inhibition of Ebola entry with selective estrogen receptor modulators by disrupting the endolysosomal calcium. *Sci Rep*, 7, 41226.

120. Farooq F., Beck K., Paolino K.M., Phillips R., Waters N.C. et al. 2016. Circulating follicular T helper cells and cytokine profile in humans following vaccination with the rVSV-ZEBOV Ebola vaccine. *Sci Rep*, 6, 27944.

121. Fed. Sel. Agent Program, 2018. Select agents and toxins list. CFR 7 Part 331, 9 Part 121, 42 Part 73, U. S. Dep. Health Hum. Serv., U. S. Dep. Agric., Washington, DC. https://www.selectagents.gov/SelectAgentsandToxinsList.html

122. Feldmann H., M€uhlberger E., Randolf A., Will C., Kiley M.P. et al. 1992. Marburg virus, a filovirus: messenger RNAs, gene order, and regulatory elements of the replication cycle. *Virus Res*, 24(1), 1-19.

123. Feldmann H., Nichol S.T., Klenk H.D., Peters C.J. and Sanchez A. 1994. Characterization of filoviruses based on differences in structure and antigenicity of the virion glycoprotein. *Virology*, 199(2), 469–473.

124. Feldmann H. and Geisbert T.W. 2011. Ebola haemorrhagic fever. *Lancet*, 377(9768), 849–862.

125. Feldmann H., Feldmann F. and Marzi A. 2018. Ebola: Lessons on Vaccine Development. *Annu Rev Microbiol*, 72, 423-446.

126. Fenstermaker R.A., Ciesielski M.J., Qiu J., Yang N., Frank C.L. et al. 2016. Clinical study of a survivin long peptide vaccine (SurVaxM) in patients with recurrent malignant glioma. *Cancer Immunol. Immunother*, 65 (11), 1339-1352.

127. Folayan M.O., Yakubu A., Haire B. and Peterson K. 2016. Ebola vaccine development plan: Ethics, concerns and proposed measures. *BMC Med Ethics*, 17, 10.

128. Formenty P., Boesch C., Wyers M., Steiner C., Donati F. et al. 1999. Ebola virus outbreak among wild chimpanzees living in a rain forest of Cote d'Ivoire. *J Infect Dis*, 179(1), S120–126.

129. Formenty P., Libama F., Epelboin A., Allarangar Y., Leroy E. et al. 2003. Outbreak of Ebola hemorrhagic fever in the Republic of the Congo, 2003: a new strategy? *Méd Trop (Marseille)*, 63(3), 291-295.

130. Forrester. 2018. Ebola virus and persistent chronic infection: when does replication cease? *Ann Transl Med*, 6(1), S39.

131. Fosgerau K. and Hoffmann T. 2015. Peptide therapeutics: current status and future directions. *Drug Discovery Today*, 20(1), 122.

132. Fowler R.A., Fletcher T., Fischer W.A., Lamontagne F., Jacob S. et al. 2014. Caring for critically ill patients with Ebola virus disease. Perspectives from West Africa. *Am J Respir Crit Care Med*, 190(7), 733–737.

133. Fradera X. and Babaoglu K. 2017. Overview of methods and strategies for conducting virtual small molecule screening. *Curr Protoc Chem Biol*, 9(3), 196–212.

134. Francica J.R., Varela-Rohena A., Medvec A., Plesa G., Riley J.L. et al. 2010. Steric shielding of surface epitopes and impaired immune recognition induced by the Ebola virus glycoprotein. *PLoS Pathog*, 6(9), e1001098.

135. Francis J.N., Bunce C.J., Horlock C., Watson J.M., Warrington S.J. et al. 2015. A novel peptide-based pan-influenza A vaccine: A double blind, randomised clinical trial of immunogenicity and safety. *Vaccine*, 33(2), 396-402.

136. Fritsche P.J., Helbling A. and Ballmer-Weber B.K. 2010. Vaccine hypersensitivity—update and overview. *Swiss Med Wkly*, 140(17–18), 238–246.

137. Fujihira H., Usami K., Matsuno K., Takeuchi H., Denda-Nagai K. et al. 2018. A Critical Domain of Ebolavirus Envelope Glycoprotein Determines Glycoform and Infectivity. *Sci Rep*, 8(1), 5495.

138. Furuta Y., Komeno T. and Nakamura T. 2017. Favipiravir (T-705), a broad spectrum inhibitor of viral RNA polymerase. *Proc Jpn Acad Ser B Phys Biol Sci*, 93(7), 449-463.

139. Garbutt M., Liebscher R., Wahl-Jensen V., Jones S., Moller P. et al. 2004. Properties of replication competent vesicular stomatitis virus vectors expressing glycoproteins of filoviruses and arenaviruses. *J Virol*, 78(10), 5458–5465.

140. Gehring G., Rohrmann K., Atenchong N., Mittler E., Becker S. et al. 2014. The clinically approved drugs amiodarone, dronedarone and verapamil inhibit filovirus cell entry. *J Antimicrob Chemother*, 69(8), 2123–2131.

141. Geisbert T. and Jahrling P. 1995. Differentiation of filoviruses by electron microscopy. *Virus Res*, 39(2-3), 129-150.

142. Geisbert T.W., Pushko P., Anderson K., Smith J., Davis K.J. et al. 2002. Evaluation in nonhuman primates of vaccines against Ebola virus. *Emerg Infect Dis* 8(5), 503–507.

143. Geisbert T.W., Daddario-Dicaprio K.M., Lewis M.G., Geisbert J.B., Grolla A. et al. 2008. Vesicular stomatitis virus-based Ebola vaccine is well-tolerated and protects immunocompromised nonhuman primates. *PLOS Pathog*, 4(11), e1000225.

144. Geisbert T.W., Geisbert J.B., Leung A., Daddario-DiCaprio K.M., Hensley L.E. et al. 2009. Single-injection vaccine protects nonhuman primates against infection with marburg virus and three species of ebola virus. *J Virol*, 83(14), 7296-7304.

145. Geisbert T.W., Lee A.C., Robbins M., Geisbert J.B., Honko A.N. et al. 2010. Postexposure protection of non-human primates against a lethal Ebola virus challenge with rna interference. A proof-of-concept study. *Lancet*, 375(9729), 1896–1905.

146. Geisbert T.W., Bailey M., Hensley L., Asiedu C., Geisbert J. et al. 2011. Recombinant adenovirus serotype 26 (Ad26) and Ad35 vaccine vectors bypass immunity

to Ad5 and protect nonhuman primates against Ebolavirus challenge. *J Virol,* 85(9):4222–4233.

147. Georges A.J., Leroy E.M., Renaut A.A., Benissan C.T., Nabias R.J. et al. 1999. Ebola hemorrhagic fever outbreaks in Gabon, 1994-1997: epidemiologic and health control issues. *J Infect Dis,* 179(1), S65-75.

148. Gerritz S.W., Cianci C., Kim S., Pearce B.C., Deminie C. et al. 2011. Inhibition of influenza virus replication via small molecules that induce the formation of higher-order nucleoprotein oligomers. *Proc Natl Acad Sci USA,* 108 (37), 15366–15371.

149. Goeijenbier M., van Kampen J.J., Reusken C.B., Koopmans M.P. and van Gorp E.C. 2014. Ebola virus disease: a review on epidemiology, symptoms, treatment and pathogenesis. *Neth J Med,* 72(9), 442-448.

150. Goncalves A.K., Giraldo P.C. and Machado P.R. 2015. Human Papillomavirus Vaccine-Induced Cytokine Messenger RNA Expression in Vaccinated Women. *Viral Immunol,* 28(6), 339-342.

151. Gonzalez-Galarza F.F., Takeshita L.Y., Santos E.J., Kempson F., Maia M.H. et al. 2015. Allele frequency net 2015 update: new features for HLA epitopes, KIR and disease and HLA adverse drug reaction associations. *Nucleic Acids Res,* 43(Database issue), D784-788.

152. Gowthaman U. and Agrewala J.N. 2008. In silico tools for predicting peptides binding to HLA-class II molecules: more confusion than conclusion. *J Proteome Res,* 7(1), 154–163.

153. Grant-Klein R.J., Altamura L.A. and Schmaljohn C.S. 2011. Progress in recombinant DNA-derived vaccines for Lassa virus and filoviruses. *Virus Res,* 162(1-2), 148–161.

154. Grant-Klein R.J., Altamura L.A., Badger C.V., Bounds C.E., Van Deusen N.M. et al. 2015. Codon-optimized filovirus DNA vaccines delivered by intramuscular electroporation protect cynomolgus macaques from lethal Ebola and Marburg virus challenges. *Hum Vaccin Immunother,* 11(8), 1991–2004.

155. Greenbaum J.A., Andersen P.H., Blythe M., Bui H.H., Cachau R.E. et al. 2007. Towards a consensus on datasets and evaluation metrics for developing B-cell epitope prediction tools. *J Mol Recognit,* 20(2), 75–82.

156. Grifoni A., Lo Presti A., Giovanetti M., Montesano C., Amicosante M. et al. 2016. Genetic diversity in Ebola virus: Phylogenetic and in silico structural studies of Ebola viral proteins. *Asian Pac J Trop Med,* 9(4), 337-343.

157. Groseth A., Str€oher U., Theriault S. and Feldmann H. 2002. Molecular characterization of an isolate from the 1989/90 epizootic of Ebola virus Reston among macaques imported into the United States. *Virus Res,* 87(2), 155-163.

158. Groseth A., Marzi A., Hoenen T., Herwig A., Gardner D. et al. 2012. The Ebola virus glycoprotein contributes to but is not sufficient for virulence in vivo. *PLoS Pathog,* 8(8), e1002847.

159. Group PIW, Multi-National PIIST, Davey Jr R.T., Dodd L., Proschan M.A. et al. 2016. A randomized, controlled trial of ZMapp for Ebola virus infection. *N Engl J Med,* 375(15), 1448–1456.

160. Guan P., Doytchinova I.A., Zygouri C. and Flower D.R. 2003. MHCPred: bringing a quantitative dimension to the online prediction of MHC binding. *Appl Bioinformatics,* 2(1), 63–66.

161. Guito J.C., Albariño C.G., Chakrabarti A.K. and Towner J.S. 2017. Novel activities by ebolavirus and marburgvirus interferon antagonists revealed using a standardized in vitro reporter system. *Virology*, 501, 147-165.

162. Gupta S., Kapoor P., Chaudhary K., Gautam A., Kumar R. et al. 2013. In silico approach for predicting toxicity of peptides and proteins. *PLoS One,* 8(9), e73957 .

163. Halfmann P., Kim J.H., Ebihara H., Noda T., Neumann G. et al. 2008. Generation of biologically contained Ebola viruses. *Proc Natl Acad Sci USA,* 105(4), 1129–1133.

164. Halfmann P., Ebihara H., Marzi A., Hatta Y., Watanabe S. et al. 2009. Replication-deficient Ebolavirus as a vaccine candidate. *J Virol,* 83(8), 3810–3815.

165. Halfmann P., Neumann G. and Kawaoka Y. 2011. The Ebolavirus VP24 protein blocks phosphorylation of p38 mitogen-activated protein kinase. *J Infect Dis,* 204(3), S953–S956.

166. Han Z., Boshra H., Sunyer J.O., Zwiers S.H., Paragas J. et al. 2003. Biochemical and Functional Characterization of the Ebola Virus VP24 Protein: Implications for a Role in Virus Assembly and Budding. *J Virol*, 77(3), 1793-1800.

167. Han Z., Lu J., Liu Y., Davis B., Lee M.S. et al. 2014. Small-molecule probes targeting the viral PPXY-host Nedd4 interface block egress of a broad range of RNA viruses. *J Virol*, 88(13), 7294–7306.

168. Hansen S.G., Vieville C., Whizin N., Coyne-Johnson L., Siess D.C. et al. 2009. Effector memory T cell responses are associated with protection of rhesus monkeys from mucosal simian immunodeficiency virus challenge. *Nat Med,* 15(3), 293–299.

169. Hansen S.G., Ford J.C., Lewis M.S., Ventura A.B., Hughes C.M. et al. 2011. Profound early control of highly pathogenic SIV by an effector memory T-cell vaccine. *Nature*, 473(7348), 523–527.

170. Hartlieb B., Muziol T., Weissenhorn W. and Becker S. 2007. Crystal structure of the C-terminal domain of Ebola virus VP30 reveals a role in transcription and nucleocapsid association. *Proc Natl Acad Sci USA*, 104(2), 624-629.

171. Hartman A.L., Bird B.H., Towner J.S., Antoniadou Z.A., Zaki S.R. et al. 2008. Inhibition of IRF-3 activation by VP35 is critical for the high level of virulence of Ebola virus. *J Virol*, 82(6), 2699–2704.

172. Hasan M.A., Hossain M. and Alam M.J. 2013. A computational assay to design an epitope-based Peptide vaccine against Saint Louis encephalitis virus. *Bioinform Biol Insights*, 7, 347–355.

173. Hasan M.A., Khan M.A., Datta A., Mazumder M.H. and Hossain M.U. 2015. A comprehensive immunoinformatics and target site study revealed the corner-stone toward Chikungunya virus treatment. *Mol Immunol*, 65(1), 189–204.

174. He Y., Rappuoli R., De Groot A.S. and Chen R. 2010. Emerging vaccine informatics. *J Biomed Biotechnol*, 2010, 218590.

175. Hensley L.E., Young H.A., Jahrling P.B. and Geisbert T.W. 2002. Proinflammatory response during Ebola virus infection of primate models: possible involvement of the tumor necrosis factor receptor superfamily. *Immunol Lett*, 80(3), 169–179.

176. Herbert A.S., Kuehne A.I., Barth J.F., Ortiz R.A., Nichols D.K. et al. 2013. Venezuelan equine encephalitis virus replicon particle vaccine protects nonhuman primates from intramuscular and aerosol challenge with ebolavirus. *J Virol* 87(9), 4952–6455.

177. Hevey M., Negley D., Geisbert J., Jahrling P. and Schmaljohn A. 1997. Antigenicity and vaccine potential of Marburg virus glycoprotein expressed by baculovirus recombinants. *Virology*, 239(1), 206–216.

178. Hevey M., Negley D., Pushko P., Smith J. and Schmaljohn A. 1998. Marburg virus vaccines based upon alphavirus replicons protect guinea pigs and nonhuman primates. *Virology*, 251(1), 28–37.

179. Heymann D.L., Weisfeld J.S., Webb P.A., Johnson K.M., Cairns T. et al. 1980. Ebola hemorrhagic fever: Tandala, Zaire, 1977-1978. *J Infect Dis*, 142(3), 372-376.

180. Hoenen, T.; Gottschalk, R.; Becker, S. Hofmann, F., editor. [Ebolavirus hemmorrhagic fever].; [Manual of infectious diseases]. 2007. p. VIII-6.11.1-15.Landsberg/Lech: ecomed Medizin.

181. Hoenen T., Groseth A. and Feldmann H. 2012. Current Ebola vaccines. *Expert Opin Biol Ther*, 12(7), 859.

182. Hoof I., Peters B., Sidney J., Pedersen L.E., Sette A. et al. 2009. NetMHCpan, a method for MHC class I binding prediction beyond humans. *Immunogenetics*, 61(1), 1–13.

183. Hopp T.P. and Woods K.R. 1981. Prediction of protein antigenic determinants from amino acid sequences. *Proc Natl Acad Sci USA*, 78(6), 3824–3828.

184. Hossain R., Yasmin T., Hosen M.I. and Nabi A.H.M.N. 2018. In silico identification of potential epitopes present in human adenovirus proteins for vaccine design and of putative drugs for treatment against viral infection. *J Immunol Methods*, 455, 55–70.

185. Hou Y., Guo Y., Wu C., Shen N., Jiang Y. et al. 2012. Prediction and identification of T cell epitopes in the H5N1 influenza virus nucleoprotein in chicken. *PLoS One*, 7(6), e39344.

186. Huang Y., Xu L., Sun Y. and Nabel G.J. 2002. The assembly of Ebola virus nucleocapsid requires virion-associated proteins 35 and 24 and posttranslational modification of nucleoprotein. *Mol Cell*, 10(2), 307-316.

187. Huang J., Gutteridge A., Honda W. and Kanehisa M. 2006. MIMOX: a web tool for phage display based epitope mapping. *BMC Bioinformatics*, 7(1), 451.

188. Huang Y.X., Bao Y.L., Guo S.Y., Wang Y., Zhou C.G. et al. 2008. Pep-3D-Search: a method for B-cell epitope prediction based on mimotope analysis. *BMC Bioinformatics*, 9, 538.

189. Huang Y., Zhang L., Jolliffe D., Sanchez B., Stjernholm G. et al. 2018. Post vaccination C-Reactive protein and C5/gp41732–744 antibody level fold-changes over baseline are independent predictors of therapeutic HIV vaccine effect in a phase 2 clinical study of Vacc-4x. *AIDS Res Hum Retroviruses*, 34(3), 307–313.

190. Hunt C.L., Kolokoltsov A.A., Davey R.A. and Maury W. 2011. The Tyro3 receptor kinase Axl enhances macropinocytosis of Zaire ebolavirus. *J Virol*, 85(1), 334–347.

191. Hutchinson K.L. and Rollin P.E. 2007. Cytokine and Chemokine Expression in Humans Infected with Sudan Ebola Virus. *J Infect Dis*, 196(2), S357-S363.

192. Huttner A., Dayer J.A., Yerly S., Combescure C., Auderset F. et al. 2015. The effect of dose on the safety and immunogenicity of the VSV Ebola candidate vaccine: a

randomised double-blind, placebo-controlled phase 1/2 trial. *Lancet Infect Dis*, 15(10), 1156–1166.

193. ICAHN, 2019. Basler Laboratory. http://labs.icahn.mssm.edu/basler-lab/research-interests/

194. Ichihashi T., Yoshida R., Sugimoto C., Takada A. and Kajino K. 2011. Cross-protective peptide vaccine against influenza A viruses developed in HLA-A*2402 human immunity model. *PLoS One*, 6(9), e24626.

195. Ikegami T., Calaor A., Miranda M., Niikura M., Saijo M. et al. 2001. Genome structure of Ebola virus subtype Reston: differences among Ebola subtypes. *Arch Virol*, 146(10), 2021-2027.

196. Ilesanmi O.S. and Alele F.O. 2014. The effect of Ebola Virus Disease outbreak on hand washing among secondary school students in Ondo State Nigeria, October, 2014. *Pan Afr Med J*, 22(1), 24.

197. Ilinykh P.A., Lubaki N.M., Widen S.G., Renn L.A., Theisen T.C. et al. 2015. Different Temporal Effects of Ebola Virus VP35 and VP24 Proteins on Global Gene Expression in Human Dendritic Cells. *J Virol*, 89(15), 7567-7583.

198. Ito H., Watanabe S., Takada A. and Kawaoka Y. 2001. Ebola virus glycoprotein: proteolytic processing, acylation, cell tropism, and detection of neutralizing antibodies. *J Virol*, 75(3), 1576–1580.

199. Ivashkiv L.B. and Donlin L.T. 2014. Regulation of type I interferon responses. *Nat Rev Immunol*, 14(1), 36–49.

200. Iversen P.L., Warren T.K., Wells J.B., Garza N.L., Mourich D.V. et al. 2012. Discovery and early development of avi-7537 and avi-7288 for the treatment of Ebola virus and Marburg virus infections. *Viruses*, 4(11), 2806–2830.

201. Jackwood M.W., Hickle L., Kapil S. and Silva R. 2008. Vaccine development using recombinant DNA technology. *Council Agric Sci Technol*, 38, 1–11.

202. Jacob L. and Vert J.P. 2008. Efficient peptide-MHC-I binding prediction for alleles with few known binders. *Bioinformatics*, 24(3), 358-366.

203. Jahrling P.B., Geisbert T.W., Dalgard D.W., Johnson E.D., Ksiazek T.G. et al. 1990. Preliminary report: isolation of Ebola virus from monkeys imported to USA. *Lancet*, 335(8688), 502–505.

204. Jain A., Tripathi P., Shrotriya A., Chaudhary R. and Singh A. 2015. In silico analysis and modeling of putative T cell epitopes for vaccine design of Toscana virus. *3 Biotech*, 5(4), 497-503.

205. Jarzab A., Skowicki M. and Witkowska D. 2013. Subunit vaccines- antigens, carriers, conjugation methods and the role of adjuvants. *Postepy Hig Med Dosw*, 67, 1128–1143.

206. Jayaratne H.E., Wijeratne D., Fernando S., Kamaladasa A., Gomes L. et al. 2018. Regulatory T-cells in acute dengue viral infection. *Immunology*, 154(1), 89–97.

207. Jeffers S.A., Sanders D.A. and Sanchez A. 2002. Covalent modifications of the Ebola virus glycoprotein. *J Virol*, 76(24), 12463–12472.

208. Jerebtsova M. and Nekhai S. 2015. Therapeutics for postexposure treatment of Ebola virus infection. *Future Virol*, 10(3), 221-232.

209. Jespersen M.C., Peters B., Nielsen M. and Marcatili P. 2017. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Res*, 45(W1), W24-W29.

210. Johansen L.M., Brannan J.M., Delos S.E., Shoemaker C.J., Stossel A. et al. 2013. FDA-approved selective estrogen receptor modulators inhibit Ebola virus infection. *Sci Transl Med*, 5(190), 190ra79.

211. Johnson R.F., Kurup D., Hagen K.R., Fisher C., Keshwara R. et al. 2016. An inactivated rabies virus-based Ebola vaccine, FILORAB1, adjuvanted with glucopyranosyl lipid A in stable emulsion confers complete protection in nonhuman primate challenge models. *J Infect Dis*, 214(3), S342–S354.

212. Jones S.M., Feldmann H., Stroher U., Geisbert J.B., Fernando L. et al. 2005. Live attenuated recombinant vaccine protects nonhuman primates against Ebola and Marburg viruses. *Nat Med*, 11(7), 786–790.

213. Jun S., Leuze M.R., Nookaew I., Uberbacher E.C., Land M. et al. 2015. Ebolavirus comparative genomics. *FEMS Microbiol Rev*, 39(5), 764-778.

214. Kaiko G.E., Horvat J.C., Beagley K.W. and Hansbro P.M. 2008. Immunological decision-making: How does the immune system decide to mount a helper T-cell response? *Immunology*, 123(3), 326-338.

215. Kajihara M., Nakayama E., Marzi A., Igarashi M., Feldmann H. et al. 2013. Novel mutations in Marburg virus glycoprotein associated with viral evasion from antibody mediated immune pressure. *J Gen Virol*, 94(Pt 4), 876–883.

216. Kang S., Brown H.M. and Hwang S. 2018. Direct Antiviral Mechanisms of Interferon-Gamma. *Immune Netw*, 18(5), e33.

217. Kao R.Y., Yang D., Lau L.S., Tsui W.H., Hu L. et al. 2010. Identification of influenza A nucleoprotein as an antiviral target. *Nat Biotechnol*, 28(6), 600–605.

218. Karlsson I., Brandt L., Vinner L., Kromann I., Andreasen L.V. et al. 2013. Adjuvanted HLA-supertype restricted subdominant peptides induce new T-cell immunity during untreated HIV-1-infection. *Clin Immunol*, 146 (2), 120-130.

219. Karp P.D., Berger B., Kovats D., Lengauer T., Linial M. et al. 2015. ISCB Ebola Award for Important Future Research on the Computational Biology of Ebola Virus. *F1000Res*, 4, 12.

220. Karplus P.A. and Schulz G.E. 1985. Prediction of chain flexibility in proteins - a tool for the selection of peptide antigens. *Naturwissenschaften*, 72(4), 212–213.

221. Karron R.A., Belshe R.B., Wright P.F., Thumar B., Burns B. et al. 2003. A live human parainfluenza type 3 virus vaccine is attenuated and immunogenic in young infants. *Pediatr Infect Dis J*, 22(5), 394–405.

222. Keller M.A. and Stiehm E.R. 2000. Passive immunity in prevention and treatment of infectious diseases. *Clin Microbiol Rev*, 13(4), 602–614.

223. Khaiboullina S.F., Levis S., Morzunov S.P., Martynova E.V., Anokhin V.A. et al. 2017. Serum cytokine profiles differentiating hemorrhagic fever with renal syndrome and hantavirus pulmonary syndrome. *Front Immunol*, 8, 567.

224. Khan A.S., Tshioko F.K., Heymann D.L., Le Guenno B., Nabeth P. et al. 1999. The Reemergence of Ebola Hemorrhagic Fever, Democratic Republic of the Congo, 1995. Commission de Lutte contre les Epidémies à Kikwit. *J Infect Dis*, 179(1), S76-S86.

225. Khan M.A., Hossain M.U., Rakib-Uz-Zaman S.M. and Morshed M.N. 2015. Epitope-based peptide vaccine design and target site depiction against Ebola viruses: an immunoinformatics study. *Scand J Immunol*, 82(1), 25-34.

226. King A.M., Adams M.J., Lefkowitz E.J. and Carstens E.B. 2012. Virus Taxonomy: Classification and Nomenclature of Viruses: Ninth Report of the International Committee on Taxonomy of Viruses Vol. 9.

227. Kirchdoerfer R.N., Abelson D.M., Li S., Wood M.R. and Saphire E.O. 2015. Assembly of the Ebola Virus Nucleoprotein from a Chaperoned VP35 Complex. *Cell reports*, 12(1), 140-149.

228. Kirchdoerfer R.N., Moyer C.L., Abelson D.M. and Saphire E.O. 2016. The Ebola Virus VP30-NP Interaction Is a Regulator of Viral RNA Synthesis. *PLoS Pathog*, 12(10), e1005937.

229. Konde M.K., Baker D.P., Traore F.A., Sow M.S., Camara A. et al. 2017. Interferon β-1a for the treatment of Ebola virus disease: a historically controlled, single-arm proof-of-concept trial. *PLoS One*, 12(2), e0169255.

230. Korber B., LaBute M. and Yusim K. 2006. Immunoinformatics comes of age. *PLoS Comput Biol*, 2(6), e71.

231. Kovjazin R., Shitrit D., Preiss R., Haim I., Triezer L. et al. 2013. Characterization of novel multiantigenic vaccine candidates with pan-HLA coverage against Mycobacterium tuberculosis. *Clin Vaccine Immunol,* 20(3), 328–340 .

232. Kringelum J.V., Lundegaard C., Lund O. and Nielsen M. 2012. Reliable B cell epitope predictions: impacts of method development and improved benchmarking. *PLoS Comput Biol*, 8(12), e1002829.

233. Kudoyarova-Zubavichene N.M., Sergeyev N.N., Chepurnov A.A. and Netesov S.V. 1999. Preparation and use of hyperimmune serum for prophylaxis and therapy of Ebola virus infections. *J Infect Dis*, 179(1), S218–S223.

234. Kuhn J.H., Radoshitzky S.R., Guth A.C., Warfield K.L., Li W. et al. 2006. Conserved receptor-binding domains of Lake Victoria Marburgvirus and Zaire Ebolavirus bind a common receptor. *J Biol Chem,* 281(23), 15951–15958.

235. Kuhn J.H. 2017. Guide to the correct use of filoviral nomenclature. *Curr Top Microbiol Immunol,* 411, 447-460.

236. Kumar A., Yadav I.S., Hussain S., Das B.C. and Bharadwaj M. 2015. Identification of immunotherapeutic epitope of E5 protein of human papillomavirus-16: An in silico approach. *Biologicals*, 43(5), 344–348.

237. Kwong P.D., Wyatt R., Robinson J., Sweet R.W., Sodroski J. et al. 1998. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature*, 393(6686), 648–659.

238. de La Vega M.A., Stein D. and Kobinger G.P. 2015. Ebolavirus evolution: Past and present. *PLoS Pathog,* 11(11), e1005221.

239. Larsen J.E., Lund O. and Nielsen M. 2006. Improved method for predicting linear B-cell epitopes. *Immunome Res*, 2, 2.

240. Larsen M.V., Lundegaard C., Lamberth K., Buus S., Lund O. et al. 2007. Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. *BMC Bioinformatics*, 8, 424.

241. Le Guenno B., Formenty P., Wyers M., Gounon P., Walker F. et al. 1995. Isolation and partial characterisation of a new strain of Ebola virus. *Lancet*, 345(8960), 1271-1274.

242. Ledgerwood J.E., DeZure A.D., Stanley D.A., Coates E.E., Novik L. et al. 2017. Chimpanzee Adenovirus Vector Ebola Vaccine. *N Engl J Med*, 376(10), 928–938.

243. Lee J.E., Fusco M.L., Hessell A.J., Oswald W.B., Burton D.R. et al. 2008. Structure of the Ebola virus glycoprotein bound to an antibody from a human survivor. *Nature*, 454(7201), 177–182 .

244. Lee J.E. and Saphire E.O. 2009. Ebolavirus glycoprotein structure and mechanism of entry. *Future Virol*, 4(6), 621–635.

245. Leroy E.M., Baize S., Volchkov V.E., Fisher-Hoch S.P., Georges-Courbot M.C. et al. 2000. Human asymptomatic Ebola infection and strong inflammatory response. *Lancet*, 355(9222), 2210–2215.

246. Leroy E.M., Kumulungui B., Pourrut X., Rouquet P., Hassanin A. et al. 2005. Fruit bats as reservoirs of Ebola virus. *Nature*, 438(7068), 575–576.

247. Leroy E.M., Epelboin A., Mondonge V., Pourrut X., Gonzalez J.P. et al. 2009. Human Ebola outbreak resulting from direct exposure to fruit bats in Luebo, Democratic Republic of Congo, 2007. *Vector Borne Zoonot Dis*, 9(6), 723–728.

248. Leung D.W., Ginder N.D., Fulton D.B., Nix J., Basler C.F. et al. 2009. Structure of the Ebola VP35 interferon inhibitory domain. *Proc Natl Acad Sci USA*, 106(2), 411-416.

249. Leung D.W., Daisy W., Borek D., Luthra P., Binning J.M. et al. 2015. An Intrinsically Disordered Peptide from Ebola Virus VP35 Controls Viral RNA Synthesis by Modulating Nucleoprotein-RNA Interactions. *Cell Rep*, 11(3), 376–389.

250. Li W., Joshi M.D., Singhania S., Ramsey K.H. and Murtl A.K. 2014. Peptide vaccine: progress and challenges. *Vaccines (Basel)*, 2(3), 515–536.

251. Liao S.J., Zhang W.N., Hu X.J., Jiang X.F., Wang C.Y. et al. 2012. Preparation of HPV18 E7 peptide plus CpG vaccine and its immunologic effects in vitro. *Zhonghua Yi Xue Za Zhi*, 92(23), 1641-1645.

252. Licata J.M., Johnson R.F., Han Z. and Harty R.N. 2004. Contribution of Ebola virus glycoprotein, nucleoprotein, and VP24 to budding of VP40 virus-like particles. *J Virol*, 78(14), 7344–7351.

253. Lin H.H., Zhang G.L., Tongchusak S., Reinherz E.L. and Brusic V. 2008. Evaluation of MHC-II peptide binding prediction servers: applications for vaccine research. *BMC Bioinformatics*, 9(12), S12-S22.

254. Ling A. and Whitton J.L. 1997. A multivalent minigene vaccine, containing B cell, CTL, and Th epitopes from several microbes, induces appropriate responses in vivo and confers protection against more than one pathogen. *J Virol*, 71(3), 2292–2302.

255. Lingemann M., Liu X., Surman S., Liang B., Herbert R. et al. 2017. Attenuated human parainfluenza virus type 1 expressing Ebola virus glycoprotein GP administered intranasally is immunogenic in African green monkeys. *J Virol,* 91(10), e02469-16.

256. Lohia N. and Baranwal M. 2015. Identification of conserved peptides comprising multiple T cell epitopes of Matrix 1 protein in H1N1 influenza virus. *Viral Immunol,* 28 (10), 570–579.

257. Lohia N. and Baranwal M. 2017. Immune response of highly conserved influenza A virus matrix 1 peptides: Matrix 1 peptides for influenza vaccine. *Microbiol Immunol,* 61(6), 225–231.

258. Lohia N. and Baranwal M. 2018. Highly conserved hemagglutinin peptides of H1N1 influenza virus elicit immune response. *3 Biotech,* 8(12), 492.

259. Lund O., Nielsen M., Kesmir C., Petersen A.G., Lundegaard C. et al. 2004. Definition of supertypes for HLA molecules using clustering of specificity matrices. *Immunogenetics,* 55(12), 797–810 .

260. Lundegaard C., Lund O. and Nielsen M. 2008. Accurate approximation method for prediction of class I MHC affinities for peptides of length 8, 10 and 11 using prediction tools trained on 9mers. *Bioinformatics,* 24(11), 1397–1398.

261. Lupton H.W., Lambert R.D., Bumgardner D.L., Moe J.B. and Eddy G.A. 1980. Inactivated vaccine for Ebola virus efficacious in guineapig model. *Lancet,* 2(8207), 1294–1295.

262. MacNeil A., Farnon E.C., Morgan O.W., Gould P., Boehmer T.K. et al. 2011. Filovirus Outbreak Detection and Surveillance: Lessons from Bundibugyo. *J Infect Dis,* 204(3), S761-S767.

263. Madara J.J., Han Z., Ruthel G., Freedman B.D. and Harty R.N. 2015. The multifunctional Ebola virus VP40 matrix protein is a promising therapeutic target. *Future Virol,* 10(5), 537–546.

264. Maganga G.D., Kapetshi J., Berthet N., Ilunga B.K., Kabange F. et al. 2014. Ebola virus disease in the Democratic Republic of Congo. *New England Journal of Medicine,* 371, 2083-2091.

265. Mahaddalkar T., Suri C., Naik P.K. and Lopus M. 2015. Biochemical characterization and molecular dynamic simulation of beta-sitosterol as a tubulin-binding anticancer agent. *Eur J Pharmacol,* 760, 154-162.

266. Malaria Site, 2019. Malaria Vaccines. https://www.malariasite.com/malaria-vaccines/

267. Malashkevich V.N., Schneider B.J., McNally M.L., Milhollen M.A., Pang J.X. et al. 1999. Core structure of the envelope glycoprotein GP2 from Ebola virus at 1.9-A resolution. *Proc Natl Acad Sci USA*, 96(6), 2662–2667.

268. Manicassamy B., Wang J., Jiang H. and Rong L. 2005. Comprehensive analysis of Ebola virus GP1 in viral entry. *J Virol*, 79(8), 4793–4805.

269. Marcinkiewicz J., Bryniarski K. and Nazimek K. 2014. Ebola haemorrhagic fever virus: pathogenesis, immune responses, potential prevention. *Folia Med Cracov*, 54(3), 39-48.

270. Martin J.E., Sullivan N.J., Enama M.E., Gordon I.J., Roederer M. et al. 2006. A DNA vaccine for Ebola virus is safe and immunogenic in a phase I clinical trial. *Clin Vaccine Immunol*, 13(11), 1267–1277.

271. Martinez M.J., Biedenkopf N., Volchkova V., Hartlieb B., Alazard-Dany N. et al. 2008. Role of Ebola virus VP30 in transcription reinitiation *J Virol*, 82(24), 12569–12573.

272. Martins K.A.O., Cooper C.L., Stronsky S.M., Norris S.L.W., Kwilas S.A. et al. 2015. Adjuvant-enhanced CD4 T cell responses are critical to durable vaccine immunity. *EBioMedicine*, 3, 67-78.

273. Martins K.A., Jahrling P.B., Bavari S. and Kuhn J.H. 2016. Ebola virus disease candidate vaccines under evaluation in clinical trials. *Expert Rev Vaccines*, 15(9), 1101-1112.

274. Marzi A., Feldmann H., Geisbert T.W. and Falzarano D. 2011. Vesicular stomatitis virus-based vaccines for prophylaxis and treatment of filovirus infections. *J Bioterror Biodef*, S1(4), 004.

275. Marzi A., Engelmann F., Feldmann F., Haberthur K., Shupert W.L. et al. 2013. Antibodies are necessary for rVSV/ZEBOV-GP-mediated protection against lethal Ebola virus challenge in nonhuman primates. *PNAS*, 110(5), 1893–1898.

276. Marzi A. and Feldmann H. 2014. Ebola virus vaccines: an overview of current approaches. *Expert Rev Vaccines*, 13(4), 521–31.

277. Marzi A., Halfmann P., Hill-Batorski L., Feldmann F., Shupert W.L. et al. 2015. Vaccines: An Ebola whole-virus vaccine is protective in nonhuman primates. *Science*, 348(6233), 439–442.

278. Marzi A., Robertson S.J., Haddock E., Feldmann F., Hanley P.W. et al. 2015. VSV-EBOV rapidly protects macaques against infection with the 2014/15 Ebola virus outbreak strain. *Science*, 349(6249), 739–742.

279. Marzi A., Murphy A.A., Feldmann F., Parkins C.J., Haddock E. et al. 2016. Cytomegalovirus-based vaccine expressing Ebola virus glycoprotein protects nonhuman primates from Ebola virus infection. *Sci Rep,* 6, 21674.

280. Mast T.C., Kierstead L., Gupta S.B., Nikas A.A., Kallas E.G. et al. 2010. International epidemiology of human pre-existing adenovirus (Ad) type-5, type-6, type-26 and type-36 neutralizing antibodies: correlates of high Ad5 titers and implications for potential HIV vaccine trials. *Vaccine,* 28(4), 950–957.

281. Matassov D., Marzi A., Latham T., Xu R., Ota-Setlik A. et al. 2015. Vaccination with a highly attenuated recombinant vesicular stomatitis virus vector protects against challenge with a lethal dose of Ebola virus. *J Infect Dis,* 212(2), S443–S451.

282. Mateo M., Reid S.P., Leung L.W., Basler C.F. and Volchkov V.E. 2010. Ebolavirus VP24 binding to karyopherins is required for inhibition of interferon signaling. *J Virol,* 84(2), 1169–1175.

283. Mathebula L., Ndwandwe D.E., Pienaar E. and Wiysonge C.S. 2019. Effects of vaccines in protecting against Ebola virus disease: protocol for a systematic review. *BMJ Open,* 9(7), e029617.

284. Matzinger P. and Kamala T. 2011. Tissue-based class control: the other side of tolerance. *Nat Rev Immunol,* 11(3), 221-230.

285. McCune J.M., Rabin L.B., Feinberg M.B., Lieberman M., Kosek J.C. et al. 1988. Endoproteolytic cleavage of gp160 is required for the activation of human immunodeficiency virus. *Cell,* 53(1), 55–67.

286. McElroy A.K., Akondy R.S., Davis C.W., Ellebedy A.H., Mehta A.K. et al. 2015. Human Ebola virus infection results in substantial immune activation. *Proc Natl Acad Sci USA,* 112(15), 4719-4724.

287. McMahon S.A., Ho L.S., Brown H., Miller L., Ansumana R. et al. 2016. Healthcare providers on the frontlines: a qualitative investigation of the social and emotional impact of delivering health services during Sierra Leone's Ebola epidemic. *Health Policy Plan,* 31(9), 1232–1239.

288. Mehedi M., Falzarano D., Seebach J., Hu X., Carpenter M.S. et al. 2011. A new Ebola virus nonstructural glycoprotein expressed through RNA editing. *J Virol,* 85(11), 5406-5414.

289. M€uhlberger E., Sanchez A., Randolf A., Will C., Kiley M.P. et al. 1992. The nucleotide sequence of the L gene of Marburg virus, a filovirus: homologies with paramyxoviruses and rhabdoviruses. *Virology,* 187(2), 534-547.

290. Meyboom R.H., Fucik H. and Edwards I.R. 1995. Thrombocytopenia reported in association with hepatitis B and A vaccines. *Lancet,* 345(8965), 1638.

291. Meyer M., Garron T., Lubaki N.M., Mire C.E., Fenton K.A. et al. 2015. Aerosolized Ebola vaccine protects primates and elicits lung-resident T cell responses. *J Clin Invest*, 125(8), 3241–3255.

292. Michelow I.C., Dong M., Mungall B.A., Yantosca L.M., Lear C. et al. 2010. A novel l-ficolin/Mannose-binding lectin chimeric molecule with enhanced activity against Ebola virus. *J Biol Chem*, 285(32), 24729–24739.

293. Michelow I.C., Lear C., Scully C., Prugar L.I., Longley C.B. et al. 2011. High-dose Mannose-binding lectin therapy for Ebola virus infection. *J Infect Dis*, 203(2), 175–179.

294. Miller Y.I., Choi S.H., Wiesner P., Fang L., Harkewicz R. et al. 2011. Oxidation-specific epitopes are danger-associated molecular patterns recognized by pattern recognition receptors of innate immunity. *Circ Res*, 108(2), 235-248.

295. Miller M.E., Adhikary S., Kolokoltsov A.A. and Davey R.A. 2012. Ebolavirus Requires Acid Sphingomyelinase Activity and Plasma Membrane Sphingomyelin for Infection. *J Virol*, 86(14), 7473–7483.

296. Miotto O., Heiny A., Tan T.W., August J.T. and Brusic V. 2008. Identification of human-to-human transmissibility factors in PB2 proteins of influenza A by large scale mutual information analysis. *BMC Bioinformatics*, 9(1), S18.

297. Miraglia C.M. 2018. Marburgviruses: An Update. *Lab Med*, 50(1), 16-28.

298. Miranda M.E. and Miranda N.L. 2011. Reston ebolavirus in humans and animals in the Philippines: a review. *J Infect Dis*, 204(3), S757–S760.

299. Mire C.E., Geisbert J.B., Marzi A., Agans K.N., Feldmann H. et al. 2013. Vesicular stomatitis virus-based vaccines protect nonhuman primates against bundibugyo Ebolavirus. *PLoS Negl Trop Dis*, 7(12), e2600.

300. Mire C.E., Matassov D., Geisbert J.B., Latham T.E., Agans K.N. et al. 2015. Single-dose attenuated Vesiculovax vaccines protect primates against Ebola Makona virus. *Nature*, 520(7549), 688–691.

301. Mire C.E., Geisbert T.W., Feldmann H. and Marzi A. 2016. Ebola virus vaccines—reality or fiction? *Expert Rev Vaccines*, 15(11), 1421–1430.

302. Mittendorf E.A., Holmes J.P., Ponniah S. and Peoples G.E. 2008. The E75 HER2/neu peptide vaccine. *Cancer Immunol Immunother*, 57(10), 1511-1521.

303. Modrof J., Moritz C., Kolesnikova L., Konakova T., Hartlieb B. et al. 2001. Phosphorylation of Marburg virus VP30 at serines 40 and 42 is critical for its interaction with NP inclusions. *Virology*, 287(1), 171-182.

304. Modrof J., Muhlberger E., Klenk H.D. and Becker S. 2002. Phosphorylation of VP30 impairs Ebola virus transcription. *J Biol Chem*, 277(36), 33099–33104.

305. Modrof J., Becker S. and Muhlberger E. 2003. Ebola virus transcription activator VP30 is a zinc-binding protein. *J Virol*, 77(5), 3334-3338.

306. Mohan G.S., Li W., Ye L., Compans R.W. and Yang C. 2012. Antigenic Subversion: A novel mechanism of host immune evasion by Ebola virus. *PLoS Pathogens*, 8(12), e1003065.

307. Monterrubio-López G.P., Y-Merchand JA G. and Ribas-Aparicio R.M. 2015. Identification of novel potential vaccine candidates against tuberculosis based on reverse vaccinology. *BioMed Res Int*, 2015, 483150 .

308. Mosmann T. 1983. Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. *J Immunol Methods*, 65(1-2), 55-63.

309. Moutaftsi M., Peters B., Pasquetto V., Tscharke D.C., Sidney J. et al. 2006. A consensus epitope prediction approach identifies the breadth of murine T(CD8+)-cell responses to vaccinia virus. *Nat Biotechnol*, 24(7), 817-819.

310. Muhlberger E., Weik M., Volchkov V.E., Klenk H.D. and Becker S. 1999. Comparison of the transcription and replication strategies of marburg virus and Ebola virus by using artificial replication systems. *J Virol*, 73(3), 2333–2342.

311. Murray M.J. 2015. Ebola Virus Disease: A Review of Its Past and Present. *Anesth Analg*, 121(3), 798-809.

312. Mustafa A.S. 2011. Comparative evaluation of MPT83 (Rv2873) for T helper-1 cell reactivity and identification of HLA-promiscuous peptides in Mycobacterium bovis BCG-vaccinated healthy subjects. *Clin Vaccine Immunol*, 18(10), 1752-1759.

313. Naik P.K., Lopus M., Aneja R., Vangapandu S.N. and Joshi H.C. 2012. In silico inspired design and synthesis of a novel tubulin-binding anti-cancer drug: folate conjugated noscapine (Targetin). *J Comput Aided Mol Des*, 26(2), 233–247.

314. Nanbo A., Imai M., Watanabe S., Noda T., Takahashi K. et al. 2010. Ebolavirus is internalized into host cells via macropinocytosis in a viral glycoprotein-dependent manner. *PLoS Pathog*, 6(9), e1001121.

315. Nardin E.H., Oliveira G.A., Calvo-Calle J.M., Castro Z.R., Nussenzweig R.S. et al. 2000. Synthetic malaria peptide vaccine elicits high levels of antibodies in vaccinees of defined HLA genotypes. *J Infect Dis*, 182(5), 1486-1496.

316. Neefjes J., Jongsma M.L., Paul P. and Bakke O. 2011. Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nat Rev Immunol,* 11(12), 823-836.

317. Negi S.S. and Braun W. 2009. Automated detection of conformational epitopes using phage display peptide sequences. *Bioinform Biol Insights*, 3, 71–81.

318. Neil S.J., Zang T. and Bieniasz P.D. 2008. Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature*, 451(7177), 425–430.

319. Neumann G., Geisbert T.W., Ebihara H., Geisbert J.B., Daddario-DiCaprio K.M. et al. 2007. Proteolytic processing of the Ebola virus glycoprotein is not critical for Ebola virus replication in nonhuman primates. *J Virol*, 81(6), 2995–2998.

320. Nevagi R.J., Toth I. and Skwarczynski M. 2018. Peptide-based vaccines. *Peptide Applications in Biomedicine, Biotechnology and Bioengineering*, 327-358.

321. Nielsen M., Lundegaard C., Worning P., Lauemøller S.L., Lamberth K. et al. 2003. Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. *Protein Sci*, 12(5), 1007–1017.

322. Nielsen M., Lundegaard C. and Lund O. 2007. Prediction of HLA class II binding affinity using SMMalign, a novel stabilization matrix alignment method. *BMC Bioinformatics*, 8, 238.

323. Nielsen M. and Lund O. 2009. NN-align. An artificial neural network-based alignment algorithm for MHC class II peptide binding prediction. *BMC Bioinformatics*, 10, 296.

324. Niikura M., Ikegami T., Saijo M., Kurane I., Miranda M.E. et al. 2001. Detection of Ebola viral antigen by enzyme-linked immunosorbent assay using a novel monoclonal antibody to nucleoprotein. *J Clin Microbiol*, 39(9), 3267–3271.

325. Nkoghe D., Kone M.L., Yada A. and Leroy E. 2011. A limited outbreak of Ebola haemorrhagic fever in Etoumbi, Republic of Congo, 2005. *Trans R Soc Trop Med Hyg*, 105(8), 466–472.

326. Noda T., Hagiwara K., Sagara H. and Kawaoka Y. 2010. Characterization of the Ebola virus nucleoprotein-RNA complex. *J Gen Virol*, 91(6), 1478–1483.

327. Oany A.R., Emran A.A. and Jyoti T.P. 2008. Design of an epitope-based peptide vaccine against spike protein of human corona virus: an in silico approach. *Drug Des Devel Ther*, 8, 1139–1149.

328. Oany A.R., Ahmad S.A.I., Hossain M.U. and Jyoti T.P. 2015. Identification of highly conserved regions in L-segment of Crimean–Congo hemorrhagic fever virus and immunoinformatic prediction about potential novel vaccine. *Adv Appl Bioinformatics Chem*, 8, 1–10.

329. Oestereich L., Ludtke A., Wurr S., Rieger T., Munoz-Fontela C. et al. 2014. Successful treatment of advanced Ebola virus infection with t-705 (favipiravir) in a small animal model. *Antiviral Res*, 105, 17–21.

330. Ohkuri T., Wakita D., Chamoto K., Togashi Y., Kitamura H. et al. 2009. Identification of novel helper epitopes of MAGE-A4 tumour antigen: useful tool for the propagation of Th1 cells. *Br J Cancer*, 100(7), 1135-1143.

331. Ohno S., Okuyama R., Aruga A., Sugiyama H. and Yamamoto M. 2012. Phase I trial of Wilms' tumor 1 (WT1) peptide vaccine with GM-CSF or CpG in patients with solid malignancy. *Anticancer Res*, 32(6), 2263-2269.

332. Okoror L., Kamara A., Kargbo B., Bangura J., Lebby M. et al. 2018. Transplacental transmission: A rare case of Ebola virus transmission. *Infect Dis Rep*, 10(3), 7725.

333. Okware S.I., Omaswa F.G., Zaramba S., Opio A., Lutwama J.J. et al. 2002. An outbreak of Ebola in Uganda. *Trop Med Int Health*, 7(12), 1068-1075.

334. Oliveira F.M., Coelho I.E., Lopes M.D., Taranto A.G., Junior M.C. et al. 2016. The use of reverse vaccinology and molecular modeling associated with cell proliferation stimulation approach to select promiscuous epitopes from Schistosoma mansoni. *Appl Biochem Biotechnol*, 179(6), 1023–1040.

335. Olson J.K., Croxford J.L., Calenoff M.A., Dal Canto M.C. and Miller S.D. 2001. A virus-induced molecular mimicry model of multiple sclerosis. *J Clin Invest*, 108(2), 311–8.

336. O'Shea M.K., Clay K.A., Craig D.G., Matthews S.W., Kao R.L. et al. 2015. Diagnosis of febrile illnesses other than Ebola virus disease at an Ebola treatment unit in Sierra Leone. *Clin Infect Dis*, 61(5), 795–798.

337. Oswald W.B., Geisbert T.W., Davis K.J., Geisbert J.B., Sullivan N.J. et al. 2007. Neutralizing antibody fails to impact the course of Ebola virus infection in monkeys. *PLoS Pathog*, 3(1), e9.

338. Parker J.M., Guo D. and Hodges R.S. 1986. New hydrophilicity scale derived from high performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites. *Biochemistry*, 25(19), 5425–5432.

339. Parker K.C., Bednarek M.A. and Coligan J.E. 1994. Scheme for ranking potential HLA-A2 binding peptides based on independent binding of individual peptide side-chains. *J Immunol*, 152(1), 163-175.

340. Pascal M., Konstantinou G.N., Masilamani M., Lieberman J. and Sampson H.A. 2013. In silico prediction of Ara h 2 T cell epitopes in peanut-allergic children. *Clin Exp Allergy*, 43(1), 116-127.

341. Patja A., Mäkinen-Kiljunen S., Davidkin I., Paunio M. and Peltola H. 2001. Allergic reactions to measles-mumps-rubella vaccination. *Pediatrics*, 107(2), E27.

342. Patronov A., Dimitrov I., Flower D.R. and Doytchinova I. 2011. Peptide binding prediction for the human class II MHC allele HLA-DP2: a molecular docking approach. *BMC Structural Biology*, 11(1), 1-10.

343. Patronov, A. and Doytchinova I. 2013. T-cell epitope vaccine design by immunoinformatics. *Open Biol*, 3(1), 120139.

344. Perdomo-Celis F., Salvato M.S., Medina-Moreno S. and Zapata J.C. 2019. T-Cell Response to Viral Hemorrhagic Fevers. *Vaccines (Basel)*, 7(1), E11.

345. Perez-Martinez A.P., Ong E., Zhang L., Marrs C.F., He Y. et al. 2017. Conservation in gene encoding Mycobacterium tuberculosis antigen Rv2660 and a high predicted population coverage of H56 multistage vaccine in South Africa. *Infect Genet Evol*, 55, 244-250.

346. Peterson A.T., Carroll D.S., Mills J.N. and Johnson K.M. 2004. Potential mammalian filovirus reservoirs. *Emerg Infect Dis*, 10(12), 2073–2081.

347. Petesch C. 2018. Alarm as Red Cross workers attacked in Congo. Medical Press. Medical Press.

348. Pharmaintelligence, 2019. Missing In Action: China's Best Shot In Ebola Vaccine Race. https://pharmaintelligence.informa.com/resources/product-content/missing-in-action-chinas-best-shot-in-ebola-vaccine-race

349. Plachouras D., Monnet D.L. and Catchpole M. 2015. Severe Ebola virus infection complicated by gram-negative septicemia. *N Engl J Med*, 372(14), 1376–1377.

350. Pleguezuelos O., Robinson S., Stoloff G.A. and Caparrós-Wanderley W. 2012. Synthetic Influenza vaccine (FLU-v) stimulates cell mediated immunity in a double-blind, randomised, placebo-controlled Phase I trial. *Vaccine,* 30(31), 4655–4660.

351. Plowright R.K., Eby P., Hudson P., Smith I.L., Westcott D. et al. 2015. Ecological dynamics of emerging bat virus spillover. *Proc Biol Sci,* 282(1798), 20142124.

352. Poch O., Blumberg B.M., Bougueleret L. and Tordo N. 1990. Sequence comparison of five polymerases (L proteins) of unsegmented negative-strand RNA viruses: theoretical assignment of functional domains. *J Gen Virol,* 71(Pt 5), 1153-1162.

353. Ponomarenko J., Bui H.H., Li W., Fusseder N., Bourne P.E. et al. 2008. ElliPro: a new structure-based tool for the prediction of antibody epitopes. *BMC Bioinformatics*, 9, 514.

354. Ponomarenko J.V. and Van Regenmortel M.H. 2009. B cell epitope prediction. *Struct Bioinfo,* 35, 849–879.

355. Prins K.C., Binning J.M., Shabman R.S., Leung D.W., Amarasinghe G.K. et al. 2010. Basic residues within the ebolavirus VP35 protein are required for its viral polymerase cofactor function. *J Virol,* 84(20), 10581-10591.

356. Pritam M., Singh G., Swaroop S., Singh A.K. and Singh S.P. 2019. Exploitation of reverse vaccinology and immunoinformatics as promising platform for genome-wide screening of new effective vaccine candidates against Plasmodium falciparum. *BMC Bioinformatics*, 19(13), 468.

357. Pushko P., Bray M., Ludwig G.V., Parker M., Schmaljohn A. et al. 2000. Recombinant RNA replicons derived from attenuated Venezuelan equine encephalitis virus protect guinea pigs and mice from Ebola hemorrhagic fever virus. *Vaccine,* 19(1), 142–153.

358. Qiu X. and Kobinger G.P. 2014. Antibody therapy for Ebola: is the tide turning around? *Hum Vac Immunother,* 10(4), 964–967.

359. Qiu X., Wong G., Audet J., Bello A., Fernando L. et al. 2014. Reversion of advanced Ebola virus disease in nonhuman primates with zmapp. *Nature*, 514(7520), 47–53.

360. Radzimanowski J., Effantin G. and Weissenhorn W. 2014. Conformational plasticity of the Ebola virus matrix protein. *Protein Sci,* 23(11), 1519-1527.

361. Rahi A., Dhiman A., Singh D., Lynn A.M., Rehan M. et al. 2018. Exploring the interaction between Mycobacterium tuberculosis enolase and human plasminogen using computational methods and experimental techniques. *J Cell Biochem*, 119(2), 2408- 2417.

362. Rammensee H., Bachmann J., Emmerich N.P., Bachor O.A. and Stevanovic S. 1999. SYFPEITHI: database for MHC ligands and peptide motifs. *Immunogenetics*, 50(3-4), 213-219.

363. Rao M., Matyas G.R., Grieder F., Anderson K., Jahrling P.B. et al. 1999. Cytotoxic T lymphocytes to Ebola Zaire virus are induced in mice by immunization with liposomes containing lipid A. *Vaccine*, 17(23–24), 2991–2998.

364. Rathakrishnan A., Wang S.M., Hu Y., Khan A.M., Ponnampalavanar S. et al. 2012. Cytokine Expression Profile of Dengue Patients at Different Phases of Illness. *PLoS One*, 7(12), e52215.

365. Raymond D.D., Piper M.E., Gerrard S.R., Skiniotis G. and Smith J.L. 2012. Phleboviruses encapsidate their genomes by sequestering RNA bases. *Proc Natl Acad Sci USA*, 109(47), 19208–19213.

366. Reche P.A. and Reinherz E.L. 2007. Prediction of peptide-MHC binding using profiles. *Methods Mol Biol*, 409, 185–200.

367. Reguera J., Malet H., Weber F. and Cusack S. 2013. Structural basis for encapsidation of genomic RNA by La Crosse Orthobunyavirus nucleoprotein. *Proc Natl Acad Sci USA*, 110(18), 7246–7251.

368. Reguera J., Cusack S. and Kolakofsky D. 2014. Segmented negative strand RNA virus nucleoprotein structure. *Curr Opin Virol*, 5, 7–15.

369. Reid S.P., Ca´rdenas W.B. and Basler C.F. 2005. Homo-oligomerization facilitates the interferon-antagonist activity of the ebolavirus VP35 protein. *Virology*, 341(2), 179–189.

370. Rhein B.A., Powers L.S., Rogers K., Anantpadma M., Singh B.K. et al. 2015. Interferon-γ Inhibits Ebola Virus Infection. *PLoS Pathog*, 11(11), e1005263.

371. Richardson J.S., Pillet S., Bello A.J. and Kobinger G.P. 2013. Airway delivery of an adenovirus-based Ebola virus vaccine bypasses existing immunity to homologous adenovirus in nonhuman primates. *J Virol*, 87(7), 3668–77.

372. Riemenschneider J., Garrison A., Geisbert J., Jahrling P., Hevey M. et al. 2003. Comparison of individual and combination DNA vaccines for *B. anthracis*, Ebola virus, Marburg virus and Venezuelan equine encephalitis virus. *Vaccine*, 21(25-26), 4071–4080.

373. Ripoll D.R., Mitchell D.A., Dupuy L.C., Wallqvist A., Schmaljohn C. et al. 2017. Combinatorial peptide-based epitope mapping from Ebola virus DNA vaccines and infections reveals residue-level determinants of antibody binding. *Hum Vaccin Immunother*, 13(12), 2953–2966.

374. Rose N.F., Roberts A., Buonocore L. and Rose J.K. 2000. Glycoprotein exchange vectors based on vesicular stomatitis virus allow effective boosting and generation of neutralizing antibodies to a primary isolate of human immunodeficiency virus type 1. *J Virol,* 74(23), 10903–10910.

375. Rose N.F., Marx P.A., Luckay A., Nixon D.F., Moretto W.J. et al. 2001. An effective AIDS vaccine based on live attenuated vesicular stomatitis virus recombinants. *Cell,* 106(5), 539–549.

376. Rosendahl Huber S., van Beek J., de Jonge J., Luytjes W. and van Baarle D. 2014. T cell responses to viral infections - opportunities for peptide vaccination. *Front Immunol*, 5, 171.

377. Rosendahl Huber S.K., Camps M.G.M., Jacobi R.H.J., Mouthaan J., van Dijken H. et al. 2015. Synthetic long peptide influenza vaccine containing conserved t and b cell epitopes reduces viral load in lungs of mice and ferrets. *PLoS One*, 10(6), e0127969.

378. Roy S., Zhi Y., Kobinger G.P., Figueredo J., Calcedo R. et al. 2006. Generation of an adenoviral vaccine vector based on simian adenovirus 21. *J Gen Virol,* 87(Pt 9), 2477–2485.

379. Rubinstein N.D., Mayrose I., Martz E. and Pupko T. 2009. Epitopia: A web-server for predicting B-cell epitopes. *BMC Bioinform*, 10, 287.

380. Ruibal P., Oestereich L., Lüdtke A., Becker-Ziaja B., Wozniak D.M. et al. 2016. Unique human immune signature of Ebola virus disease in Guinea. *Nature*, 533(7601), 100–104.

381. Ruigrok R.W.H., Crépin T. and Kolakofsky D. 2011. Nucleoproteins and nucleocapsids of negative-strand RNA viruses. *Curr Opin Microbiol,* 14(4), 504–510.

382. Saeed M.F., Kolokoltsov A.A., Albrecht T. and Davey R.A. 2010. Cellular entry of ebola virus involves uptake by a macropinocytosis-like mechanism and subsequent trafficking through early and late endosomes. *PLoS Pathog*, 6(9), e1001110.

383. Safari S., Baratloo A., Rouhipour A., Ghelichkhani P. and Yousefifard M. 2015. Ebola hemorrhagic fever as a public health emergency of international concern; A review article. *Emerg (Tehran),* 3(1), 3-7.

384. Saha S., Bhasin M. and Raghava G.P.S. 2005. Bcipep: A database of B-cell epitopes. *BMC Genomics*, 6, 79.

385. Saha S. and Raghava G. 2006. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. *Proteins,* 65(1), 40–8.

386. Saha S. and Raghava G.P. 2006. AlgPred: prediction of allergenic proteins and mapping of IgE epitopes. *Nucleic Acids Res*, 34, W202–W209.

387. Saijo M., Niikura M., Morikawa S., Ksiazek T.G., Meyer R.F. et al. 2001. Enzyme-linked immunosorbent assays for detection of antibodies to Ebola and Marburg viruses using recombinant nucleoproteins. *J Clin Microbiol*, 39(1), 1–7.

388. Sakabe S., Sullivan B.M., Hartnett J.N., Robles-Sikisaka R., Gangavarapu K. et al. 2018. Analysis of CD8+ T cell response during the 2013–2016 Ebola epidemic in West Africa. *Proc Natl Acad Sci USA*, 115(32), E7578-E7586.

389. Sakib M.S., Islam M.R., Hasan A.K.M.M. and Nabi A.H.M.N. 2014. Prediction of epitope-based peptides for the utility of vaccine development from fusion and glycoprotein of nipah virus using in silico approach. *Adv Bioinformatics*, 2014, 402492.

390. Sakurai Y., Kolokoltsov A.A., Chen C.C., Tidwell M.W., Bauta W.E. et al. 2015. Ebola Virus. Two-pore channels control Ebola virus host cell entry and are drug targets for disease treatment. *Science*, 347(6225), 995–998.

391. Samykannu G., Vijayababu P., Antonyraj C.B., Perumal P., Narayanan S. et al. 2019. In Silico Characterization of B Cell and T Cell Epitopes for Subunit Vaccine Design of Salmonella typhi PgtE: A Molecular Dynamics Simulation Approach. *J Comput Biol*, 26(2), 105-116.

392. Sanchez A., Trappier S.G., Mahy B.W., Peters C.J. and Nichol S.T. 1996. The virion glycoproteins of Ebola viruses are encoded in two reading frames and are expressed through transcriptional editing. *Proc Natl Acad Sci USA*, 93(8), 3602–3607.

393. Sanchez A., Yang Z.Y., Xu L., Nabel G.J., Crews T. et al. 1998. Biochemical analysis of the secreted and virion glycoproteins of Ebola virus. *J Virol*, 72(8), 6442–6447.

394. Sanchez A. and Rollin P.E. 2005. Complete genome sequence of an Ebola virus (Sudan species) responsible for a 2000 outbreak of human disease in Uganda. *Virus Res*, 113(1), 16-25.

395. Sanchez A., Geisbert T.W. and Feldmann H. 2007. Filoviridae: Marburg and Ebola Viruses. *Fields Virology*, 1409-1448.

396. Sankar S., Ramamurthy M., Suganya S., Nandagopal B. and Sridharan G. 2018. Design of peptide epitope from the neuraminidase protein of influenza A and influenza B towards short peptide vaccine development. *Bioinformation*, 14(5), 183-189.

397. Santos-Lima E.K.N., Oliveira Y.A., Santos R.P.B., Sampaio G.P., Pimentel A.C.M. et al. 2019. Production of interferon-gamma, interleukin-6, and interleukin-1β by human

peripheral blood mononuclear cells stimulated with novel lys-gingipain synthetic peptides. *J Periodontol*, 18-0626.

398. Schieffelin J.S., Shaffer J.G., Goba A., Gbakie M., Gire S.K. et al. 2014. Clinical illness and outcomes in patients with Ebola in Sierra Leone. *N Engl J Med*, 371(22), 2092–2100.

399. Schmidt M.L. and Hoenen T. 2017. Characterization of the catalytic center of the Ebola virus L polymerase. *PLoS Negl Trop Dis*, 11(10), e0005996.

400. Schmitt M., Schmitt A., Rojewski M.T., Chen J., Giannopoulos K. et al. 2008. RHAMM-R3 peptide vaccination in patients with acute myeloid leukemia, myelodysplastic syndrome, and multiple myeloma elicits immunologic and clinical responses. *Blood*, 111(3), 1357-1365.

401. Schneider-Futschik E.K., Hoyer D., Khromykh A.A., Baell J.B., Marsh G.A. et al. 2018. Contemporary Anti-Ebola Drug Discovery Approaches and Platforms. *ACS Infect Dis*, 5(1), 35-48.

402. Schornberg K., Matsuyama S., Kabsch K., Delos S., Bouton A. et al. 2006. Role of endosomal cathepsins in entry mediated by the Ebola virus glycoprotein. *J Virol*, 80(8), 4174–4178.

403. Schumann M., Gantke T. and Muhlberger E. 2009. Ebola virus VP35 antagonizes PKR activity through its C-terminal interferon inhibitory domain. *J Virol*, 83(17), 8993–8997.

404. Shabman R.S., Gulcicek E.E., Stone K.L. and Basler C.F. 2011. The Ebola virus VP24 protein prevents hnRNP C1/C2 binding to karyopherin alpha1 and partially alters its nuclear import. *J Infect Dis*, 204(3), S904–S910.

405. Shen Y., Maupetit J., Derreumaux P. and Tuffery P. 2014. Improved PEP-FOLD approach for peptide and miniprotein structure prediction. *J Chem Theory Comput*, 10(10), 4745–4758 .

406. Shoemaker T., MacNeil A., Balinandi S., Campbell S., Wamala J.F. et al. 2012. Reemerging Sudan Ebola Virus Disease in Uganda, 2011. *Emerg Infect Dis*, 18(9), 1480-1483.

407. Simmons G., Lee A., Rennekamp A.J., Fan X., Bates P. et al. 2004. Identification of murine T-cell epitopes in Ebola virus nucleoprotein. *Virology*, 318(1), 224–230.

408. Simmons J.A., D'Souza R.S., Ruas M., Galione A., Casanova J.E. et al. 2016. Ebolavirus Glycoprotein Directs Fusion through NPC1+ Endolysosomes. *J Virol*, 90(1), 605–610.

409. Simpson J.C., Joggerst B., Laketa V., Verissimo F., Cetin C. et al. 2012. Genome-wide RNAi screening identifies human proteins with a regulatory function in the early secretory pathway. *Nat Cell Biol*, 14(7), 764–774.

410. Singh H. and Raghava G.P. 2001. ProPred: prediction of HLA-DR binding sites. *Bioinformatics*, 17(12), 1236-1237.

411. Singh B. 2004. PepBuild: a web server for building structure data of peptides/proteins. *Nucleic acids res.* 32, W559-61.

412. Singh S., Bajpai U. and Lynn A.M. 2014. Structure based virtual screening to identify inhibitors against MurE enzyme of *Mycobacterium tuberculosis* using AutoDock Vina. *Bioinformation*, 10(11), 697-702.

413. Singh T.R. and Pardasani K.R. 2009. In silico analysis of evolutionary patterns in restriction endonucleases. *In Silico Biol*, 9(1-2), 45–53.

414. Sirima S.B., Tiono A.B., Oue´draogo A., Diarra A., Oue´draogo A.L. et al. 2009. Safety and immunogenicity of the malaria vaccine candidate MSP3 long synthetic peptide in 12-24 months-old Burkinabe children. *PLoS One*, 4(10), e7549.

415. Sirskyj D., Diaz-Mitoma F., Golshani A., Kumar A. and Azizi A. 2011. Innovative bioinformatic approaches for developing peptidebased vaccines against hypervariable viruses. *Immunol Cell Biol*, 89(1), 81–89.

416. Skwarczynski M. and Toth I. 2016. Peptide-based synthetic vaccines. *Chem Sci*, 7(2), 842-854.

417. Smith L.M., Hensley L.E., Geisbert T.W., Johnson J., Stossel A. et al. 2013. Interferon-beta therapy prolongs survival in Rhesus macaque models of Ebola and Marburg hemorrhagic fever. *J Infect Dis*, 208(2), 310–318.

418. Soema P.C., van Riet E., Kersten G. and Amorij J.P. 2015. Development of cross-protective influenza a vaccines based on cellular responses. *Front Immunol*, 6, 237.

419. Soni S.P. and Stahelin R.V. 2014. The Ebola virus matrix protein VP40 selectively induces vesiculation from phosphatidylserine-enriched membranes. *J Biol Chem*, 289(48), 33590–33597.

420. Soria-Guerra R.E., Nieto-Gomez R., Govea-Alonso D.O. and Rosales-Mendoza S. 2015. An overview of bioinformatics tools for epitope prediction: implications on vaccine development. *J Biomed Inform*, 53, 405-414.

421. Sridhar S. 2015. Clinical development of Ebola vaccines. *Ther Adv Vaccines*, 3(5-6), 125.

422. Staff C., Mozaffari F., Fro¨din J.E., Mellstedt H. and Liljefors M. 2014. Telomerase (GV1001) vaccination together with gemcitabine in advanced pancreatic cancer patients. *Int J Oncol,* 45(3), 1293-1303.

423. Stanley D.A., Honko A.N., Asiedu C., Trefry J.C., Lau-Kilby A.W. et al. 2014. Chimpanzee adenovirus vaccine generates acute and durable protective immunity against ebolavirus challenge. *Nat Med,* 20(10), 1126–1129.

424. STAT, 2017. As foreign powers approve Ebola vaccines, U.S. drug makers lag in development pipeline. https://www.statnews.com/2017/12/08/ebola-vaccine-development/

425. Steer A.C., Carapetis J.R., Dale J.B., Fraser J.D., Good M.F. et al. 2016. Status of research and development of vaccines for Streptococcus pyogenes. *Vaccine*, 34(26), 2953-2958.

426. Stoloff G.A. and Caparros-Wanderley W. 2007. Synthetic multi-epitope peptides identified in silico induce protective immunity against multiple influenza serotypes. *Eur J Immunol*, 37(9), 2441–2449.

427. Strober W. 2001. Trypan blue exclusion test of cell viability. *Curr Protoc Immunol,* 111(A3.B), 1-3.

428. Sturniolo T., Bono E., Ding J., Raddrizzani L., Tuereci O. et al. 1999. Generation of tissue-specific and promiscuous HLA ligand databases using DNA microarrays and virtual HLA class II matrices. *Nat Biotechnol,* 17(6), 555-561.

429. Su Z., Wu C., Shi L., Luthra P., Pintilie G.D. et al. 2018. Electron Cryo-microscopy Structure of Ebola Virus Nucleoprotein Reveals a Mechanism for Nucleocapsid-like Assembly. *Cell,* 172(5), 966-978.

430. Subbarao K. and Matsuoka Y. 2013. The prospects and challenges of universal vaccines for influenza. *Trends Microbiol*, 21(7), 350-358.

431. Sullivan N.J., Sanchez A., Rollin P.E., Yang Z.Y. and Nabel G.J. 2000. Development of a preventive vaccine for Ebola virus infection in primates. *Nature,* 408(6812), 605–609.

432. Sullivan N.J., Geisbert T.W., Geisbert J.B., Xu L., Yang Z.Y. et al. 2003. Accelerated vaccination for Ebola virus haemorrhagic fever in non-human primates. *Nature,* 424(6949), 681–684.

433. Sun D.X., Seyer J.M., Kovari I., Sumrada R.A. and Taylor R.K. 1991. Localization of protective epitopes within the pilin subunit of the Vibrio cholerae toxin-coregulated pilus. *Infect Immun*, 59(1), 114–118.

434. Sun J., Wu D., Xu T., Wang X., Xu X. et al. 2009. SEPPA: A computational server for spatial epitope prediction of protein antigens. *Nucleic Acids Res,* 37, W612–W616.

435. Sun Y., Guo Y. and Lou Z. 2012. A versatile building block: the structures and functions of negative-sense single-stranded RNA virus nucleocapsid proteins. *Protein Cell*, 3(12), 893–902.

436. Sun P., Ju H., Liu Z., Ning Q., Zhang J. et al. 2013. Bioinformatics resources and tools for conformational B-cell epitope prediction. *Comput Math Methods Med*, 2013, 943636.

437. Sundar K., Boesen A. and Coico R. 2007. Computational prediction and identification of HLA-A2.1-specific Ebola virus CTL epitopes. *Virology*, 360(2), 257-263.

438. Swenson D.L., Wang D., Luo M., Warfield K.L., Woraratanadharm J. et al. 2008. Vaccine to confer to nonhuman primates complete protection against multistrain Ebola and Marburg virus infections. *Clin. Vaccine Immunol,* 15(3), 460–467.

439. Sweredoski M.J. and Baldi P. 2008. PEPITO: Improved discontinuous B-cell epitope prediction using multiple distance thresholds and half sphere exposure. *Bioinformatics*, 24(12), 1459–1460.

440. Szakonyi G., Klein M.G., Hannan J.P., Young K.A., Ma R.Z. et al. 2006. Structure of the Epstein-Barr virus major envelope glycoprotein. *Nat Struct Mol Biol,* 13(11), 996-1001.

441. Tariq A., Roosa K., Mizumoto K. and Chowell G. 2019. Assessing reporting delays and the effective reproduction number: The Ebola epidemic in DRC, May 2018-January 2019. *Epidemics*, 26, 128-133.

442. Tawar R.G., Duquerroy S., Vonrhein C., Varela P.F., Damier-Piolle L. et al. 2009. Crystal structure of a nucleocapsid-like nucleoprotein-RNA complex of respiratory syncytial virus. *Science*, 326(5957), 1279-83.

443. Theaker S.M., Rius C., Greenshields-Watson A., Lloyd A., Trimby A. et al. 2016. T-cell libraries allow simple parallel generation of multiple peptide-specific human T-cell clones. *J Immunol Methods*, 430, 43–50.

444. Thude H., Hardt C., Schorner U., Ferencik S., Helfricht C. et al. 2005. Identification of a new HLA-B allele, HLA-B*4443, in a German family. *Tissue Antigens*, 66(6), 696–699.

445. Tipu H.N. 2016. Immunoinformatic Analysis of Crimean Congo Hemorrhagic Fever Virus Glycoproteins and Epitope Prediction for Synthetic Peptide Vaccine. *J Coll Physicians Surg Pak*, 26(2), 108-112.

446. Topaloglu H., Berker M., Kansu T., Saatci U. and Renda Y. 1992. Optic neuritis and myelitis after booster tetanus toxoid vaccination. *Lancet*, 339(8786), 178–179.

447. Towner J.S., Sealy T.K., Khristova M.L., Albariño C.G., Conlan S. et al. 2008. Newly discovered Ebola virus associated with hemorrhagic fever outbreak in Uganda. *PLoS Pathog*, 4(11), e1000212.

448. Trost B., Bickis M. and Kusalik A. 2007. Strength in numbers: achieving greater accuracy in MHC-I binding prediction by combining the results from multiple prediction tools. *Immunome Res*, 3, 5.

449. Trott O. and Olson A.J. 2010. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem*, 31(2), 455-461.

450. Trunschke M., Conrad D., Enterlein S., Olejnik J., Brauburger K. et al. 2013. The L-VP35 and L-L interaction domains reside in the amino terminus of the Ebola virus L protein and are potential targets for antivirals. *Virology*, 441(2), 135–145.

451. Twiddy S.S., Holmes E.C. and Rambaut A. 2003. Inferring the rate and time-scale of dengue virus evolution. *Mol Biol Evol*, 20(1), 122–129.

452. Uddin S., Majchrzak B., Woodson J., Arunkumar P., Alsayed Y. et al. 1999. Activation of the p38 mitogenactivated protein kinase by type I interferons. *J Biol Chem*, 274(42), 30127–30131.

453. UNIPROT, 1998. UniProtKB - Q05322 (VP24_EBOZM). https://www.uniprot.org/uniprot/Q05322

454. Usmani S.S., Kumar R., Bhalla S., Kumar V. and Raghava G.P.S. 2018. In Silico Tools and Databases for Designing Peptide-Based Vaccine and Drugs. *Adv Protein Chem Struct Biol*, 112, 221-263.

455. van Doorn E., Liu H., Ben-Yedidia T., Hassin S., Visontai I. et al. 2017. Evaluating the immunogenicity and safety of a BiondVax-de- veloped universal influenza vaccine (Multimeric-001) either as a standalone vaccine or as a primer to H5N1 influenza vaccine: Phase IIb study protocol. *Medicine (Baltimore)*, 96 (11), e6339 .

456. Vanderzanden L., Bray M., Fuller D., Roberts T., Custer D. et al. 1998. DNA vaccines expressing either the GP or NP genes of Ebola virus protect mice from lethal challenge. *Virology*, 246(1), 134–144.

457. Vani J., Shaila M.S., Chandra N.R. and Nayak R. 2006. A combined immuno-informatics and structure-based modeling approach for prediction of T cell epitopes of secretory proteins of Mycobacterium tuberculosis. *Microbes Infect*, 8(3), 738–746.

458. Vici P., Pizzuti L., Mariani L., Zampa G., Santini D. et al. 2016. Targeting immune response with therapeutic vaccines in premalignant lesions and cervical cancer: hope or reality from clinical studies. *Expert Rev Vaccines,* 15(10), 1327-1336.

459. Vita R., Overton J.A., Greenbaum J.A., Ponomarenko J., Clark J.D. et al. 2015. The immune epitope database (IEDB) 3.0. *Nucleic Acids Res,* 43, D405-D412.

460. Volchkov V.E., Blinov V.M. and Netesov S.V. 1992. The envelope glycoprotein of Ebola virus contains an immunosuppressive-like domain similar to oncogenic retroviruses. *FEBS Lett,* 305(3), 181–184.

461. Volchkov V.E., Becker S., Volchkova V.A., Ternovoj V.A., Kotov A.N. et al. 1995. GP mRNA of Ebola Virus Is Edited by the Ebola Virus Polymerase and by T7 and Vaccinia Virus Polymerases. *Virology,* 214(2), 421-430.

462. Volchkov V.E., Volchkova V.A., Chepurnov A.A., Blinov V.M., Dolnik O. et al. 1999. Characterization of the L gene and 5' trailer region of Ebola virus. *J Gen Virol,* 80(Pt 2), 355-362.

463. Wald A., Koelle D.M., Fife K., Warren T., LeClair K. et al. 2011. Safety and immunogenicity of long HSV-2 peptides complexed with rhHsc70 in HSV-2 seropositive persons. *Vaccine,* 29(47), 8520-8529.

464. Wang L.F., Harcourt B.H., Yu M., Tamin A., Rota P.A. et al. 2001. Molecular biology of Hendra and Nipah viruses. *Microbes Infect*, 3(4), 279-287.

465. Wang Y., Wu W., Negre N.N., White K.P., Li C. et al. 2011. Determinants of antigenicity and specificity in immune response for protein sequences. *BMC Bioinformatics*, 12, 251.

466. Wang Y., Li J., Hu Y., Liang Q., Wei M. et al. 2017. Ebola vaccines in clinical trial: the promising candidates. *Hum Vaccin Immunother,* 13(1), 153–168.

467. Wang W., Wang F., Zhang W., He J. and Zhu F.M. 2019. Description of two new HLA alleles: HLA-A*30:118 and HLA-C*03:02:17. *HLA,* doi: 10.1111/tan.13625.

468. Warfield K.L., Olinger G., Deal E.M., Swenson D.L., Bailey M. et al. 2005. Induction of humoral and CD8+ T cell responses are required for protection against lethal Ebola virus infection. *J Immunol,* 175(2), 1184–1191.

469. Warfield K.L., Swenson D.L., Olinger G.G., Kalina W.V., Aman M.J. et al. 2007. Ebola virus-like particlebased vaccine protects nonhuman primates against lethal Ebola virus challenge. *J Infect Dis,* 196(2), S430–S437.

470. Warren T.K., Wells J., Panchal R.G., Stuthman K.S., Garza N.L. et al. 2014. Protection against filovirus diseases by a novel broad-spectrum nucleoside analogue bcx4430. *Nature*, 508(7496), 402–405.

471. Watanabe S., Noda T. and Kawaoka Y. 2006. Functional mapping of the nucleoprotein of Ebola virus. *J Virol*, 80(8), 3743–3751.

472. Watt A., Moukambi F., Banadyga L., Groseth A., Callison J. et al. 2014. A novel life cycle modeling system for Ebola virus shows a genome length-dependent role of VP24 in virus infectivity. *J Virol*, 88(18), 10511.

473. Wei H., Wang S., Chen Q., Chen Y., Chi X. et al. 2016. Suppression of interferon lambda signaling by SOCS-1 results in their excessive production during influenza virus infection. *PLoS Pathog*, 12(1), e1005402.

474. Weichert H., Blechschmidt I., Schroder S. and Ambrosius H. 1991. The MTT-assay as a rapid test for cell proliferation and cell killing: application to human peripheral blood lymphocytes (PBL). *Allerg Immunol (Leipz)*, 37(3-4), 139–144.

475. Weissenhorn W., Carfi A., Lee K.H., Skehel J.J. and Wiley D.C. 1998. Crystal structure of the Ebola virus membrane fusion subunit, GP2, from the envelope glycoprotein ectodomain. *Mol Cell*, 2(5), 605–616.

476. Wherry E.J. and Ahmed R. 2004. Memory CD8 T-Cell Differentiation during Viral Infection. *J Virol*, 78(11), 5535-5545.

477. White J.M. and Schornberg K.L. 2012. A new player in the puzzle of filovirus entry. *Nat Rev Microbiol*, 10(5), 317-322.

478. World Health Organization, 1978. Ebola haemorrhagic fever in Sudan, 1976. Report of a WHO/International Study Team. *Bulletin of the World Health Organization*, 56(2), 247-270.

479. World Health Organization, 1978. Ebola haemorrhagic fever in Zaire, 1976. Report of an International Commission. *Bulletin of the World Health Organization*, 56(2), 271-293.

480. World Health Organization, 1996. Ebola haemorrhagic fever – South Africa. *Weekly Epidemiological Record*, 71(47), 359.

481. World Health Organization, 2003. Outbreak(s) of Ebola haemorrhagic fever, Congo and Gabon, October 2001- July 2002. *Weekly Epidemiological Report*, 78(26), 223-225.

482. World Health Organization, 2004. Ebola haemorrhagic fever in the Republic of the Congo – Update 6. *Weekly Epidemiological Record.* https://www.who.int/csr/don/2004_01_06/en/

483. World Health Organization, 2005. Outbreak of Ebola haemorrhagic fever in Yambio, south Sudan, April-June 2004. *Weekly Epidemiological Record*, 80(43), 370-375.

484. World Health Organization, 2007. Ebola virus haemorrhagic fever, Democratic Republic of the Congo – Update. *Weekly Epidemiological Record*, 82(40), 345-346.

485. World Health Organization, 2009. End of the Ebola Outbreak in the Democratic Republic of the Congo. *Global Alert and Response*. https://www.who.int/csr/don/2009_02_17/en/

486. World Health Organization, 2014. Ebola Virus Disease. https://www.who.int/news-room/fact-sheets/detail/ebola-virus-disease

487. World Health Organization, 2015. Ebola Virus Disease – Italy. Disease Outbreak News. https://www.who.int/csr/don/13-may-2015-ebola/en/

488. WHO, 2019. Ebola virus disease – Democratic Republic of the Congo Disease outbreak news: Update. https://www.who.int/csr/don/18-july-2019-ebola-drc/en/

489. Wiersma L., Rimmelzwaan G.F. and de Vries R.D. 2015. Developing Universal Influenza Vaccines: Hitting the Nail, Not Just on the Head. *Vaccines (Basel)*, 3(2), 239-262.

490. Willet M., Kurup D., Papaneri A., Wirblich C., Hooper J.W. et al. 2015. Preclinical development of inactivated rabies virus-based polyvalent vaccine against rabies and filoviruses. *J Infect Dis*, 212(2), S414–S424.

491. Wilson J.A. and Hart M.K. 2001. Protection from Ebola virus mediated by cytotoxic T lymphocytes specific for the viral nucleoprotein. *J Virol*, 75(6), 2660–2664.

492. Wong G., Audet J., Fernando L., Fausther-Bovendo H., Alimonti J.B. et al. 2014. Immunization with vesicular stomatitis virus vaccine expressing the Ebola glycoprotein provides sustained long-term protection in rodents. *Vaccine*, 32(43), 5722–5729.

493. Woodland D.L. 2004. Jump-starting the immune system: prime-boosting comes of age. *Trends Immunol*, 25(2), 98-104.

494. Wool-Lewis R.J. and Bates P. 1999. Endoproteolytic processing of the Ebola virus envelope glycoprotein: cleavage is not required for function. *J Virol*, 73(2), 1419–1426.

495. Wu S., Yu T., Song X., Yi S., Hou L. et al. 2012. Prediction and identification of mouse cytotoxic T lymphocyte epitopes in Ebola virus glycoproteins. *Virol J*, 9, 111.

496. Wyatt R., Kwong P.D., Desjardins E., Sweet R.W., Robinson J. et al. 1998. The antigenic structure of the HIV gp120 envelope glycoprotein. *Nature*, 393(6686), 705–711.

497. Yabukarski F., Lawrence P., Tarbouriech N., Bourhis J.M., Delaforge E. et al. 2014. Structure of Nipah virus unassembled nucleoprotein in complex with its viral chaperone. *Nat Struct Mol Biol*, 21(9), 754–759.

498. Yao B., Zheng D., Liang S. and Zhang C. 2013. Conformational B-cell epitope prediction on antigen protein structures: a review of current algorithms and comparison with common binding site prediction methods. *PLoS One*, 19(8), e62249.

499. Yasmin T., Akter S., Debnath M., Ebihara A., Nakagawa T. et al. 2016. In silico proposition to predict cluster of B- and T-cell epitopes for the usefulness of vaccine design from invasive, virulent and membrane associated proteins of C. jejuni. *In Silico Pharmacology*, 4(1), 1-10.

500. Yassin G.M., Amin M.A. and Attia A.S. 2016. Immunoinformatics identifies a lactoferrin binding protein a peptide as a promising vaccine with a global protective prospective against moraxella catarrhalis. *J Infect Dis*, 213(12), 1938–1945.

501. Yin H.S., Wen X., Paterson R.G., Lamb R.A. and Jardetzky T.S. 2006. Structure of the parainfluenza virus 5F protein in its metastable, prefusion conformation. *Nature*, 439(7072), 38–44.

502. Yiu, H.H., Graham, A.L. and Stengel R.F. 2012. Dynamics of a Cytokine Storm. *PLoS One*, 7(10), e45027.

503. Yoshimura K., Minami T., Nozawa M. and Uemura H. 2013. Phase I clinical trial of human vascular endothelial growth factor receptor 1 peptide vaccines for patients with metastatic renal cell carcinoma. *Br J Cancer*, 108(6), 1260-1266.

504. Yu D.S., Weng T.H., Wu X.X., Wang F.X.C., Lu X.Y. et al. 2017. The lifecycle of the Ebola virus in host cells. *Oncotarget*, 8(33), 55750-55759.

505. Zampieri C.A., Sullivan N.J. and Nabel G.J. 2007. Immunopathology of highly virulent pathogens: Insights from Ebola virus. *Nat Immunol*, 8(11), 1159–1164.

506. Zhang H., Wang L., Compans R.W. and Wang B.Z. 2014. Universal influenza vaccines, a dream to be realized soon. *Viruses*, 6(5), 1974-1991.

507. Su Z., Wu C., Shi L., Luthra P., Pintilie G.D. et al. 2018. Electron Cryo-microscopy Structure of Ebola Virus Nucleoprotein Reveals a Mechanism for Nucleocapsid-like Assembly. *Cell*, 172(5), 966-978.

508. Zhou H., Sun Y., Guo Y. and Lou Z. 2013. Structural perspective on the formation of ribonucleoprotein complex in negative-sense single-stranded RNA viruses. *Trends Microbiol*, 21(9), 475–484.

509. Zinzula L., Nagy I., Orsini M., Weyher-Stingl E., Bracher A. et al. 2018. Structures of Ebola and Reston Virus VP35 Oligomerization Domains and Comparative Biophysical Characterization in All Ebolavirus Species. *Structure*, 27(1), 39-54.

510. Zinzula L. and Tramontano E. 2013. Strategies of highly pathogenic RNA viruses to block dsRNA detection by RIG-I-like receptors: hide, mask, hit. *Antiviral Res*, 100(3), 615–635.

Annexure – I

GP unique sequence IDs			NP unique sequence IDs		
>gb|AJP1545	>gb|AKG957	>gb|AKI840	>gb|AGB566	>gb|ALG014	>gb|AJP1436
>gb|AER597	>gb|AKG959	>gb|AKI841	>gb|ABY753	>gb|AKG958	>gb|ALX320
>gb|AKU751	>gb|AKG962	>gb|ALF045	>gb|AFP282	>gb|AKG653	>gb|AKG958
>gb|AKU751	>gb|AKG956	>gb|AKG656	>gb|AGL734	>gb|AKC364	>gb|AJP1400
>gb|AKU752	>gb|AIE1192	>gb|AJZ747	>gb|ACI2862	>gb|AKC364	>gb|AKU752
>gb|AKU752	>gb|AKC359	>gb|AKI831	>gb|AGL734	>gb|ALX339	>gb|AKG960
>gb|AKU752	>gb|AKC359	>gb|AKI842	>gb|AKB095	>gb|AKG960	>gb|AKU754
>gb|AKU752	>gb|AKC359	>gb|AJZ745	>gb|AGL509	>gb|ABW34	>gb|AKG657
>gb|AKU752	>gb|AKC359	>gb|AJZ745	>gb|AKN201	>gb|ALX328	>gb|AKI8277
>gb|AKU752	>gb|AKC359	>gb|ALG015	>gb|AKB095	>gb|AKU755	>gb|AJP1532
>gb|AKU753	>gb|AKC360	>gb|AJZ746	>gb|AKB095	>gb|AKI8419	>gb|AHX246
>gb|AKU754	>gb|AKC360	>gb|ALG010	>gb|AKU754	>gb|AJP1445	>gb|ALH214
>gb|AKU754	>gb|AKC360	>gb|ALG009	>gb|ALX335	>gb|AKI8268	>gb|ALG010
>gb|AKU754	>gb|AKC361	>gb|ALG018	>gb|AKG957	>gb|ALT667	>gb|ALG011
>gb|AKU755	>gb|AKC361	>gb|ALG013	>gb|ALX332	>gb|ALX312	>gb|ALG012
>gb|AKU755	>gb|AKC362	>gb|ALG015	>gb|ALX320	>gb|AJP1516	>gb|AND812
>gb|AKU755	>gb|AKC363	>gb|ALG016	>gb|ALX320	>gb|AKG960	>gb|ABW34
>gb|AGB568	>gb|AKC363	>gb|ALG016	>gb|AKC369	>gb|ALG018	>gb|AKC367
>gb|AGB567	>gb|AKC364	>gb|ALG010	>gb|AKC364	>gb|AKC368	>gb|AKC368
>gb|AGB567	>gb|AKC364	>gb|ALG017	>gb|AKC364	>gb|AKC370	>gb|AJP1466
>gb|AGB567	>gb|AKC364	>gb|ALG013	>gb|AJP1455	>gb|AKC371	>gb|AJP1546
>gb|AGB568	>gb|AKC364	>gb|ALG019	>gb|AKI8341	>gb|AKI8352	>gb|AKI8328
>gb|AGB568	>gb|AKC364	>gb|ALG008	>gb|ALG013	>gb|AKC360	>gb|AKU752
>gb|AKA593	>gb|AKC364	>gb|ALG018	>gb|AKC371	>gb|AKC361	>gb|AKG657
>gb|AKC014	>gb|AKC364	>gb|AJZ747	>gb|AKC371	>gb|AJP1538	>gb|ALX317
>gb|AKC014	>gb|AKC365	>gb|ALX333	>gb|ALX315	>gb|AKG657	>gb|AMY60
>gb|AKC014	>gb|AKC365	>gb|ALX344	>gb|AKI8414	>gb|ALX311	>gb|AKC359
>gb|ALP300	>gb|AKC365	>gb|AKG959	>gb|ALX319	>gb|ALX310	>gb|AKG656
>gb|AKC014	>gb|AKC366	>gb|AKG961	>gb|AJP1496	>gb|ALX339	>gb|AKI8329

GP unique sequence IDs			NP unique sequence IDs		
>gb|AKT088	>gb|AKC367	>gb|AKG961	>gb|AJP1403	>gb|ALX336	>gb|AIG9624
>gb|ALF045	>gb|AKC367	>gb|AKG955	>gb|AKC365	>gb|ALG018	>gb|AKC363
>gb|AIO117	>gb|AKC367	>gb|AKI842	>gb|AKC364	>gb|AIE1186	>gb|AJP1413
>gb|AIR940	>gb|AKC367	>gb|AKI842	>gb|ABW34	>gb|AKC366	>gb|AKC369
>gb|AIY291	>gb|AKC368	>gb|AGL734	>gb|AKG653	>gb|ALX337	>gb|AKG651
>gb|AJA043	>gb|AKC368	>gb|AGL734	>gb|ALX325	>gb|ALX340	>gb|AND812
>gb|AJA044	>gb|AKC368	>gb|ACI286	>gb|AKC364	>gb|AKC359	>gb|AMY60
>gb|AIW659	>gb|AKC368	>gb|AFP282	>gb|AKI8275	>gb|AIE1180	>gb|AKG959
>gb|AKL910	>gb|AKC368	>gb|ABY753	>gb|AKG961	>gb|AKC369	>gb|ABW34
>gb|AKL911	>gb|AKC369	>gb|AGL734	>gb|AKI8304	>gb|AKI8363	>gb|AKC365
>gb|AKL911	>gb|AKC369	>gb|AGL509	>gb|AKI8347	>gb|AND812	>gb|AKC367
>gb|AKG657	>gb|AKC369	>gb|AGB566	>gb|ALX316	>gb|AKC360	>gb|AKC368
>gb|AKG652	>gb|AKC369	>gb|ACR331	>gb|AKC365	>gb|AKC362	>gb|ALG014
>gb|AKG657	>gb|AKC369	>gb|AAU438	>gb|AKC364	>gb|ALG012	>gb|AKI8343
>gb|AKG653	>gb|AKC369		>gb|ABW34	>gb|AKI8278	>gb|ALX310
>gb|AKG654	>gb|AKC370		>gb|ALX330	>gb|ALX312	>gb|AKC359
>gb|AKG657	>gb|AKC371		>gb|AKU755	>gb|AIG9636	>gb|AKC359
>gb|AKG652	>gb|AKC371		>gb|AJP1428	>gb|ABW34	>gb|AKI8275
>gb|AKG653	>gb|AKC371		>gb|AKC369	>gb|AKG653	>gb|ALX309
>gb|AKI826	>gb|AKC371		>gb|AKC363	>gb|ABW34	>gb|AKU755
>gb|AKI827	>gb|AKC372		>gb|ALX336	>gb|ALG016	>gb|AKA593
>gb|AKI827	>gb|AKC372		>gb|ALX325	>gb|ALG016	>gb|ALF045
>gb|AKI828	>gb|ALX343		>gb|ALX325	>gb|AKC366	>gb|AKC361
>gb|AKI828	>gb|AJP1481		>gb|ALX324	>gb|ALG011	>gb|AKC372
>gb|AKI830	>gb|AJP1519		>gb|AKG955	>gb|ALB071	>gb|ALX323
>gb|AKI833	>gb|AJP1528		>gb|ALX323	>gb|ALG015	>gb|AKC366
>gb|AKI834	>gb|AJP1532		>gb|AKC367	>gb|ALX322	>gb|AKC367
>gb|AKI834	>gb|AJP1443		>gb|ALX337	>gb|AKC359	>gb|AKC369
>gb|AKI834	>gb|AJP1542		>gb|ALX336	>gb|AKI8269	>gb|ALX317
>gb|AKI835	>gb|AJP1447		>gb|ALX343	>gb|AKU755	>gb|AKC359
>gb|AKI836	>gb|AJP1549		>gb|AKG955	>gb|ALX321	>gb|AKU753

GP unique sequence IDs		NP unique sequence IDs		
>gb|AKI836	>gb|AJP1397	>gb|ALX329	>gb|ABW34	>gb|AKI8338
>gb|AKI837	>gb|AJP1467	>gb|AKC368	>gb|AKC365	>gb|ALG013
>gb|AKI838	>gb|AJP1485	>gb|AKC368	>gb|AIE1192	>gb|ALG020
>gb|AKI838	>gb|AJP1486	>gb|AKC371	>gb|ALH214	>gb|AKC363
>gb|AKI839	>gb|ALH214	>gb|AKC369	>gb|AKC366	>gb|ALG015

Journal of Theoretical Biology 465 (2019) 34–44

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/jtb

Computational analysis in designing T cell epitopes enriched peptides of Ebola glycoprotein exhibiting strong binding interaction with HLA molecules

Sahil Jain, Manoj Baranwal*

Department of Biotechnology, Thapar Institute of Engineering and Technology, Patiala, 147004 Punjab, India

ARTICLE INFO

Article history:
Received 20 June 2018
Revised 28 December 2018
Accepted 9 January 2019
Available online 10 January 2019

Keywords:
Epitope mapping
In silico analysis
Conservation analysis
Docking
Ebola glycoprotein

ABSTRACT

Computational approach has shown remarkable progress in epitope mapping, paving the way to finding vaccine candidates against different viruses. In the current study, prediction algorithms and molecular docking were applied to select peptides containing multiple Ebola glycoprotein epitopes showing interaction with different HLA molecules. Six peptides containing overlapping multiple HLA I ($CD8^+$) and II ($CD4^+$) restricted T cell epitopes were generated via consensus approach applying six different prediction tools. Four (P1, P2, P5 and P6) out of six peptides were selected after screening for absence of undesirable responses and presence of B cell epitopes. Peptide-HLA interaction analysis based on Autodock Vina and CABS-dock showed strong binding of these four peptides with eighteen HLA molecules. HLA coverage analysis from each prediction tool showed that these peptides were able to bind to diverse HLA-A, HLA-B, HLA-DP, HLA-DQ and HLA-DR alleles. Population coverage analysis of peptides for expected immune response in four different continents (Africa, America, Asia and Europe) have shown average population coverage viz, P1 (95%), P2 (96%), P5 (91%) and P6 (94%). Further, these peptides were found to be nearly 100% conserved in Zaire Ebola virus while LANETTQALQLF (P5) was found to be 100% conserved in Zaire, Sudan, Bundibugyo and Tai Forest species. Therefore, these peptides capable of inducing T and B cell response and being presented by a wide range of HLA molecules have a strong potential to be part of diagnostic and preventive tools against Ebola virus disease.

1. Introduction

Ebola virus (EBOV) is an enveloped, non-segmented, negative-stranded RNA virus which belongs to the Filoviridae family (Lee et al., 2008). Five Ebola virus species (Zaire, Reston, Sudan, Bundibugyo, and Tai Forest) are known of which, Zaire is the most pathogenic and Reston is non-pathogenic in humans (Groseth et al., 2012). The mortality rate of Ebola virus disease (EVD) varies from 25%–90% with an average of 50% (Peters et al., 1999). More than 28,000 cases with over 11,000 deaths have been reported during its latest breakout (Cherpillod et al., 2016).

The 19 kb genome of Ebola virus consists of seven genes which are present in a defined order viz. nucleoprotein (NP), polymerase cofactors (VP35 and VP40), glycoprotein (GP), transcription activators (VP30 and VP24) and RNA-dependent RNA polymerase (L) (Brauburger et al., 2014). The fourth gene from the 3′ end of EBOV genome results in the formation of three proteins viz. secreted glycoprotein (sGP), soluble secreted glycoprotein (ssGP) and surface glycoprotein (GP) (Ito et al., 2001) which is cleaved post-translationally to yield GP1 and GP2 subunits connected by a disulfide link (Jeffers et al., 2002). GP1 plays an important role in attachment to host cells with the help of 54-201 residues that form a highly conserved putative receptor-binding site (RBS) (Kuhn et al., 2006; Manicasammy et al., 2005; Lee et al., 2008). GP2 is indispensable for viral and host membrane fusion (Volchkov et al., 1998; Feldmann et al., 1993; Takada et al., 1997). Presence of both GP1 and GP2 is critical for recognition by an antibody. GP1 is required to maintain the proper pre-fusion conformation of GP2 for antibody binding. Studies have revealed that antibodies bind to a non-glycosylated epitope at the base of the GP2 subunit where they interact with residues 42-43 at the N terminus of GP1 and 505-514 and 549-556 at the N terminus of GP2 (Lee et al., 2008). As EBOV GP is the first protein to interact with the immune system, it is considered as a potential candidate for vaccine and anti-filoviral therapeutic development (Madara et al., 2015).

sGP is the most abundant product of the glycoprotein gene (Iwasa et al., 2011) and it has been proven to result in antigenic subversion (a process wherein sGP acts as a barrier between the virus GP and anti-GP12 antibodies, thus, preventing contact

* Corresponding author.
E-mail address: manoj.baranwal@thapar.edu (M. Baranwal).

https://doi.org/10.1016/j.jtbi.2019.01.016
0022-5193/

between the virus GP and the antibodies) (Mohan et al., 2012). It has been reported that sGP shares 295 amino acids with GP1 (Iwasa et al., 2011). Hence, sGP may be an important target for vaccine development.

An effective prophylaxis or treatment for EVD is not yet available (Takada et al., 2015). Among the vaccines in various stages of clinical trials are recombinant adenoviruses, recombinant vesicular stomatitis viruses, recombinant human para influenza viruses and virus-like particles (Sykes et al., 2015; Sridhar, 2015). Efficient and prolonged response to the vaccine candidate and protection against different strains of the virus represent the current challenges faced by the scientific community (Sridhar, 2015; Hoenen et al., 2012).

A safer alternative route to live-attenuated or inactivated vaccines is the development of peptide based vaccines. They are relatively easier to produce and handle as there is no need to culture any pathogenic organisms. With the advancement in immunoinformatics tools focusing on epitope mapping, it has become easy to select peptides containing immunogenic epitopes. Immunoinformatics utilizes sequence and structure based approaches and has shown promising results in peptide based vaccine development. The immunogenic peptides obtained using this approach were validated in *in vitro* and *in vivo* system (Lohia and Baranwal, 2017; Oliveira et al., 2016; Kovjazin et al., 2013; Duvvuri et al., 2013). One of the noteworthy examples is *in silico* based selected influenza peptides which are the constituents of an influenza vaccine currently undergoing phase II clinical trial with the concept name of FluV (Pleguezuelos et al., 2012; van Doorn et al., 2017).

In the present study, consensus based predictions and molecular docking tools were employed to obtain the peptides containing multiple epitopes of Ebola glycoprotein which have the potential to interact with a wide range of HLA molecules.

2. Method

2.1. Conserved peptides identification

1092 sequences belonging to various Ebola species (1976–30th January 2018) pathogenic to humans were downloaded from Ebola virus database (viprbrc). 173 unique full length (676 amino acids) Ebola glycoprotein sequences (164, 5, 3 and 1 sequences of Zaire, Sudan, Bundibugyo and Tai Forest species respectively) were screened out after removing redundancy.

22 sGP sequences (364 amino acids) belonging to various Ebola species (1976–April 2015) pathogenic to humans were downloaded from the Ebola virus database (viprbrc). Six sequences found to be unique and considered for this study included three sequences for Zaire and one sequence each for Sudan, Bundibugyo and Tai Forest species.

Multiple sequence comparison by log-expectation (MUSCLE) is an iteration based multiple sequence alignment tool. It is faster than other available tools such as T-Coffee, MAFFT and CLUSTAL W and maintains the same accuracy (Edgar, 2004). It was employed for multiple sequence alignment of the obtained glycoprotein sequences which is a prerequisite for running AVANA.

Entropy is a measure of sequence variability and therefore, antigenic variability analyser (AVANA) tool was employed. It analyses multiple sequence alignments based on entropy and measures variability at a given position (Miotto et al., 2008). It is used to find conserved fragments and the criteria for conservation was set to $\geq 90\%$.

2.2. Prediction of $CD4^+$ and $CD8^+$ T cell epitopes

A consensus approach using six prediction tools to define $CD4^+$ and $CD8^+$ T cell epitopes has been utilized as mentioned in previous study (Lohia and Baranwal, 2014). Various tools were selected to consider different parameters and algorithms in order to increase the prediction stringency.

SYFPEITHI (Rammensee et al., 1999) and NetCTL 1.2 (Larsen et al., 2007) were used for HLA class I binding ($CD8^+$ T cells) epitopes and their threshold value was taken as 20 and 0.75 respectively. SYFPEITHI relies on published motifs for T cell epitope prediction. It considers the positions of amino acids and scores them based on whether they are anchor or auxiliary anchor position preferring residues (Rammensee et al., 1999). NetCTL 1.2 uses artificial neural network and weight matrix based method for predicting the HLA class I binding epitope. The server integrates the HLA class I epitope binding prediction with antigen processing steps including C-terminal proteasomal cleavage and TAP (transporter associated with antigen processing) transport efficiency (Larsen et al., 2007). It predicts epitopes for 12 HLA class I supertypes. Each supertype is a cluster of functionally related HLA alleles that share binding specificities towards the same panel of peptides owing to similar structural features of HLAs peptide binding groove (Lund et al., 2004).

MHC2Pred (Bhasin et al., 2004) and Propred (Singh et al., 2001) were employed for HLA class II binding ($CD4^+$ T cells) epitopes by taking criteria as 0.5 and 3% respectively. MHC2Pred is a support vector machine (SVM)-based method for prediction of promiscuous binders for 42 MHC class II alleles (Bhasin et al., 2004). ProPred is a matrix-based prediction algorithm for 51 HLA-DR alleles (HLA class II) (Singh et al., 2001). It employs amino acid /position coefficient table derived from pocket profile database described by Sturniolo and coworkers (Sturniolo et al., 1999).

IEDB ANN (Nielsen et al., 2003) and IEDB NN Align (Nielsen and Lund, 2009) tool was used for HLA I and II binding epitopes respectively which are based on artificial neural network. These tools calculate the IC50 value which shows the binding ability of peptides to specific HLA molecules and the threshold for IC50 values was taken as less than 500 nm.

The epitopes commonly predicted by all tools catering to a specific HLA class were considered. The overlapping epitopes were merged to obtain peptide fragments containing more than one epitope. In this way, peptides containing multiple epitopes were generated for both, HLA class I and class II. Further, the peptide fragments which were present commonly in both, HLA class I and class II peptide fragments were selected with the intention to consider peptides with a potential to induce both, $CD4^+$ and $CD8^+$ T cell response.

2.3. Presence of B cell epitopes

ABCpred was used to predict linear B cell epitopes by keeping the window length as 10. This tool utilizes a recurring neural network method to predict B cell epitopes (Saha et al., 2006a,b).

2.4. Screening of peptides for autoimmune, allergic and toxic response

BLAST analysis was carried out to eliminate peptides having 7 out of 9 consecutive identical amino acids to human proteome (Altschul et al., 1990). AlgPred and ToxinPred were used for screening of peptides having allergic and toxic nature respectively. AlgPred provides the option to follow various prediction approaches such as mapping of IgE epitopes and MEME/MAST motif to measure allergenicity of the query sequence (Saha et al., 2006a,b). ToxinPred classifies toxic and non-toxic peptides based on dataset of non-toxic or random peptides (Gupta et al., 2013).

2.5. Molecular docking

Molecular docking was put into service to understand peptide-HLA interactions. High-resolution crystallographic structures of eighteen class I and II HLA molecules (nine each) representing eighteen different HLA alleles were retrieved from the protein data bank (PDB) (Berman et al., 2000). The native peptides (peptides bound to the HLA molecule downloaded from PDB) were separated from their HLA molecule using the Discovery studio visualizer v4.1. The resultant HLA molecules were used as receptors to dock the separated native peptides (control) and *in silico* identified Ebola epitopes/peptides as ligands using Autodock Vina and CABS-dock.

The structure of identified epitopes/peptides was generated using the PEP-FOLD server. It is based on optimized potential for efficient protein structure prediction (OPEP) coarse grained force field for molecular simulation and provides a fast and user friendly platform for *de novo* design of small (9 to 36 residues) peptides (Shen et al., 2014). The top model structures predicted for the peptides were used for docking.

Autodock Vina helps to estimate the binding energy (Kcal/mol) of the peptide-HLA complex. The higher the negative value of the binding energy, the more stable is the peptide-HLA interaction. Positive binding energies indicate highly unstable and undesirable interaction. Docking with AutoDock vina involves various steps such as molecule preparation and grid optimization which are carried out with the help of AutoDockTools (ADT), a graphical user interface (GUI) employed by AutoDock (Trott and Olson, 2010). The receptor and ligands are prepared in PDBQT format which stores the atomic coordinates, partial charges and AutoDock atom types. In order to define the space (grid) where peptides are supposed to bind for presentation to T cells by HLA molecule, the thumbwheels in grid box option of ADT grid widget were used. Epitope/peptides showing positive binding energy were eliminated.

CABS-dock performs docking by allowing full flexibility of ligand and small fluctuations of the receptor backbone (Kurcinski et al., 2015). Also, it does not require a predefined epitope/peptide structure and therefore, only the pdb file of receptor (prepared in Discovery Studio Visualizer 4.1 by editing the HLA-native peptide file downloaded from PDB) and ligand sequences were provided. CABS-dock was used to estimate the root mean square deviation (RMSD) value of the native peptides (control) and identified epitope/peptides. Ligand-RMSD value gives an estimate of the docking pose quality. The lower the RMSD value, the better are the key binding interactions which indicate a high quality docking pose. RMSD values <3 are considered high quality predictions while $3 \leq \text{RMSD} \geq 5.5$ are considered moderate quality predictions (Blaszczyk et al., 2016).

2.6. Population coverage analysis

Population coverage analysis tool provides the expected response of each peptide in different geographic regions based on peptide-HLA data and HLA genotypic frequency. The selected peptides and their HLA alleles obtained from prediction tools were used as input for this tool. Four major geographical continents considered were Africa (Central, East, North, West and South Africa), America (Central, North and South America), Asia (East, South, Southwest, Northeast and Southeast Asia) and Europe. Average of population coverage for different regions of same continent was taken. Analysis was also carried out by considering the whole world.

2.7. Structural analysis

The three-dimensional structure of Ebola glycoprotein was obtained from PDB (PDB id 5JQ3) and used as base structure for mapping identified peptides. Discovery Studio Visualizer tool v4.1 was used to locate the peptide fragments on the base structure.

3. Results

The goal of this study was *in silico* elucidation of conserved peptide fragments capable of eliciting T ($CD8^+$ and $CD4^+$) and B cell response. Also, the identified fragments must be devoid of undesirable (autoimmune, allergic and toxic) responses. Further, these peptides should be presented by diverse HLA molecules and exhibit wide population coverage.

3.1. Conserved peptides containing T cell epitopes

Ten overlapping peptide fragments of surface glycoprotein (GP) with $\geq$90% conservancy were identified with the help of MUSCLE and AVANA (Supplementary Table 1). In soluble glycoprotein (sGP), two peptide fragments (YEAGEWAENCYNL and GAFFLYDRLAST) with $\geq$90% conservancy were identified which were present in conserved peptide C2 of GP (Supplementary Table 1). Therefore, the identified sGP conserved fragments were not treated separately henceforth. Three different tools each for HLA class I and II were applied to predict epitopes on conserved peptides and those epitopes were considered which were predicted commonly by all three tools. 84 $CD8^+$ T cell epitopes were obtained (data not shown) and overlapping epitopes were identified to generate peptide fragments containing multiple $CD8^+$T cell epitopes (minimum two epitopes). Fifteen peptide fragments containing multiple $CD8^+$ T cell epitopes were generated by merging the overlapping epitopes (Supplementary Table 2). Similarly, from the 77 $CD4^+$ T cell epitopes obtained (data not shown), thirteen peptide fragments containing multiple $CD4^+$ T cell epitopes were constructed (Supplementary Table 3). Now, to generate peptides which have the potential to elicit both $CD8^+$ and $CD4^+$ T cell response, the peptide fragments commonly present in fifteen peptides enriched with $CD8^+$ T cell epitopes and thirteen peptides enriched with $CD4^+$ T cell epitopes were chosen (Table 1). In this way, six peptide fragments (P1-P6) containing overlapping $CD8^+$ and $CD4^+$ T cell epitopes were selected.

3.2. Peptide screening for undesired responses and presence of B cell epitopes

Linear B cell epitope prediction (ABCpred) tool resulted into a total of 152 B cell epitopes and nine fragments were generated by merging the overlapping epitopes (Supplementary Table 4). The six identified fragments containing T cell epitopes (Table 1) were examined for the presence of predicted B cell epitopes and four peptides were found to contain B cell epitopes excluding VVAFLILPQAKK (P3) and YLFEVDNLTYVQLESRFT (P4). BLAST analysis was carried out and one peptide fragment (YLFEVDNLTYVQLESRFT) was eliminated owing to its similarity to human proteome. None of the remaining peptides were found to be allergic (AlgPred) or toxic (ToxinPred). After screening, four conserved peptides (P1, P2, P5 and P6) containing multiple T and B cell epitopes were selected (Table 2). Interestingly, P2 was found to contain one of conserved fragments **(GAFFLYDRLAST)** identified for sGP.

3.3. Peptide-HLA interactions

Peptides are presented to T cells by HLA molecules which results in T cell immune response. HLA molecules are found to be highly polymorphic in nature; hence, peptides capable of interacting with wide HLA molecules are potent vaccine candidates. Docking was carried out to analyze the binding behavior of peptides

Table 1
Glycoprotein peptides containing overlapping $CD8^+$ and $CD4^+$ T-cell epitopes. The fragments marked in bold are the ones which are commonly present in the $CD8^+$ and $CD4^+$T cell epitope enriched peptide fragments. **VVAFLILPQAKK** has been underlined to denote that it is not a part of **FHKEGAFFLYDRLASTVIYRGTTFAEG** fragment.

$CD8^+$ T cell epitope enriched peptides	No. of epitopes	$CD4^+$T cell epitope enriched peptides	No. of epitopes	Common peptides containing $CD8^+$ and $CD4^+$ T-cell epitopes
LPRDR**FKRTSFFLWVIILFQRTFSIPL**GV	10	**FKRTSFFLWVIILFQRTFSIPLG** VIHNSTLQVSDVDKLVCRDKLSS	14	**FKRTSFFLWVIILFQRTFSIPL (P1)**
GPCAGDFA**FHKEGAFFLYDRLASTVIYRGTTFA EGVVAFLILPQAKK**DFFSSHPL	14	**FHKEGAFFLYDRLASTVIYRGTTFAEG**	7	**FHKEGAFFLYDRLASTVIYRGTTFAEG (P2)**
		VVAFLILPQAKK	3	**VVAFLILPQAKK (P3)**
E**YLFEVDNLTYVQLESRFT**PQFLLQLNETIY	10	**YLFEVDNLTYVQLESRFT**	3	**YLFEVDNLTYVQLESRFT (P4)**
LANETTQALQLF	3	LICGLRQ**LANETTQALQLF**L**RA TTELRTFSILNRKAIDF**	9	**LANETTQALQLF (P5)**
RATTELRTFSILNRKAIDFLL	6			**RATTELRTFSILNRKAIDF (P6)**

Table 2
Peptides representing the presence of different T and B cell epitopes. The epitopes in bold represent the epitopes also present in the sGP fragment "**GAFFLYDRLAST**" which is a part of P2.

Peptide fragments	$CD8^+$T cell epitopes	$CD4^+$T cell epitopes	B-cell epitopes
FKRTSFFLWVIILFQRTFSIPL (P1)	KRTSFFLWV RTSFFLWVI FFLWVIILF VIILFQRTF ILFQRTFSI FQRTFSIPL	WVIILFQRT VIILFQRTF ILFQRTFSI FQRTFSIPL IILFQRTFS FLWVIILFQ FFLWVIILF LWVIILFQR FKRTSFFLW	RTSFFLWVII SFFLWVIILF LWVIILFQRT LFQRTFSIPL
FHKEGAFFLYDRLASTVIYRGTTFAEG (P2)	FHKEGAFFL HKEGAFFLY **FLYDRLAST** DRLASTVIY RLASTVIYR TVIYRGTTF	FHKEGAFFL YDRLASTVI IYRGTTFAE YRGTTFAEG **FFLYDRLAS** **FLYDRLAST** VIYRGTTFA	HKEGAFFLYD **GAFFLYDRLA** VIYRGTTFAE LASTVIYRGT ASTVIYRGTT FHKEGAFFLY STVIYRGTTF LYDRLASTVI RLASTVIYRG
LANETTQALQLF (P5)	LANETTQAL NETTQALQL ETTQALQLF	LANETTQAL	NETTQALQLF
RATTELRTFSILNRKAIDF (P6)	RATTELRTF TELRTFSIL LRTFSILNR RTFSILNRK FSILNRKAI	FSILNRKAI LRTFSILNR ILNRKAIDF	ELRTFSILNR SILNRKAIDF RATTELRTFS LRTFSILNRK RTFSILNRKA TTELRTFSIL FSILNRKAID

with different HLA molecules. Class I HLA binds to shorter peptides as the binding groove is closed while class II accommodates larger peptide owing to open ended grooves. Considering this difference in the binding pockets of HLA molecules, twenty nonamer CD8 + T cell epitopes which are a part of considered peptides and the four selected peptides (Table 2) were docked with HLA class I and class II alleles respectively. The length of native peptides (naturally bound peptides separated from their respective HLA crystal structure) was 8-11 and 9-20 amino acids for HLA class I and II respectively which were re-docked with their respective HLA to obtain a standard for comparison.

For HLA class I, average binding energy and RMSD value of $CD8^+$ T cell epitopes which are a part of each peptide were taken for comparison with native peptides. The average binding energy of the four peptides with HLA class I was found to be comparable with native peptide binding energy (Fig. 1(a) and (b)). The average binding energies considering all HLA class I alleles were −8.1 (P1), −8.1 (P2), −7.6 (P5), −7.5 (P6) and −9.3 (native peptides). In majority of HLA class I-peptide interactions, the RMSD value was found to be comparable with native peptides with variation in some cases (Fig. 1(c) and (d)). The average RMSD values were 2.0 (P1), 2.5 (P2), 2.5 (P5), 2.1 (P6) and 1.9 (native peptides) which are less than 3 representing high quality predictions.

Docking for HLA class II molecules was carried out by taking the four identified common peptides (P1, P2, P5 and P6). The binding energy for all peptides with HLA class II was found to be similar with native peptide binding energy (Fig. 2(a) and (b)). Tellingly, positive binding energy of P2 indicated highly unstable interactions with HLA-DRA1 and hence, it was not plotted in Fig. 2(a). The average binding energies considering all HLA class II molecules were −5.4 (P1), −5.5 (P2), −7.1 (P5), −5.7 (P6) and −6.9 (native peptides). The RMSD values of peptide P1 (DQ0602, DQ8, DR1, DRB1*0101, DRB1*1501 and DR4), P5 (DPA1, DQ8, DR1, DRB1*0101and DRB1*1501) and P6 (DPA1, DQ0602 and DR5) were found to be better than native peptide RMSD values (Fig. 2(c) and (d)). RMSD values of peptides for some HLA class II molecules were not plotted as either the value was found to be >5 or peptides did not bind in the binding cavity. The average RMSD values were 1.4 (P1), 3.4 (P2), 2.0 (P5), 2.6 (P6) and 2.4 (native peptides). Except P2, three peptides have shown good quality predictions as RMSD value is less than 3.

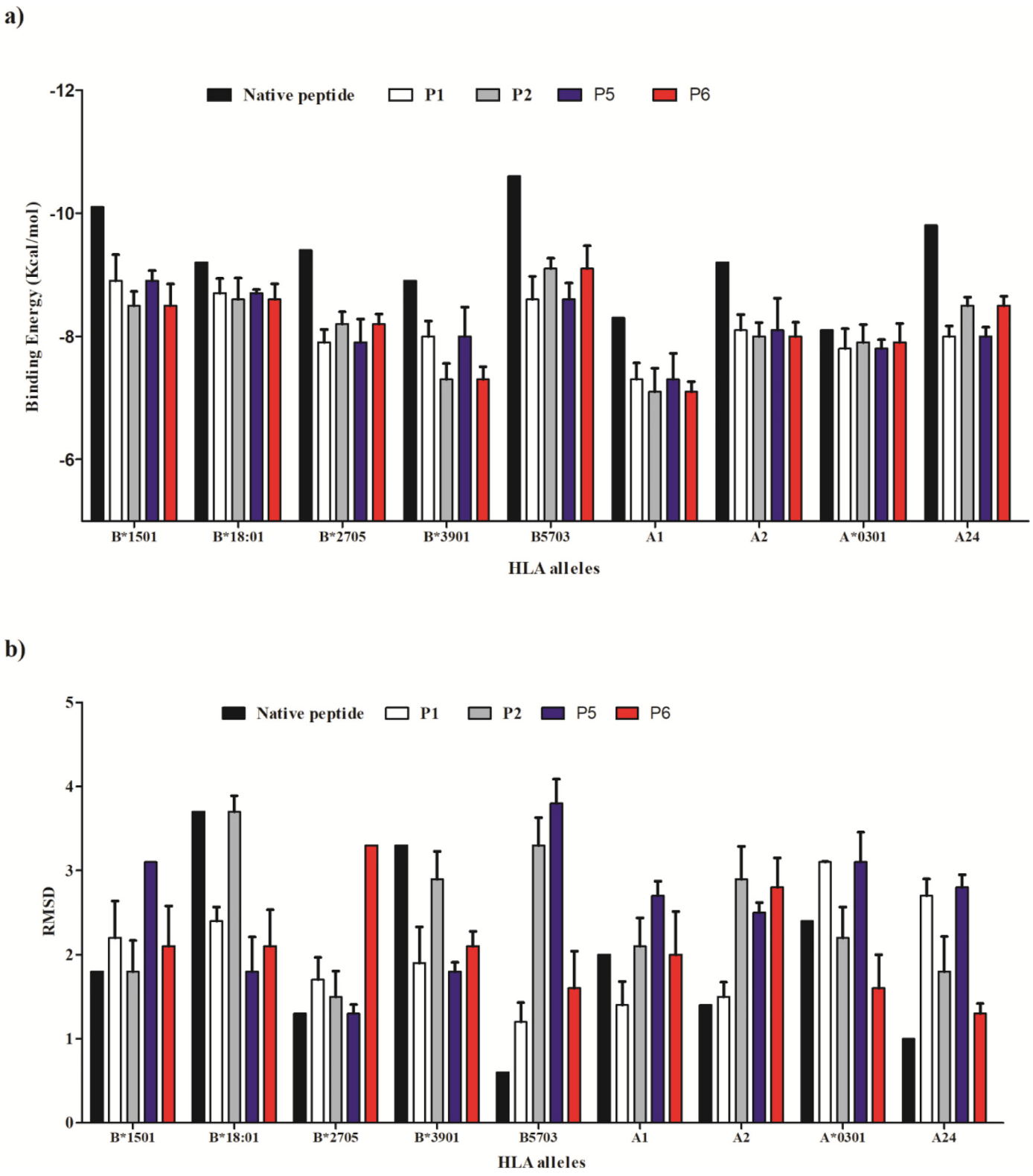


Fig. 1. (a) Binding energy (Autodock Vina) and (b) RMSD (CABS-dock) of the native peptide (control) and Ebola peptides for HLA class I. For all four peptides, the mean binding energy and RMSD of epitopes belonging to respective peptide were considered and hence, standard error of mean (SEM) has been plotted. Native peptides represent the peptide that already existed in the crystallographic structures of the HLA molecules. They were separated and docked with their respective HLA molecule.

Prediction tools have shown that these peptides were predicted to bind with large number of HLA-A, HLA-B, HLA-DP, HLA-DQ and HLA-DR alleles (Table 3). All HLA alleles predicted by the predictions tools have been mentioned in supplementary Table 5. Population coverage analysis (an indication of the fraction of individuals expected to respond to a given epitope set) for different continents was carried out for the identified Ebola peptides by taking all the HLA alleles (supplementary Table 5). The results showed that the expected response of these peptides is in the range of 77%–100% for populations residing in different continents (Fig. 3.). Population

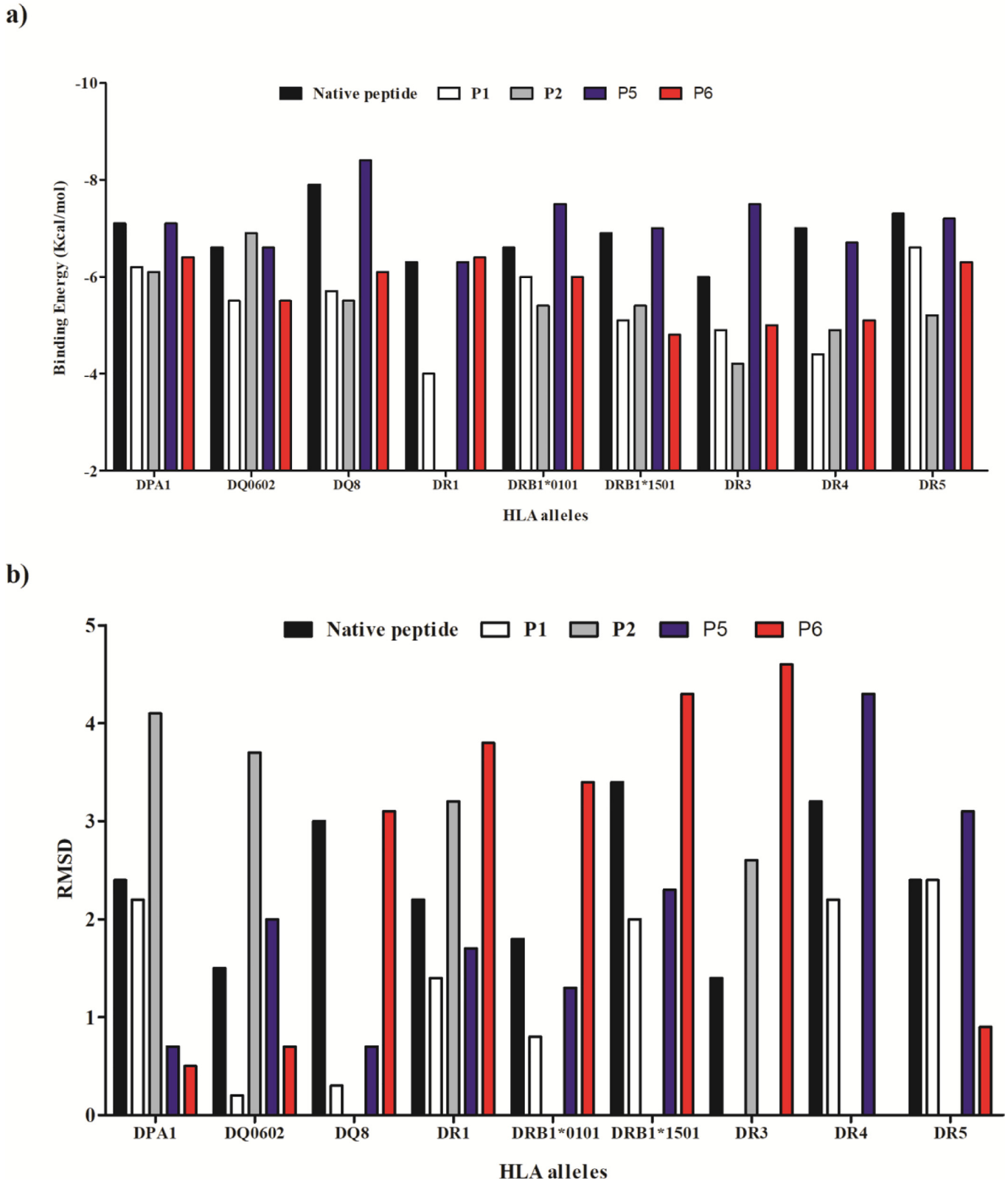


Fig. 2. (a) Binding energy (Autodock Vina) and (b) RMSD (CABS-dock) of the native peptide (control) and Ebola peptides for HLA class II. The four identified peptides were docked with each HLA II molecule. Native peptides represent the peptide that already existed in the crystallographic structures of the HLA molecules. They were separated and docked with their respective HLA molecule. P2 showed highly unstable interactions (positive binding energy) with HLA-DRA1, hence it was not plotted in Fig. 2(a). RMSD value of some peptides was found to be >5 representing poor quality prediction and hence, they were not plotted in Fig. 2(b).

Table 3
Peptides containing multiple epitopes binding to diverse HLA alleles.

Peptide fragment	HLA alleles covered				
	Class I		Class II		
	HLA-A	HLA-B	HLA DP	HLA DQ	HLA DR
FKRTSFFLWVIILFQRTFSIPL (P1)	100	75	11	20	121
FHKEGAFFLYDRLASTVIYRGTTFAEG (P2)	115	74	10	18	120
LANETTQALQLF (P5)	17	75	4	11	3
RATTELRTFSILNRKAIDF (P6)	7	72	2	7	63

Table 4
Peptide conservation in various Ebola virus species. The amino acid variation of identified peptides among different species has been highlighted in bold and underlined.

Ebola species	FKRTSFFLWVIILFQRTFSIPL (P1)	FHKEGAFFLYDRLASTVIYRGTTFAEG (P2)	LANETTQALQLF (P5)	RATTELRTFSILNRKAIDF (P6)
ZAIRE	99.39%	100%	100%	100%
SUDAN	F**RKS**SFF**V**WVIILFQ**KA**FS**M**PL	FHK**D**GAFFLYDRLASTVIYRG**VN**FAEG	100%	RATTELRT**YT**ILNRKAIDF
TAI FOREST	F**RK**TSFF**V**WVIILF**HKV**FSIPL	FHKEGAFFLYDRLAST**I**IYRGTTFAEG	100%	100%
BUNDIBUGYO	F**RK**TSFF**V**WVIILF**HKV**F**P**IPL	FHKEGAFFLYDRLAST**I**IYR**S**TTF**S**EG	100%	100%

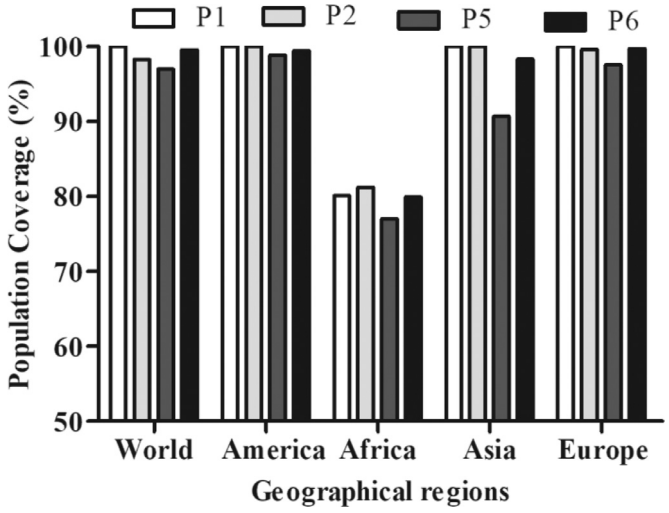


Fig. 3. Population coverage of the identified peptides in whole world and four different continents. Bars indicate the percentage of population coverage obtained from IEDB population coverage analysis tool. Population coverage represents the fraction of individuals expected to respond to a given epitope set.

coverage for the whole world was found to be 96%–100% (Fig. 3.). These computational results merely indicate the probability of the potential candidates to generate immune response in different geographical regions but exact immune response needs to be validated *in vitro/in vivo.*

3.4. Conservation analysis of peptides amongst different Ebola virus species

Each selected peptide was looked for conservancy in all considered glycoprotein sequences of Ebola virus species infecting humans (Zaire, Sudan, Bundibugyo and Tai Forest). All peptides showed more than 99% conservancy amongst Zaire Ebola glycoprotein sequences (Table 4). P5 was found to be 100% conserved in all Ebola virus species infecting humans. P6 was found to be 100% conserved in all Ebola virus species except Sudan. P1 and P2 are mostly conserved except for a few variations in different species (Table 4).

3.5. Mapping of peptide fragments

As the protein has a trimeric structure, only monomer was selected from the downloaded PDB structure with the help of Discovery Studio Visualizer tool v4.1 to highlight selected peptide fragments enriched with multiple T and B-cell epitopes (Fig. 4.). In the PDB structure of glycoprotein, 1-31 residues were missing so P1 could not be mapped. Further, the location of all the peptides was schematically shown in different domains of glycoprotein. Peptide P1 is a part of the signal peptide, P2 is a part of GP1 domain whereas P5 and P6 are parts of GP2 domain (Fig. 4.).

4. Discussion

Peptide as a choice of vaccine candidate has made remarkable headway in vaccine development against virus, bacteria and cancer. Peptides can be a recombinant or synthetic construct of epitopes which are targeted against cell surface or intracellular proteins (Kumai et al., 2017). GMP grade peptides are easy to synthesize and cost effective as compared to whole organism vaccine or recombinant proteins. Although peptide vaccines are still not translated into clinics, the promising results have led to different stages of clinical trials. A multimeric recombinant peptide consisting of nine conserved B and T cell epitopes from hemagglutinin, nucleoprotein and matrix 1 protein has been produced by BiondVax Pharmaceuticals Ltd. (Israel). It has shown the ability to induce cell mediated and humoral immunity and is presently in clinical trials phase II (van Doorn et al., 2017; Atsmon et al., 2014). Vacc-4x and Vacc-C5 are peptide based vaccines against HIV infections and have shown encouraging results in phase II clinical study (Huang et al., 2018; Brekke et al., 2017). Peptide based vaccines have also been developed against different types of cancer such as gastric cancer, colorectal cancer and myeloid leukemia and are either in phase I or phase II clinical trials (Maslak et al., 2018; Sundar et al., 2018; Taniguchi et al., 2017). Identification of immunogenic peptide regions (epitopes) solely based on experimental approach is time consuming and expensive. Rapid development of computational approaches has made it easier to screen potential epitopes, thus, reducing the number of epitopes requiring validation based on *in vitro* and *in vivo* experiments. In the present study, four highly conserved peptides containing multiple epitopes of Ebola virus glycoprotein were identified based on different *in silico* approaches that include multiple epitope prediction algorithms and docking tools.

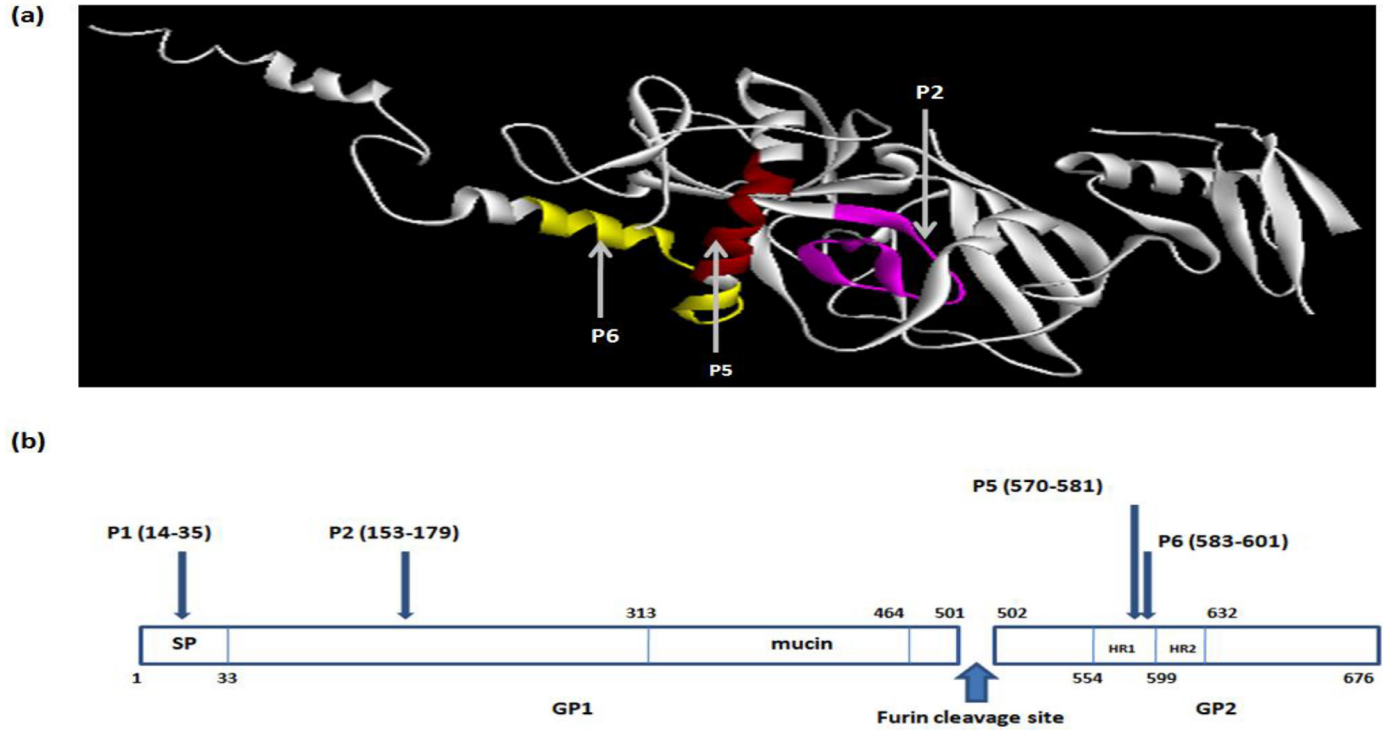


Fig. 4. (a) Mapping of identified peptide fragments onto the crystal structures of Ebola Glycoprotein (PDB id: 5JQ3). P2, P5 and P6 are depicted in pink, red and yellow respectively. P1 could not me located in the structure owing to missing residues. (b) Schematic presentation of peptides in different regions of Ebola glycoprotein. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

A number of epitopes have been reported based on different computational prediction algorithms for influenza virus, adenovirus, papilloma virus, tuberculosis and cancer (Hossain et al., 2018; Dhiman et al., 2016; Kumar et al., 2015; Lohia and Baranwal, 2015; Monterrubio-Lopez and Ribas-Aparicio, 2015). Several prediction algorithms have been developed for mapping potential epitopes (Patronov and Doytchinova, 2013; Ponomarenko and Van Regenmortel, 2009). A consensus approach where multiple epitope prediction tools are applied is more effective as different tools allow consideration of multiple algorithms, immunological factors and databases (Lohia and Baranwal, 2014). In the current study, peptides containing multiple $CD4^+$ and $CD8^+$ epitopes with a potential to induce both helper and cytotoxic T cell response were identified by employing six different (three for each $CD4^+$ and $CD8^+$ T cell epitopes) immunoinformatics tools.

HLA class II binding groove is open ended and accommodates larger and flexible peptides as compared to closed HLA class I binding groove (Andreatta and Nielsen, 2016). Hence, the prediction tools of HLA class II are not as accurate as those of HLA class I. X-ray studies have shown that the binding groove of HLA class II is occupied by large peptides in which the core peptide is of nine amino acids (also called peptide registers) and rest of amino acids extend to different sides of groove (Dimitrov et al., 2010). There is existence of multiple peptide registers and as a result, long peptides adopt different conformations which contribute to the poor prediction accuracy of class II binding epitopes. However, class II predicted epitopes have shown encouraging results in *in vitro* or *in vivo* system (Kovjazin et al., 2013; Burgosa et al., 2010; Brintnell et al., 2007). Further, larger peptides were found to elicit both, HLA class I and II immune response in different studies (Krug et al., 2010; Maslak et al., 2010). In one of our earlier studies, peptides containing multiple $CD8^+$ and $CD4^+$T cell epitopes have shown induced proliferation and IFN-γ secretion in peripheral blood mononuclear cells (Lohia and Baranwal, 2017).

Most of the B cell epitopes are discontinuous and nearly 10% epitopes are linear (Wang et al., 2011). *In silico* prediction tools for B cell epitopes (mainly linear B cell epitopes) (Ponomarenko and Van Regenmortel, 2009) have been developed but are not as reliable as T cell epitope prediction tools. In one recent study on malarial antigen, some epitopes predicted by ABCpred were found to match with experimentally identified epitopes in immune rabbit antiserum (Bergmann-Leitner et al., 2013). Hence, in the present study, identified peptides containing multiple T cell epitopes were looked for the presence of B cell epitopes and four out of six peptides were found to contain B cell epitopes.

Development of autoimmune diseases such as Gullian Barr syndrome (influenza and oral polio vaccine), systemic lupus anaemia and mysthemia gravis (hepatitis B), diabetes mellitus (Mumps) and multiple sclerosis (swine flu) after vaccination has been reported in several studies (Cohen and Shoenfeld, 1996). Hence, it becomes critical to eliminate the vaccine candidates which resemble human proteome to avoid autoimmunity. BLAST results reveal that one peptide fragment (YLFEVDNLTYVQLESRFT) resembles the human cytoskeletal linker protein epiplakin1 and hence, it was eliminated from further studies. Urticarial, vasculitis, bronchospasm and gastrointestinal disorders are some of the clinical manifestations associated with hypersensitivity reported after vaccination (Fritsche et al., 2010). Current *in silico* analysis have shown that none of the selected peptides have a chance to induce either allergic or toxic response.

Peptide protein interaction has gained attention in recent days that resulted in advancement of docking approaches focused on peptide-protein docking. It is significant to find peptide vaccine candidates which can be presented by diverse array of HLA molecules so as to have efficacy in different populations across the globe. Nine protein structures each of HLA class I and II obtained from PDB and two docking tools (Autodock Vina and CABS-dock) were employed in the current study. Docking studies reveal that

selected peptides and epitopes have comparable binding energy and RMSD with native peptide. In our previous studies, peptides containing multiple T cell epitopes of H1N1 influenza virus matrix protein and metadherin protein (tumor antigen) have shown good binding potential with different HLA class I and II alleles (Dhiman et al., 2016; Lohia and Baranwal, 2015). In other study on human baculo virus, selected epitopes of capsid protein have shown strong binding with HLA class I alleles using Autodock Vina (Kalyanaraman, 2018). In addition to Autodock Vina, CABS-dock was used to get better perspective for HLA-peptide interactions. Average RMSD of all peptides except P2 (HLA class II) was found to be less than 3 indicating highly favorable binding interactions with selected HLA molecules. Although there was unstable interaction of P2 with some HLA class II alleles but it was found to have comparable binding energy and RMSD value to native peptides for majority of HLA class I alleles. P2 is a longer peptide and therefore, may adopt different conformations in open binding groove of HLA class II which could be the reason of unstable interactions. P2 is an interesting candidate also because it was found to contain a peptide fragment (GAFFLYDRLAST) common with soluble glycoprotein (sGP). Further, all peptides were predicted to bind to a large number of class I and II HLA alleles. The expected immune response of all these peptides in different continents (Africa, America, Asia and Europe) was found to be noteworthy (range of population coverage was 77%–100%) indicating that these peptides may be effective vaccine candidates for different populations across the world.

On performing a comparative analysis with the help of IEDB to elucidate previous reports, it was found that the exact sequence of these peptides was not reported for T and B cell response in the *in vitro* or *in vivo* system. A fragment (LYDRLASTV) of P2 was reported to be capable of stimulating splenocytes in mice to secrete interferon-gamma (Wu et al., 2012) which supports the potential of P2 to evoke immune response. LANETTQALQLF (P5) was found to be part of a large peptide (TEGLMHNQDGLICGLRQLANETTQALQLFLRATTELRT) which was reported to react with antisera obtained after vaccinating the Ebola DNA vaccine (Ripoll et al., 2017). TELRTFSI peptide which is a part of P6 (RATTELRTFSILNRKAIDF) has shown cytotoxic T lymphocytes response in mice model (Rao et al., 1999). Hence, these studies support the potential of selected peptides to induce T and B cell response.

Further, these peptides were found to be highly conserved in most of Ebola virus strains infecting humans. Interestingly, one peptide was 100% conserved in all four species of Ebola virus. Additionally, these peptides were found to be parts of crucial domains of glycoprotein. Peptide P1 formed a part of signal peptide at the glycoprotein N terminus which directs the nascent glycoprotein to the endoplasmic reticulum where carbohydrates are added on its surface (Feldmann et al., 1994; Lee et al., 2008). A thick layer of oligosaccharide coats majority of the GP trimer (Lee et al., 2008) enabling formation of a shield that protects it from humoral immune responses. It is similar to glycan shields of HIV-1 gp120 (Kwong et al., 1998; Wyatt et al., 1998) and Epstein-Barr virus gp350 (Szakonyi et al., 2006) thus, indicating a common immune evasion method. Peptides P5 and P6 were found to be located in the heptad repeat regions (HR1 and HR2) of GP2. HRI can be segmented into HR1A, HR1B, HR1C and HR1D. HR1A and HR1B (residues 554-575) form an alpha helix. The bend between HR1A and HR1B contains an atypical stutter similar to that in parainfluenza virus (Yin et al., 2006), which may act as a conformational switch (Weissenhorn et al., 1998). Cys511 and Cys556 form a highly conserved disulfide bond similar to the two cysteine residues bordering the internal fusion loop in avian sarcoma leukosis virus (Delos et al., 2008). HR1C (residues 576-582) forms an extended coil linking to HR1D (residues 583-598) (Lee et al., 2008).

In an interesting study, a cocktail of therapeutic monoclonal antibodies (MAbs) was used to find the critical residue of GP responsible for interacting with these MAbs (Davidson et al., 2015). It was reported that these MAbs binds to different residues which are present in base of GP, glycan cap or mucin like domain. The identified peptides of current study are different from these experimentally identified epitopes and they are located in different regions of GP protein. These epitopes are highly conserved and have potential to generate cross protective immunity against different strains of EBOV. The identified epitopes are projecting majorly to T cell epitopes and continuous B cell epitopes. It is proposed to use the cocktails of these four peptides for generating cell mediated and humoral immune response.

More than 15 vaccine candidates are in preclinical development while eight vaccine candidates are dispersed in various stages of clinical tests. Two DNA Ebola vaccines under consideration are a vaccine encoding a transmembrane-deleted GP sequence from the Zaire and Sudan strain (SUDV) along with the NP from the Zaire strain of EBOV and a second DNA vaccine containing two DNA plasmids in equal proportion expressing the WT GP of EBOV and SUDV (Sridhar, 2015). A multivalent filovirus vaccine namely Modified Vaccinia Ankara (MVA) that encodes the glycoproteins from Zaire, Sudan and Marburg virus and a nucleoprotein from Tai Forest Ebola species is under clinical trials. All these efforts to develop a vaccine against Ebola virus are a long way from being translated into practical applications. Thus, these peptides having the potential to induce T and B cell response may be considered in designing synthetic vaccine against the Ebola virus.

5. Conclusion

Four peptides containing multiple T and B cell epitopes of Ebola glycoprotein which are highly conserved in different Ebola strains infecting humans were identified. These peptides were found to be predicted for a large number of HLA molecules and have shown strong binding interactions with 18 HLA molecules. Hence, these peptides are proposed as candidates for synthetic vaccine design against Ebola virus and need to be validated experimentally *in vitro* or *in vivo*.

Disclosure and funding

This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors.

Declarations of interest

None.

Author's contribution

Dr. Manoj Baranwal was involved in design of study, results analysis and writing the manuscript. Sahil Jain carried out all the work and helped in writing the manuscript.

Acknowledgement

I express my sincere thanks to the scientific community for developing *in silico* tools.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jtbi.2019.01.016.

References

Andreatta, M., Nielsen, M., 2016. Gapped sequence alignment using artificial neural networks: Application to the MHC class I system. Bioinformatics 32 (4), 511–517.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215 (3), 403–410.

Atsmon, J., Caraco, Y., Ziv-Sefer, S., Shaikevich, D., Abramov, E., Volokhov, I., Bruzil, S., Haima, K.Y., Gottlieb, T., Ben-Yedidia, T., 2014. Priming by a novel universal influenza vaccine (multimeric-001)—a gateway for improving immune response in the elderly population. Vaccine 32 (44), 5816–5823.

Bergmann-Leitner, E.S., Chaudhury, S., Steers, N.J., Sabato, M., Delvecchio, V., Wallqvist, A.S., Ockenhouse, C.F., Angov, E., 2013. Computational and experimental validation of B and T-cell epitopes of the *in vivo* immune response to a novel malarial antigen. PLoS One 8 (8), e71610.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., Bourne, P.E., 2000. The protein data bank. Nucleic Acids Res. 28 (1), 235–242.

Bhasin, M., Raghava, GP., 2004. SVM based method for predicting HLA-DRB1*0401 binding peptides in an antigen sequence. Bioinformatics 20 (3), 421–423.

Blaszczyk, M., Kurcinski, M., Kouza, M., Wieteska, L., Debinski, A., Kolinski, A., Kmiecik, S., 2016. Modeling of protein–peptide interactions using the CABS-dock web server for binding site search and flexible docking. Methods 93, 72–83.

Brauburger, K., Boehmann, Y., Tsuda, Y., Hoenen, T., Olejnik, J., Schumann, M., Ebihara, H., Muhlberger, E., 2014. Analysis of the highly diverse gene borders in Ebola virus reveals a distinct mechanism of transcriptional regulation. J. Virol. 88 (21), 12558–12571.

Brekke, K., Sommerfelt, M., Ökvist, M., Dyrhol-Riise, A.M., Kvale, D., 2017. The therapeutic HIV Env C5/gp41 vaccine candidate Vacc-C5 induces specific T cell regulation in a phase I/II clinical study. BMC Infect. Dis. 17 (1), 228.

Brintnell, W., Bell, D.A., Hill, J.A., Jevnikar, A.M., Sette, A., Sidney, J., Doege, K., Cairns, E., 2007. The influence of MHC class II molecules containing the rheumatoid arthritis shared epitope on the immune response to aggrecan G1 and its peptides. Scand. J. Immunol. 65 (5), 444–452.

Sánchez-Burgos, G., Ramos-Castañeda, J., Cedillo-Rivera, R., Dumonteil, E., 2010. Immunogenicity of novel Dengue virus epitopes identified by bioinformatic analysis. Virus Res. 153 (1), 113–120.

Cherpillod, P., Schibler, M., Vieille, G., Cordey, S., Mamin, A., Vetter, P., Kaiser, L., 2016. Ebola virus disease diagnosis by real-time RT-PCR: A comparative study of 11 different procedures. J. Clin. Virol. 77, 9–14.

Cohen, A.D., Shoenfeld, Y., 1996. Vaccine-induced autoimmunity. J. Autoimmun. 9 (6), 699–703.

Davidson, E., Bryan, C., Fong, R.H., Barnes, T., Pfaff, J.M., Mabila, M., Rucker, J.B., Doranz, B.J., 2015. Mechanism of binding to Ebola Virus glycoprotein by the ZMapp, ZMAb, and MB-003 cocktail antibodies. J Virol. 89 (21), 10982–10992.

Delos, S.E., Brecher, M.B., Chen, Z., Melder, D.C., Federspiel, M.J., White, J.M., 2008. Cysteines flanking the internal fusion peptide are required for the avian sarcoma/ leukosis virus glycoprotein to mediate the lipid mixing stage of fusion with high efficiency. J. Virol. 82 (6), 3131–3134.

Dimitrov, I., Garnev, P., Flower, D.R., Doytchinova, I., 2010. MHC class II binding prediction-A little help from a friend. J. Biomed. Biotechnol. 2010, 705821.

Dhiman, G., Lohia, N., Jain, S., Baranwal, M., 2016. Metadherin peptides containing $CD4^+$ and $CD8^+$ T cell epitopes as a therapeutic vaccine candidate against cancer. Microbiol. Immunol. 60 (9), 646–652.

van Doorn, E., Liu, H., Ben-Yedidia, T., Hassin, S., Visontai, I., Norley, S., Frijlink, H.W., Hak, E., 2017. Evaluating the immunogenicity and safety of a BiondVax-developed universal influenza vaccine (Multimeric-001) either as a standalone vaccine or as a primer to H5N1 influenza vaccine: Phase IIb study protocol. Medicine (Baltimore) 96 (11), e6339.

Duvvuri, V.R., Duvvuri, B., Jamnik, V., Gubbay, J.B., Wu, J., Wu, G.E., 2013. T Cell memory to evolutionarily conserved and shared hemagglutinin epitopes of H1N1 viruses: A pilot scale study. BMC Infect. Dis. 13, 204.

Edgar, R.C., 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32 (5), 1792–1797.

Feldmann, H., Klenk, H.D., Sanchez, A., 1993. Molecular biology and evolution of filoviruses. Arch. Virol. Suppl. 7, 81–100.

Feldmann, H., Nichol, S.T., Klenk, H.D., Peters, C.J., Sanchez, A., 1994. Characterization of filoviruses based on differences in structure and antigenicity of the virion glycoprotein. Virology 199 (2), 469–473.

Fritsche, P.J., Helbling, A., Ballmer-Weber, B.K., 2010. Vaccine hypersensitivity—update and overview. Swiss Med. Wkly 140 (17–18), 238–246.

Groseth, A., Marzi, A., Hoenen, T., Herwig, A., Gardner, D., Becker, S., Ebihara, H., Feldmann, H., 2012. The Ebola virus glycoprotein contributes to but is not sufficient for virulence i*n vivo*. PLoS Pathog. 8 (8), e1002847.

Gupta, S., Kapoor, P., Chaudhary, K., Gautam, A., Kumar, R., Raghava, G.P., 2013. In silico approach for predicting toxicity of peptides and proteins. PLoS One 8 (9), e73957.

Hoenen, T., Groseth, A., Feldmann, H., 2012. Current Ebola vaccines. Expert Opin. Biol. Ther. 12 (7), 859–872.

Hossain, R., Yasmin, T., Hosen, M.I., Nabi, A.H.M.N., 2018. In silico identification of potential epitopes present in human adenovirus proteins for vaccine design and of putative drugs for treatment against viral infection. J. Immunol. Methods 455, 55–70.

Huang, Y., Zhang, L., Jolliffe, D., Sanchez, B., Stjernholm, G., Jelmert, Ø., Ökvist, M., Sommerfelt, M.A., 2018. Postvaccination C-Reactive protein and C5/gp41732–744 antibody level fold-changes over baseline are independent predictors of therapeutic HIV vaccine effect in a phase 2 clinical study of Vacc-4x. AIDS Res. Hum. Retroviruses 34 (3), 307–313.

Ito, H., Watanabe, S., Takada, A., Kawaoka, Y., 2001. Ebola virus glycoprotein: Proteolytic processing, acylation, cell tropism, and detection of neutralizing antibodies. J. Virol. 75 (3), 1576–1580.

Iwasa, A., Shimojima, M., Kawaoka, Y., 2011. sGP serves as a structural protein in Ebola virus infection. J. Infect. Dis. 204 (3), S897–S903.

Jeffers, S.A., Sanders, D.A., Sanchez, A., 2002. Covalent modifications of the Ebola virus glycoprotein. J. Virol. 76 (24), 12463–12472.

Kalyanaraman, N., 2018. In silico prediction of potential vaccine candidates on capsid protein of human bocavirus 1. Mol. Immunol. 93, 193–205.

Kovjazin, R., Shitrit, D., Preiss, R., Haim, I., Triezer, L., Fuks, L., Nader, A.R., Raz, M., Bardenstein, R., Horn, G., Smorodinsky, N.I., Carmon, L., 2013. Characterization of novel multiantigenic vaccine candidates with pan-HLA coverage against Mycobacterium tuberculosis. Clin. Vaccine Immunol. 20 (3), 328–340.

Krug, L.M., Dao, T., Brown, A.B., Maslak, P., Travis, W., Bekele, S., Korontsvit, T., Zakhaleva, V., Wolchok, J., Yuan, J., Li, H., Tyson, L., Scheinberg, D.A., 2010. WT1 peptide vaccinations induce CD4 and CD8 T cell immune responses in patients with mesothelioma and non-small cell lung cancer. Cancer Immunol. Immunother 59 (10), 1467–1479.

Kuhn, J.H., Radoshitzky, S.R., Guth, A.C., Warfield, K.L., Li, W., Vincent, M.J., Towner, J.S., Nichol, S.T., Bavari, S., Choe, H., Aman, M.J., Farzan, M., 2006. Conserved receptor-binding domains of Lake Victoria Marburgvirus and Zaire Ebolavirus bind a common receptor. J. Biol. Chem. 281 (23), 15951–15958.

Kumai, T., Kobayashi, H., Harabuchi, Y., Celis, E., 2017. Peptide vaccines in cancer—old concept revisited. Curr. Opin. Immunol. 45, 1–7.

Kumar, A., Yadav, I.S., Hussain, S., Das, B.C., Bharadwaj, M., 2015. Identification of immunotherapeutic epitope of E5 protein of human papillomavirus-16: An in silico approach. Biologicals 43 (5), 344–348.

Kurcinski, M., Jamroz, M., Blaszczyk, M., Kolinski, A., Kmiecik, S., 2015. CABS-dock web server for the flexible docking of peptides to proteins without prior knowledge of the binding site. Nucleic Acids Res. 43 (W1), W419–W424.

Kwong, P.D., Wyatt, R., Robinson, J., Sweet, R.W., Sodroski, J., Hendrickson, W.A., 1998. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. Nature 393 (6686), 648–659.

Larsen, M.V., Lundegaard, C., Lamberth, K., Buus, S., Lund, O., Nielsen, M., 2007. Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. BMC Bioinformatics 8, 424.

Lee, J.E., Fusco, M.L., Hessell, A.J., Oswald, W.B., Burton, D.R., Saphire, E.O., 2008. Structure of the Ebola virus glycoprotein bound to an antibody from a human survivor. Nature 454 (7201), 177–182.

Lohia, N., Baranwal, M., 2014. Conserved peptide containing overlapping CD4+ and CD8+ T cell epitopes in H1N1 influenza virus: An immunoinformatics approach. Viral Immunol. 27 (5), 225–234.

Lohia, N., Baranwal, M., 2015. Identification of conserved peptides comprising multiple T cell epitopes of Matrix 1 protein in H1N1 influenza virus. Viral Immunol. 28 (10), 570–579.

Lohia, N., Baranwal, M., 2017. Immune response of highly conserved influenza A virus matrix 1 peptides: Matrix 1 peptides for influenza vaccine. Microbiol. Immunol. 61 (6), 225–231.

Lund, O., Nielsen, M., Kesmir, C., Petersen, A.G., Lundegaard, C., Worning, P., Sylvester-Hvid, C., Lamberth, K., Røder, G., Justesen, S., Buus, S., Brunak, S., 2004. Definition of supertypes for HLA molecules using clustering of specificity matrices. Immunogenetics 55 (12), 797–810.

Madara, J.J., Han, Z., Ruthel, G., Freedman, B.D., Harty, R.N., 2015. The multifunctional Ebola virus VP40 matrix protein is a promising therapeutic target. Future Virol. 10 (5), 537–546.

Manicassamy, B., Wang, J., Jiang, H., Rong, L., 2005. Comprehensive analysis of Ebola virus GP1 in viral entry. J. Virol. 79 (8), 4793–4805.

Maslak, P.G., Dao, T., Bernal, Y., Chanel, S.M., Zhang, R., Frattini, M., Rosenblat, T., Jurcic, J.G., Brentjens, R.J., Arcila, M.E., Rampal, R., Park, J.H., Douer, D., Katz, L., Sarlis, N., Tallman, M.S., Scheinberg, D.A., 2018. Phase 2 trial of a multivalent WT1 peptide vaccine (galinpepimut-S) in acute myeloid leukemia. Blood Adv. 2 (3), 224–234.

Maslak, P.G., Dao, T., Krug, L.M., Chanel, S., Korontsvit, T., Zakhaleva, V., Zhang, R., Wolchok, J.D., Yuan, J., Pinilla-Ibarz, J., Berman, E., Weiss, M., Jurcic, J., Frattini, M.G., Scheinberg, D.A., 2010. Vaccination with synthetic analog peptides derived from WT1 oncoprotein induces T-cell responses in patients with complete remission from acute myeloid leukemia. Blood 116 (2), 171–179.

Miotto, O., Heiny, A., Tan, T.W., August, J.T., Brusic, V., 2008. Identification of human-to-human transmissibility factors in PB2 proteins of influenza A by large scale mutual information analysis. BMC Bioinformatics 9 (1), 18.

Mohan, G.S., Li, W., Ye, L., Compans, R.W., Yang, C., 2012. Antigenic subversion: A novel mechanism of host immune evasion by Ebola virus. PLoS Pathlog. 8 (12), e1003065.

Monterrubio-López, G.P., Jorge, A., Y-Merchand, G., Ribas-Aparicio, R.M., 2015. Identification of novel potential vaccine candidates against tuberculosis based on reverse vaccinology. BioMed Res. Int. 2015, 483150.

Nielsen, M., Lundegaard, C., Worning, P., Lauemøller, S.L., Lamberth, K., Buus, S., Brunak, S., Lund, O., 2003. Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. Protein Sci. 12 (5), 1007–1017.

Nielsen, M., Lund, O., 2009. NN-align. An artificial neural network-based alignment algorithm for MHC class II peptide binding prediction. BMC Bioinformatics 10, 296.

Oliveira, F.M., Coelho, I.E., Lopes, M.D., Taranto, A.G., Junior, M.C., Santos, L.L., Villar, J.A., Fonseca, C.T., Lopes, D.D., 2016. The use of reverse vaccinology and

molecular modeling associated with cell proliferation stimulation approach to select promiscuous epitopes from Schistosoma mansoni. Appl. Biochem. Biotechnol. 179 (6), 1023–1040.
Patronov, A., Doytchinova, I., 2013. T-cell epitope vaccine design by immunoinformatics. Open Biol. 3 (1), 120139.
Peters, C.J., LeDuc, J.W., 1999. An introduction to Ebola: The virus and the disease. J. Infect. Dis. 179 (1), 9–16.
Pleguezuelos, O., Robinson, S., Stoloff, G.A., Caparrós-Wanderley, W., 2012. Synthetic Influenza vaccine (FLU-v) stimulates cell mediated immunity in a double-blind, randomised, placebo-controlled Phase I trial. Vaccine 30 (31), 4655–4660.
Ponomarenko, J.V., Van Regenmortel, M.H., 2009. B cell epitope prediction. Struct. Bioinfo. 35, 849–879.
Rammensee, H., Bachmann, J., Emmerich, N.P., Bachor, O.A., Stevanovic, S., 1999. SYFPEITHI: Database for MHC ligands and peptide motifs. Immunogenetics 50 (3–4), 213–219.
Rao, M., Matyas, G.R., Grieder, F., Anderson, K., Jahrling, P.B., Alving, C.R., 1999. Cytotoxic T lymphocytes to Ebola Zaire virus are induced in mice by immunization with liposomes containing lipid A. Vaccine 17 (23–24), 2991–2998.
Ripoll, D.R., Mitchell, D.A., Dupuy, L.C., Wallqvist, A., Schmaljohn, C., Chaudhury, S., 2017. Combinatorial peptide-based epitope mapping from Ebola virus DNA vaccines and infections reveals residue-level determinants of antibody binding. Hum. Vaccin. Immunother 13 (12), 2953–2966.
Saha, S., Raghava, G.P.S., 2006a. AlgPred: prediction of allergenic proteins and mapping of IgE epitopes. Nucleic Acids Res. 34, W202–W209.
Saha, S., Raghava, G.P., 2006b. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. Proteins 65 (1), 40–48.
Shen, Y., Maupetit, J., Derreumaux, P., Tuffery, P., 2014. Improved PEP-FOLD approach for peptide and miniprotein structure prediction. J. Chem. Theory Comput. 10 (10), 4745–4758.
Singh, H., Raghava, G.P., 2001. ProPred: Prediction of HLA-DR binding sites. Bioinformatics 17 (12), 1236–1237.
Sridhar, S., 2015. Clinical development of Ebola vaccines. Ther. Adv. Vaccines 3 (5–6), 125–138.
Sturniolo, T., Bono, E., Ding, J., Raddrizzani, L., Tuereci, O., Sahin, U., Braxenthaler, M., Gallazzi, F., Protti, M.P., Sinigaglia, F., Hammer, J., 1999. Generation of tissue-specific and promiscuous HLA ligand databases using DNA microarrays and virtual HLA class II matrices. Nat. Biotechnol. 17 (6), 555–561.
Sundar, R., Rha, S.Y., Yamaue, H., Katsuda, M., Kono, K., Kim, H.S., Kim, C., Mimura, K., Kua, L.-F., Yong, W.P., 2018. A phase I/Ib study of OTSGC-A24 combined peptide vaccine in advanced gastric cancer. BMC Cancer 18, 332.
Sykes, C., Reisman, M., 2015. Ebola: Working toward treatments and vaccines. P T. 40 (8), 521–525.
Szakonyi, G., Klein, M.G., Hannan, J.P., Young, K.A., Ma, R.Z., Asokan, R., Holers, V.M., Chen, X.S., 2006. Structure of the Epstein-Barr virus major envelope glycoprotein. Nat. Struct. Mol. Biol. 13 (11), 996–1001.
Takada, A., Robison, C., Goto, H., Sanchez, A., Murti, K.G., Whitt, M.A., Kawaoka, Y., 1997. A system for functional analysis of Ebola virus glycoprotein. Proc. Natl. Acad. Sci. USA 94 (26), 14764–14769.
Takada, A., 2015. Ebola vaccine and treatment. Uirusu 65 (1), 61–70.
Taniguchi, H., Iwasa, S., Yamazaki, K., Yoshino, T., Kiryu, C., Naka, Y., Liew, E.L., Sakata, Y., 2017. Phase 1 study of OCV-C02, a peptide vaccine consisting of two peptide epitopes for refractory metastatic colorectal cancer. Cancer Sci. 108 (5), 1013–1021.
Trott, O., Olson, A.J., 2010. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J. Comput. Chem. 31 (2), 455–461.
Volchkov, V.E., Feldmann, H., Volchkova, V.A., Klenk, H.D., 1998. Processing of the Ebola virus glycoprotein by the proprotein convertase furin. Proc. Natl. Acad. Sci. USA 95 (10), 5762–5767.
Wang, Y., Wu, W., Negre, N.N., White, K.P., Li, C., Shah, P.K., 2011. Determinants of antigenicity and specificity in immune response for protein sequences. BMC Bioinformatics 12, 251.
Weissenhorn, W., Carfi, A., Lee, K.H., Skehel, J.J., Wiley, D.C., 1998. Crystal structure of the Ebola virus membrane fusion subunit, GP2, from the envelope glycoprotein ectodomain. Mol. Cell 2 (5), 605–616.
Wu, S., Yu, T., Song, X., Yi, S., Hou, L., Chen, W., 2012. Prediction and identification of mouse cytotoxic T lymphocyte epitopes in Ebola virus glycoproteins. Virol. J. 9, 111.
Wyatt, R., Kwong, P.D., Desjardins, E., Sweet, R.W., Robinson, J., Hendrickson, W.A., Sodroski, J.G., 1998. The antigenic structure of the HIV gp120 envelope glycoprotein. Nature 393 (6686), 705–711.
Yin, H.S., Wen, X., Paterson, R.G., Lamb, R.A., Jardetzky, T.S., 2006. Structure of the parainfluenza virus 5 F protein in its metastable, prefusion conformation. Nature 439 (7072), 38–44.

Medical Microbiology and Immunology
https://doi.org/10.1007/s00430-019-00584-y

ORIGINAL INVESTIGATION

Conserved peptide vaccine candidates containing multiple Ebola nucleoprotein epitopes display interactions with diverse HLA molecules

Sahil Jain[1] · **Manoj Baranwal**[1]

Received: 18 September 2018 / Accepted: 11 February 2019

Abstract
Immunoinformatics has come by leaps and bounds to finding potent vaccine candidates against various pathogens. In the current study, a combination of different T ($CD4^+$ and $CD8^+$) and B cell epitope prediction tools was applied to find peptides containing multiple epitopes against Ebola nucleoprotein (NP) and the presentation of peptides to human leukocyte antigen (HLA) molecules was analyzed by prediction, docking and population coverage tools. Further, potential peptides were analyzed by ELISA for peptide induced IFN-γ secretion in peripheral blood mononuclear cells isolated from healthy volunteers. Six peptides were obtained after merging the overlapping multiple HLA I ($CD8^+$) and II ($CD4^+$) restricted T cell epitopes as well as B cell epitopes and eliminating the peptides liable to generate autoimmune and allergic response. All peptides displayed 100% conservancy in *Zaire ebolavirus*. In other Ebola virus species (Sudan, Bundibugyo and Taï forest) and *Filoviridae* members (*Lloviuvirus* and *Margburgvirus*), some peptides were found to be conserved with minor variations. Prediction tools confirmed the ability of predicted peptides to bind with diverse HLA (HLA-A, HLA-B, HLA-DP, HLA-DQ and HLA-DR) alleles. CABS-dock results displayed that the average root mean square deviation (RMSD) value was less than three in majority of cases representing strong binding affinity with HLA alleles. Population coverage analysis predicted high coverage (> 85%) for expected immune response in four continents (Africa, America, Asia and Europe). Nine out of ten blood samples exhibited enhanced IFN-γ secretion for two peptides (P2 and P3). Thus, the identified NP peptides can be considered as potential synthetic vaccine candidates against Ebola virus.

Keywords Epitope-based vaccine · Conservation analysis · HLA alleles · Molecular docking · Ebola nucleoprotein

Introduction

Ebolavirus (EBOV) is a *Filoviridae* member [1] responsible for Ebola virus disease (EVD) which leads to ungoverned viral replication and multi-organ failure [2]. The virus is known to multiply in various cell types (hepatocytes, macrophages, endothelial and epithelial cells) and speedily makes its way into the vital organs of the host [3]. Maximum cases of EVD happen due to person to person transmission [4]. Approximately, 30,000 cases of Ebola have been reported till date since 1976 with North Kivu province being the site of latest outbreak in 2018 [5].

EBOV is a single non-segmented negative-stranded RNA virus with an unusual, variable-length, filamentous morphology. It consists of seven proteins viz. nucleoprotein (NP), polymerase cofactors (VP35 and VP40), glycoprotein (GP), transcription activators (VP30 and VP24) and RNA-dependent RNA polymerase (L) [6]. Like its family members, EBOV RNA is incapable of existing in naked form [7]. Nucleoprotein (NP) serves as scaffold for assembly of filovirus nucleocapsid (NC) which includes VP35, V30, VP24 and L [8]. NP interacts with VP35 and VP30 [8] which in turn interact with polymerase and help in assembling viral replication complex [9]. NC plays role

Edited by: Stephan Becker.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00430-019-00584-y) contains supplementary material, which is available to authorized users.

✉ Manoj Baranwal
manoj.baranwal@thapar.edu

[1] Department of Biotechnology, Thapar Institute of Engineering and Technology, Patiala, Punjab 147004, India

Published online: 21 February 2019

in viral RNA synthesis during proliferation cycle [10–12]. Therefore, NP is essential in viral RNA synthesis and virus assembly [10] as ssRNA binding is likely dependent on oligomerization and proper orientation of NP [10, 11]. NP also protects the virus from host innate immune responses and provides resistance to host ribonucleases. When NP-specific CTLs were given to naive mice challenged with a lethal EBOV dose, they helped to induce protection against EBOV indicating the role of cell-mediated immunity against NP [13]. In other recent study, analysis of T cell response was carried out for seven proteins of EBOV in 30 individuals who survived after EBOV infection and it was observed that the maximum survivors (96%) responded against the NP protein as compared to other proteins [14]. Hence, NP is a highly critical protein and, thus, presents itself as a lucrative vaccine design target.

Vaccine development against EBOV is still in development phase and various trials include viral vector-based vaccines [15], protein-based vaccines [16] and subunit vaccines [17]. In recent years, there has been a remarkable progress in peptide-based vaccines which are fragments of protein antigen sequences assembled into a single molecule capable of inducing an immune response. Immunoinformatics tools have shown success in elucidating potent peptide vaccine candidates against influenza virus [18], hepatitis C [19], West Nile virus [20] and EBOV [21]. The immunogenic peptides obtained using this approach were validated in in vitro system for influenza [22] and in in vivo system for Brucella abortus [23] and EBOV [24].

In the present investigation, peptides containing multiple epitopes against EBOV NP were selected based on different epitope prediction tools and examined for their conservation among the EBOV species and other *Filoviridae* members. These peptides were looked for binding potential to diverse HLA molecules based on different prediction tools, docking and population coverage analysis. Further, in vitro validation of immunogenic response of three potential peptides was carried out by measuring IFN-γ secreted by peptide-stimulated peripheral blood mononuclear cells (PBMC) isolated from healthy blood samples.

Method

Conserved peptides identification

195 unique Ebola nucleoprotein sequences (739 amino acids) out of a total of 2407 entries (1976–May 2018) were downloaded from viprbrc and NCBI databases. These sequences comprised 187 (Zaire), five (Sudan), two (Bundibugyo) and one (Taï Forest) sequences belonging to various Ebola species pathogenic to humans. MUSCLE [25] and AVANA [26] tools were employed to identify peptide fragments showing at least 90% conservancy.

Prediction of T and B cell epitopes

T cell epitopes were predicted based on a consensus approach [27] that includes three prediction tools (SYFPEITHI, NetCTL 1.2 and IEDB consensus) for $CD8^+$ T cell epitopes (HLA class I) and three tools (MHC2Pred, Propred and IEDB consensus) for $CD4^+$ T cell epitopes (HLA class II). The detailed information of each tool is mentioned in Table 1. The epitopes showing overlaps were further joined to obtain peptide fragments containing both $CD4^+$ and $CD8^+$ T cell epitopes.

Linear, 10 amino acid long B cell epitopes were identified with the help of ABCpred at default threshold (0.51). This tool utilizes a recurring neural network method to predict B cell epitopes with 65.93% accuracy [33].

Screening of peptides for autoimmune and allergic response

Peptides having seven out of nine consecutive amino acids identical to human proteome were eliminated using BLAST analysis. Allergenicity of the peptides was predicted using the online tool AlgPred, which is based on screening IgE epitopes in query protein sequence and Motif Alignment & Search Tool [34].

Table 1 T cell epitope prediction tools

T cell	Tools	Algorithm	Threshold	Reference
$CD8^+$	SYFPEITHI	Motifs search approach	20	Rammensee et al. [28]
	NetCTL 1.2	Weight-matrix method	0.75	Larsen et al. [29]
	IEDB consensus	Artificial neural network	$IC_{50} < 500$ nm	Nielsen et al. [30]
$CD4^+$	MHC2Pred	Support vector machine (SVM) method	0.5	Bhasin et al. [31]
	Propred	Quantitative matrices	3%	Singh et al. [32]
	IEDB consensus	Artificial neural network	$IC_{50} < 500$ nm	Nielsen et al. [30]

Conservancy analysis amongst Ebola species and other filoviridae members

The identified peptides were looked for conservancy in human-pathogenic Ebola species (Zaire, Sudan, Bundibugyo and Taï forest) sequences as well as in 18 unique out of a total of 79 *Marburgvirus* nucleoprotein sequences and the only available unique nucleoprotein sequence of *Lloviuvirus* obtained from viprbrc and NCBI databases.

Molecular docking

HLA–peptide interaction analysis was done with the help of CABS-dock which allows for flexibility of peptide and receptor backbone [35]. High-resolution crystal structures of eighteen HLA class I and II alleles (nine each) bound to their respective native peptide were obtained from PDB. The HLA crystal structures without their native peptides (peptides already bound to the HLA) were obtained using Discovery Studio Visualizer (version 4.1). RMSD values obtained by docking the native peptides to their respective HLA structures served as standard. Peptides showing RMSD > 5 or found to be interacting outside the binding groove were eliminated.

Population coverage analysis

IEDB population coverage analysis tool, based on peptide-HLA data and HLA genotypic frequency, plays an important role in a bid to develop a globally protective vaccine. The selected peptides and their HLA alleles obtained from prediction tools were used as input for this tool. For this analysis, four different geographical continents (Africa, America, Asia and Europe) were chosen. Africa, America and Asia comprised 13 different geographical regions and, therefore, the average of population coverage for these regions was considered. Analysis was also carried out by taking into account the whole world.

Statistical analysis

One-way Anova followed by Tukey's multiple comparison test using GraphPad Prism was carried out to analyze the docking data.

Measurement of IFN-γ secreted by peptide-stimulated peripheral blood mononuclear cells

P2, P3 and P5 were commercially synthesized by GL Biochem (Shanghai) Ltd. Healthy blood samples were obtained from Nitin Hospital, Patiala and Rajindra Hospital, Patiala (India) after informed consent from all volunteers. The study was approved by the institutional ethical committee. Peripheral blood mononuclear cells (PBMC) were isolated via ficoll density gradient method [22]. Restimulation assay was carried out for measuring peptide-induced IFN-γ secretion with certain modifications to the previous report [36]. In a 96-well cell culture plate, 2×10^5 cells were seeded per well in a total volume of 200 µL complete media (RPMI-1640 supplemented with 10% fetal bovine serum, 100 µg/mL streptomycin, 100 I.U./mL penicillin and 10 mM HEPES) and stimulated with each peptide (50 µg/mL). Unstimulated cells served as negative control while cells stimulated with 10 µg/mL of concanavalin A (ConA, Sigma-Aldrich) served as positive control. Restimulation was done on 3rd day with each peptide. On 5th day, IFN-γ secreted by unstimulated, peptide-stimulated and ConA-stimulated cells was measured by performing ELISA with the help of human IFN-γ mini Elisa development kit (Peprotech, USA). A microplate reader (Tecan Austria) was used to take absorbance at 405 nm with 630 nm as reference wavelength. All experiments were carried out in triplicates. IFN-γ production was expressed as fold change which is the ratio of absorbance of peptide-stimulated cells and unstimulated cells.

Results

Conserved peptides containing T and B cell epitopes and having no autoimmune and allergic properties

Four overlapping fragments (C1–C4) with ≥ 90% conservancy in 195 Ebola nucleoprotein sequences were obtained after multiple sequence alignment via MUSCLE and conservancy analysis via AVANA (Online Resource 1). Next, epitopes commonly predicted in the identified fragments by six epitope prediction tools (three each for HLA class I and II) were considered. Initially, 105 and 79 HLA class I ($CD8^+$ T cell) and II ($CD4^+$ T cell) binding epitopes were obtained, respectively (Online Resource 2). Twelve peptide fragments containing multiple $CD8^+$ and $CD4^+$ T cell epitopes were obtained by merging overlapping epitopes (Online Resource 3).

A total of 201 linear B cell epitopes were obtained after analyzing the four conserved fragments (C1–C4) via ABCpred (Online Resource 2). The predicted B cell epitopes were present only in eight identified fragments. These eight peptides were checked for autoimmune and allergic responses. Two peptide fragments (VGHMMVIFRLMRTNFLIKFLLIHQGMHMV and YAPFARLLNLSGV) exhibited similarity to intrinsic human proteins based on BLAST analysis and, hence, were eliminated. Algpred tool confirmed none of the peptides to be allergic in nature. Thus, six non-self and non-allergic peptide candidates possessing multiple B and T cell epitopes were selected (Table 2).

Table 2 Peptides representing the presence of different T and B cell epitopes

Peptide fragments	$CD8^+$ epitopes	$CD4^+$ epitopes	B-cell epitopes
MDYHKILTAGL (P1)	YHKILTAGL	MDYHKILTA	DYHKILTAGL
		YHKILTAGL	
IVRQRVIPVYQVNNLEEI (P2)	IVRQRVIPV	IVRQRVIPV	VYQVNNLEEI
	VRQRVIPVY	IPVYQVNNL	RVIPVYQVNN
	IPVYQVNNL	VYQVNNLEE	RQRVIPVYQV
	YQVNNLEEI	YQVNNLEEI	
FLSFASLFLPKL (P3)	FLSFASLFL	FLSFASLFL	FLSFASLFLP
	SFASLFLPK		LSFASLFLPK
	FASLFLPKL		
ERGVRLHPLARTAK (P4)	ERGVRLHPL	LHPLARTAK	RLHPLARTAK
	GVRLHPLAR	VRLHPLART	ERGVRLHPLA
			VRLHPLARTA
LFPQLSAIALGVATAHGS (P5)	FPQLSAIAL	LFPQLSAIA	AIALGVATAH
	QLSAIALGV	FPQLSAIAL	LSAIALGVAT
		LGVATAHGS	FPQLSAIALG
			ALGVATAHGS
			QLSAIALGVA
VLYYHMMKDEPVVF (P6)	HMMKDEPVV	VLYYHMMKD	YHMMKDEPVV
		LYYHMMKDE	VLYYHMMKDE
		YHMMKDEPV	HMMKDEPVVF
		MMKDEPVVF	

Conservation analysis of peptides amongst different Ebola virus species and related family members

Six identified fragments were investigated for their conservancy in different species of Ebola virus (Zaire, Sudan, Bundibugyo and Taï Forest) and other members (*Marburgvirus* and *Lloviuvirus)* of *Filoviridae* to judge the potential of these candidates to develop cross protective immunity. Interestingly, all selected peptides showed 100% conservancy amongst *Zaire ebolavirus* nucleoprotein sequences (Table 3). P3 was found to be 100% conserved in all Ebola virus species and *Lloviuvirus* and there was a single amino acid variation in case of Marburgvirus. P5 was found to be 100% conserved in three Ebola virus species (Zaire, Bundibugyo and Taï Forest) while one amino acid variation at same position was observed in Sudan, Lloviu and Marburg viruses (Table 3). Rest of the peptides were found with few variations in different Ebola virus species. In most cases, one

Table 3 Peptide conservation in Ebola virus species and other filoviridae members

Virus species	MDYHKILTAGL (P1)	IVRQRVIPVYQVNNLEEI (P2)	FLSFASLFLPKL (P3)	ERGVRLHPLARTAK (P4)	LFPQLSAIALGVATAHGS (P5)	VLYYHMKDEPVVF (P6)
ZAIRE	100%	100%	100%	100%	100%	100%
[a]SUDAN	**L**DYHKILTAGL	IVRQR**I**IPVY**V**VS**-** **D**LE**G**I/IVRQRVIPVY**-** **V**VN**D**LE**G**I	100%	**DL**GVRLHPLARTAK	L**Y**PQLSAIALGVATAHGS	**IN**YYH**L**MS**-** DEP**IA**F
TAÏ FOREST	**T**DYHKILTAGL	IVRQRVI**Q**V**H**QV**T**NLEEI	100%	E**H**GVRLHPLARTAK	100%	**I**LYY**Y**M M**TE**EP**I**VF
BUNDIBUGYO	**A**DYHKILTAGL	IVRQR**I**IPVYQ**IS**NLEE**V**	100%	E**H**GVRLHPLARTAK	100%	**I**LYYHM MK**E**EP**II**F
LLOVIU			100%	E**A**GV**Q**LHPLART**S**K	L**Y**PQLSAIALGVATAHGS	
[a]**MARBURG**			FLSF**C**SLFLPKL	**DS**GV**T**LHPL**V**RT**S**K/ **DS**GV**A**LHPL**V**RT**S**K	L**Y**PQLSAIALGVATAHGS	

[a]P2 and P4 of Sudan and Marburg virus, respectively, have two variable sequences

Table 4 Peptides containing multiple epitopes binding to diverse HLA alleles

Tools	Class I						Class II					
	SYFPEITHI		NetCTL 1.2		IEDB		ProPred	MHC2Pred		IEDB		
HLA type	A	B	A	B	A	B	DR	DQ	DR	DP	DQ	DR
Number[a]	7	26	5	7	18	32	51	11	27	5	6	14
MDYHKILTAGL (P1)	–	4	–	2	–	1	22	5	5	–	1	10
IVRQRVIPVY-QVNNLEEI (P2)	1	5	1	5	5	6	51	9	18	3	2	14
FLSFASLFLPKL (P3)	3	2	3	1	9	–	18	2	4	5	1	7
ERGVRLH-PLARTAK (P4)	3	5	1	1	2	3	24	–	7	–	1	7
LFPQLSAIALGVA-TAHGS (P5)	2	4	1	3	3	6	15	3	6	1	2	10
VLYYHMMK-DEPVVF (P6)	1	–	1	–	2	–	18	4	7	4	2	13

**Both alleles and supertypes (a supertype is a group consisting of similar alleles) have been mentioned in this table. A complete list of alleles predicted by various computational tools for all the identified peptides has been provided (Online Resource 4)

[a]Total number of HLA alleles/supertypes in each tool

variable sequence was observed but P2 (*Sudan ebolavirus*) and P4 (*Marburgvirus*) were found to be with two variable sequences (Table 3). P1, P2 and P6 could not be located in NP sequence of Lloviu and Marburg viruses.

Peptide–HLA interactions

Peptides are presented by HLA molecules to induce immune response and HLA polymorphism is well known [37, 38]. Therefore, it is desirable for potent peptide vaccine candidates to exhibit interaction with a wide range of HLA molecules. During epitope prediction, all HLA alleles/supertypes available in various tools (Table 4) were considered. All selected peptides were found to bind with diverse and large number of HLA alleles which are of HLA-A, HLA-B, HLA-DP, HLA-DQ and HLA-DR types (Table 4) and the complete detail of HLA restrictions of all peptides has been mentioned (Online Resource 4). P2, P3 and P5 peptides were found to be predicted for most HLA types (Table 4) as well as maximum number of HLA alleles (Online Resource 4).

Docking study gives a wider prospective to understand the actual binding interaction of peptides with HLA molecules. Eighteen HLA alleles belonging to different HLA categories were chosen for docking with CABS-dock. Crystal structures for eighteen HLA alleles were obtained from PDB and native peptides, ranging from 8 to 11 residues in length for HLA class I and 9–20 residues in length for HLA class II, were separated using Discovery studio Visualizer 4.1. Nonamer $CD8^+$ T cell epitopes which are part of selected six peptides (Table 2) were docked with HLA class I molecules as the binding groove of HLA class I is closed and can accommodate 8–10 residue peptides. HLA class II has open grooves and is capable of presenting 13–25 residue peptides; thus, six peptides as such were docked with class II molecules [39]. RMSD values (CABS-dock) obtained by docking native peptides to their respective HLA molecules were used as test control.

The average of RMSD values of $CD8^+$ T cell epitopes (HLA class I) which are associated with respective peptide is plotted in Fig. 1. The models with RMSD value < 3 are considered high-quality predictions while those with $3 \leq \text{RMSD} \geq 5.5$ as moderate quality predictions [40]. In some HLA–peptide interactions such as P1 (B*1801 and DRB*0101), P4-DQ8 and P6-DQ8, the RMSD value was > 5 showing poor quality prediction and, hence, these values were not plotted. In majority of the cases, the RMSD value was less than 3 for peptide-HLA (class I and II) (Fig. 1). The mean and median RMSD value of each peptide was also found to be less than 3 which confirmed the strong binding interaction of predicted peptide with eighteen HLA molecules (Table 5). It was observed that RMSD values of all peptides were not significantly different from native peptides and also within the range of native peptides with few variations. The majority of RMSD data display either positive skewness or a skewness value close to zero (normal distribution) indicating stable interactions between a greater number of HLA alleles and identified peptides. P2, P4 and P6 displayed negative skewness for HLA class I indicating a highly stable interaction between them and only a few HLA class I alleles while these peptides displayed positive skewness for HLA class II alleles. Based on mean binding energy,

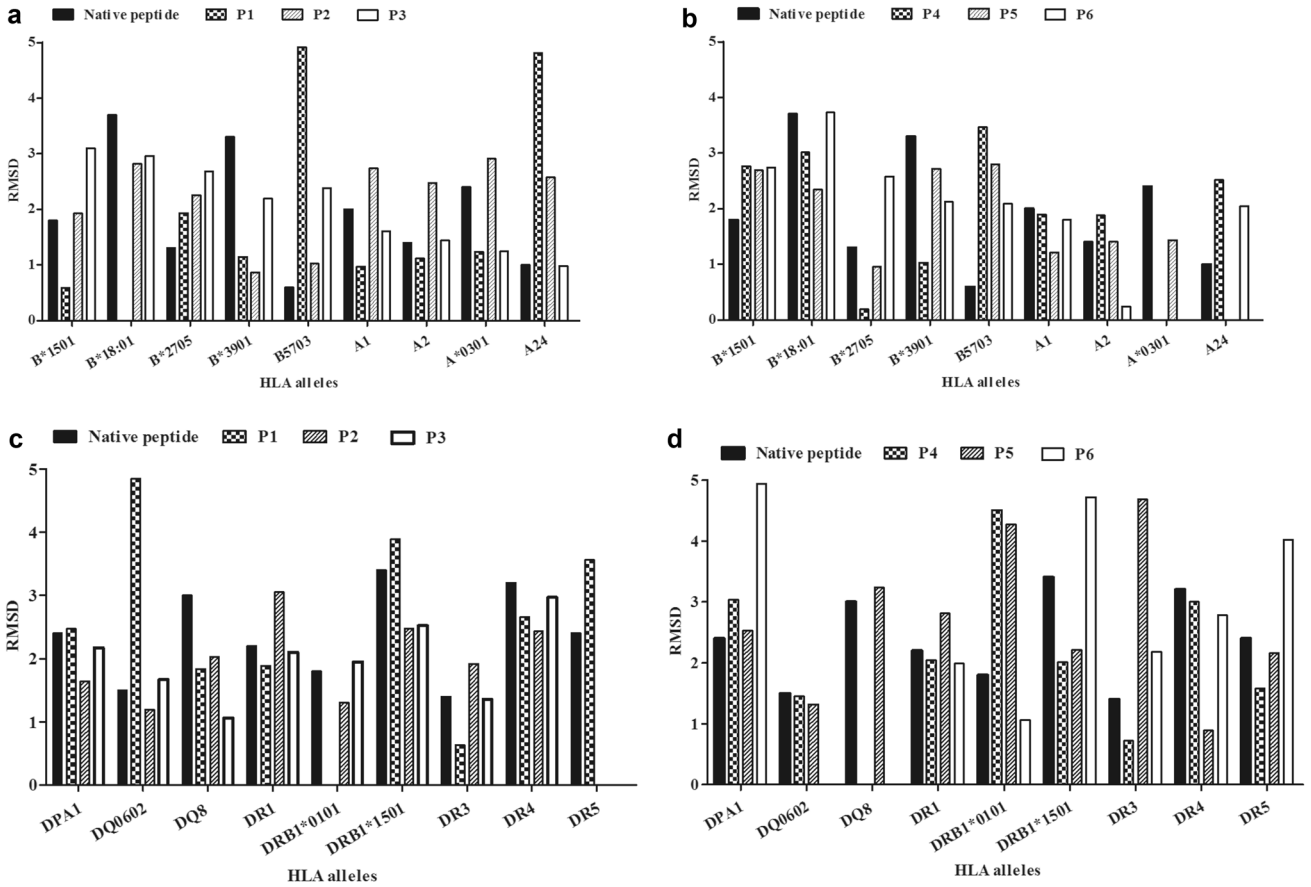


Fig. 1 RMSD of the native peptide (NP) and identified Ebola nucleoprotein peptides obtained by CABS-dock analysis. For HLA class I, RMSD of **a** Native peptide, P1 P2 and P3 peptides and **b** Native peptide, P4, P5 and P6 peptides. For HLA class II, RMSD of **c** Native peptide, P1, P2 and P3 peptides and **d** Native peptide, P4, P5 and P6 peptides. For **a** and **b**, the mean RMSD of epitopes belonging to the respective peptide was considered except for peptides consisting of a single $CD8^+$ T cell specific epitope (P1 and P6). For **c** and **d**, RMSD value obtained after docking-identified peptides with various HLA alleles was considered. Native peptides represent the peptide that already existed in the crystallographic structures of the HLA molecules. They were separated and docked with their respective HLA molecule. RMSD values found to be >5 represent poor quality predictions and, hence, they were not plotted

P5 displayed best interaction with HLA class I molecules amongst the identified peptides while P2 and P3 displayed better interaction ability with HLA class II molecules as compared to native peptides (Table 5).

In addition to docking and HLA coverage in tools, population coverage analysis was carried out which provides the expected response of the peptides to various HLA molecules in different geographic regions. Encouraging results were observed as the all peptides exhibited more than 95% coverage for American, Asian and European populations (Fig. 2) and the expected response was found to be 90–100% when the analysis was done by taking whole world. The average of response for each peptide to four continents (Africa, America, Asia and Europe) comprising 14 total geographical areas was P1 (94.7%), P2 (96%), P3 (95.4%), P4 (85%), P5 (95%) and P6 (94.3%).

Mapping of peptide fragments

P2–P5 were found to be located in the core domain of NP protein. P1 was found to be near the N tail of the protein while P6 was found to be a part of C tail (Fig. 3).

Peptide-induced IFN-γ secretion

Three peptides (P2, P3 and P5) performed better during docking analysis as well as exhibited greater HLA allele coverage. Hence, these peptides were tested in vitro to validate their immunogenic potential. PBMC isolated from healthy blood samples were incubated as such (unstimulated cells), with peptides (peptide-stimulated cells) and ConA (positive control). 9 out of 10 samples showed enhanced IFN-γ secretion (fold change >= 1.0) for peptide-stimulated cells in case of P2 and P3. P5 was less responsive as only 2 out of 10 samples clearly showed enhanced IFN-γ secretions as

Table 5 Comparative analysis of RMSD value (CABS-Dock) of selected peptides with native peptides

	Minimum	25% percentile	Median	75% percentile	Maximum	Mean	SD	Skewness	Significant? $P<0.05$?
Class I									
NP	0.60	1.15	1.80	2.85	3.70	1.94	1.03	0.62	
P1	0.60	1.02	1.20	4.07	4.90	2.08	1.74	1.27	ns
P2	0.90	1.45	2.50	2.75	2.90	2.16	0.75	−1.01	ns
P3	1.0	1.30	2.20	2.85	3.10	2.06	0.79	−0.01	ns
P4	0.20	1.40	1.90	2.90	3.50	2.06	1.02	−0.52	ns
P5	1.0	1.25	1.90	2.70	2.80	1.95	0.76	−0.02	ns
P6	0.20	1.87	2.10	2.67	3.70	2.16	0.98	−0.73	ns
Class II									
NP	1.40	1.65	2.40	3.10	3.40	2.36	0.72	0.08	
P1	0.60	1.82	2.60	3.82	4.90	2.73	1.36	0.08	ns
P2	1.20	1.37	1.95	2.47	3.10	2.00	0.64	0.42	ns
P3	1.10	1.47	2.05	2.42	3.00	2.00	0.60	0.13	ns
P4	0.70	1.52	2.00	3.00	4.50	2.28	1.17	0.78	ns
P5	0.90	1.75	2.50	3.75	4.70	2.67	1.25	0.34	ns
P6	1.10	2.00	2.80	4.70	4.90	3.10	1.45	0.04	ns

NP native peptide (Peptides which were bound to the PDB HLA structure)

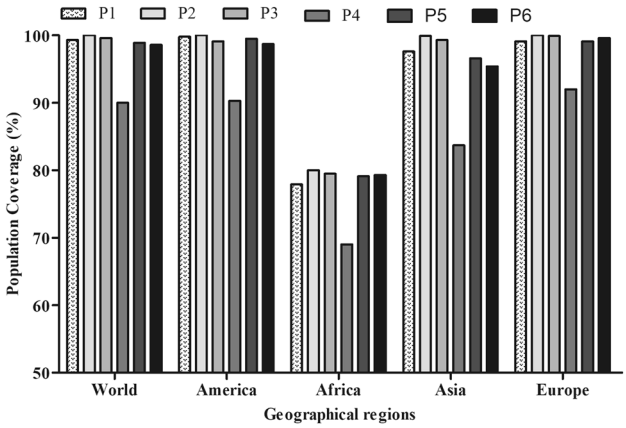


Fig. 2 Population coverage of the identified peptides in four different continents and whole world. The mean population coverage for three different geographical continents (America, Africa and Asia) has been plotted

compared to unstimulated cells (Fig. 4). ConA-treated cells showed more IFN-γ production for all the samples.

Discussion

Peptides as a choice for vaccine formulation is one of the recent developments and many peptide-based vaccines are in different stages of clinical trials. More than ten anti-cancer peptide vaccine candidates have made it to phase III trials [41]. Phase II trials are being conducted for peptide vaccines against influenza [42] and HPV-induced cancer [43]. Relying solely on in vitro or in vivo analysis for immunogenic peptide identification is cumbersome and not feasible in all facilities around the world. Computational immunology offers advantages in downsizing the number of peptide candidate to be validated in vitro or in vivo. In the current study, six EBOV NP peptides containing multiple epitopes having potential to interact with an array of HLA molecules were identified. Numerous computational works have been done in identifying epitopes against different infectious organisms

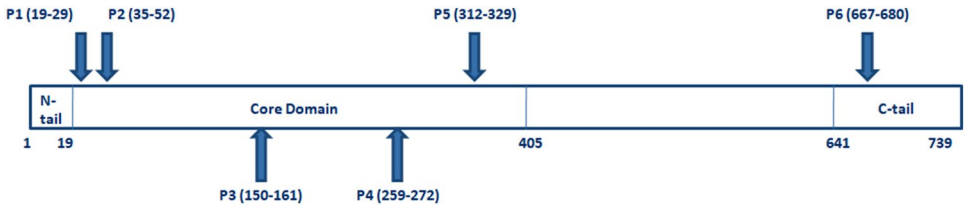


Fig. 3 Schematic presentation of identified peptides in different regions of Ebola nucleoprotein

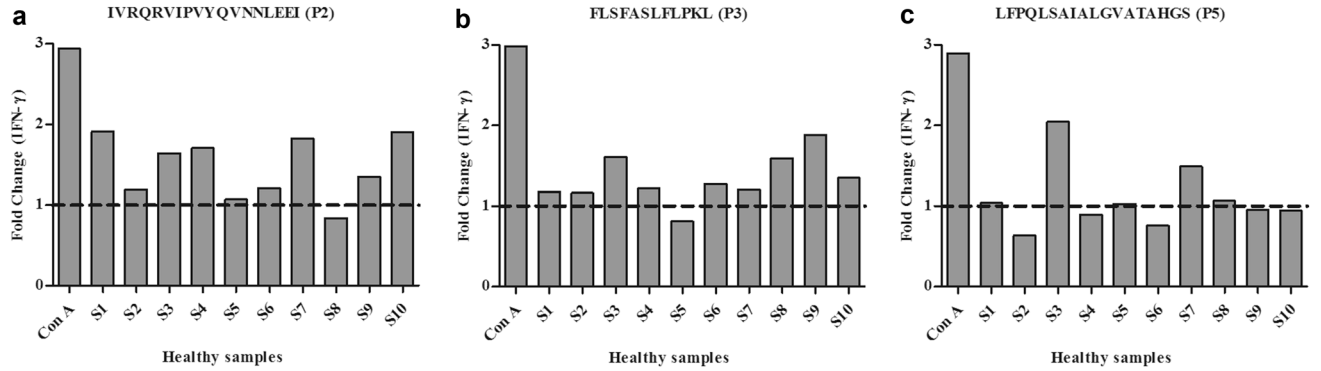


Fig. 4 IFN-γ secretion by peripheral blood mononuclear cells of 10 healthy blood samples has been presented and expressed as fold change for **a** P2, **b** P3 and **c** P5. Fold change is the ratio of absorbance of peptide-stimulated cells and unstimulated cells. S1–S10 represent the 10 healthy blood samples. ConA:Concanavalin A

such as Leishmania [44], Mycobacterium tuberculosis [45] and influenza virus [46]. Immunoinformatically identified peptides have shown enhanced proliferation and IFN-γ production in peptide-induced peripheral blood mononuclear cells [22]. Also, in vivo validation of computationally generated peptide vaccine candidates with mice as subject for Leishmania donovani [47], Moraxella catarrhalis [48] and Brucella abortus [23] showed promising results. Hence, computational identification of potential peptide vaccine candidates against different infections has picked pace in recent years. EBOV represents a serious concern for human health and, thus, with the application of various computational tools, six EBOV NP peptides containing multiple T and B cell epitopes were designed. Owing to the presence of different epitopes, these peptides may be capable of generating both humoral and cell-mediated immunity.

One of the interesting aspects of this study is the usage of consensus approach in which six prediction tools (three each for $CD8^+$ and $CD4^+$ T cell epitopes) were employed. The advantage of consensus approach is consideration of multiple prediction algorithms, immunological factors and databases in predicting the epitopes. In contrast to the present study, Sundar et al. used one prediction tool [49] while two prediction tools were used by Dutta et al. [50]. An approach similar to the current study was employed by Dikhit et al. where three different prediction algorithms were used to define only $CD8^+$ T cell epitopes but not $CD4^+$ T cell epitopes [51]. In all these studies [49–51], the work was carried out only for identifying $CD8^+$ T cell epitopes. The present study focuses on the identification of peptides containing both $CD8^+$ and $CD4^+$ T cell epitopes. Further, in contrast to previous studies [49–51] where protein sequences were taken from single EBOV strain, the present study has considered the sequences from 1976 to May 2018 of all EBOV strains infecting humans. The conserved sequences obtained from 195 unique Ebola nucleoprotein sequences were taken for predicting the epitopes.

As per the comparative analysis performed with the help of IEDB, none of the six peptide fragments were found to be reported exactly in any of the previous studies. YQVNNLEEI, FLSFASLFL and FPQLSAIAL are the partial fragments of P2, P3 and P5 peptides, respectively (Table 2), that have been reported to induce $CD8^+$ T cell response in earlier studies [49, 52–55]. In another study, mice injected with EBOV NP vaccine responded by producing protective CTLs against VYQVNNLEEIC (consisting of a part of P2) [56]. A large peptide sequence, HILRSQGPFDAVLYYHMMKDEPVVFSTSDGKEYTYP (consisting of P6), was reported to induce $CD8^+$ T cell response in human survivors [14].

Any component of a vaccine may contribute to adverse effects after vaccine administration [57]. There is a chance of autoimmune reactions after vaccination [58]. Association of polyarthritis and thrombocytopenia with hepatitis B vaccine administration has been reported [59]. Also, neural complications on administering anti-rabies and tetanus vaccinations have been presented in previous studies [60]. Hence, BLAST analysis was applied in the current study to remove potential autoimmune response elucidators. VGHMMVIFRLMRTNFLIKFLLIHQGMHMV was eliminated owing to its similarity to human Anaphase-promoting complex/cyclosome and YAPFARLLNLSGV was eliminated as it showed similarity to homeobox protein, Hox-B9. Angioedema, bronchospasm, shock and a drop in blood pressure are some of the allergic responses to vaccines [57]. Measles–Mumps–Rubella (MMR) vaccination has been reported to induce anaphylactic reactions in some individuals and has been a subject of great debate for administration to egg allergic children [61, 62]. Current in silico analysis established that all peptides under consideration were incapable of inducing any allergic responses.

To validate the in silico selected peptides in the in vitro system, three peptides were checked for IFN-γ secretion in peptide-stimulated PBMC. In previous studies also, extracellular release of IFN-γ to represent antigen-induced proliferation of T cells has been measured by ELISA [22, 63, 64]. Interestingly, nine out of ten samples responded with enhanced IFN-γ production for two peptides (P2 and P3), thus validating the immunogenic potential of these computationally identified peptides.

Conservation analysis presented a wider prospective to identify peptide candidates which may have the potential to provide immunity against existing and future viral strains as well as cross protective immunity. As per various previous studies, 36–351 residues (P2-P5 lie in this polypeptide) are highly conserved amongst all EBOV strains, *Marburgvirus* as well as *Lloviuvirus* [11] belonging to the same family (*Filoviridae*) and other Mononegavirales members such as respiratory syncytial virus [65], Nipah virus [66] and Parainfluenza virus 5 [67] have shown structural similarity to EBOV. The identified peptides have shown conservation and similarity with EBOV species and other filovirus (*Marburgvirus* and *Lloviuvirus*) but not with other Mononegavirales members. Interestingly, all selected peptides were 100% conserved amongst *Zaire ebolavirus* strains. Further, P3 and P5 were conserved (100%) in most human infecting Ebola species and considered filoviridae members barring a single amino acid variation in a few species. These results supported the idea of development of cross-protective peptide vaccine.

One of the challenges in the development of peptide-based vaccine is to identify peptide candidates which provide immunity to different populations across the world. Adaptive immune response is directly associated with peptides presented by HLA molecules which are highly polymorphic. 13,680 HLA I and 5091 HLA II alleles belonging to different populations of the world have been reported in IPD-IMGT/HLA Database (release version 3.33, July 2018) [68, 69]. Peptide selection in the current study was done by accounting for multiple epitopes which provide opportunity to include large number of HLA alleles belonging to HLA-A, HLA-B, HLA-DP, HLA-DQ and HLA-DR. Further, docking studies have been carried out previously to ascertain the same in case of influenza virus and human baculo virus [18, 70]. Docking analysis was also carried out for EBOV epitopes with only HLA A0201 allele earlier [51]. In the current study, high-resolution nine PDB protein structures, each of HLA class I and II belonging to different HLA alleles, were considered. Statistical analysis indicated that the average RMSD value of selected peptides was not significantly different from that of native peptides indicating their similar binding potential. Moreover, RMSD value is less than 3 in most of the cases indicating highly favorable binding interactions with selected HLA molecules. Population coverage analysis was employed as an additional computational tool to judge the expected immune response of peptides in different geographical populations. Average population coverage lied in the range 90–100% for America, Asia, Europe and whole world which further confirmed the potential of the selected peptides as global vaccine candidates.

Further, these peptides were found to be parts of crucial nucleoprotein domains. P1–P5 are a part of N terminal 1-450 residues, a polypeptide which in itself is sufficient for viral genome replication [71]. P2–P5 are a part of N-terminal 36–351 residues which are needed for NP oligomerization and RNA binding [12]. Most of P1 form an oligomerization arm of NP [11]. A part of P1 lies in N-terminal 1–24 residues which enhance ssRNA binding as well as control NP intermolecular interactions [10, 12]. One residue of P3 (160th) is amongst the four important nucleoprotein residues (160, 171, 174 and 248) responsible for RNA encapsidation [72] and their deletion impairs EBOV replication [73]. Residues of P4 and P5 (264, 268 and 316) are involved in the formation of a highly conserved hydrophobic pocket significant for RNA formation [12].

Vaccine development against EBOV is in progress and nearly eight ebola vaccines candidates have made it to clinical trials [74]. A replication competent vesicular stomatitis virus vector-based vaccine named rVSVΔG-ZEBOV-GP has shown encouraging results in random and non-random trials conducted in Guinea, West Africa [75] and was proven to be effective with tolerable side effects [76]. Another vaccine named ChAd3-EBO is a non-replicating chimpanzee adenovirus vector-based vaccine which has exhibited efficacy in non-human primates [77] and an increased response with a modified vaccinia Ankara (MVA) booster [78]. Although some EBOV vaccines are in clinical trials, there is still a long way to reach the market. Peptides identified in this study offer advantages of being small molecules with a potential to provide immunity against all EBOV strains and other filoviruses and, hence, might be considered for designing a globally protective vaccine.

Conclusion

Six non-self and non-allergic peptides having multiple T and B cell epitopes were obtained and found to be 100% conserved in Zaire EBOV species. These peptides were predicted for diverse HLA alleles and found to have strong binding affinity with eighteen different HLA molecules. Also, the peptides exhibited strong population coverage among the different geographical regions across the globe. Two out of three potential peptides tested for in vitro immune response showed enhanced IFN-γ production for peptide-stimulated PBMC. Thus, these peptides are proposed to be validated further to inculcate these peptides in design of a synthetic peptide vaccine against EBOV and related species.

Acknowledgements We express our sincere thanks to the scientific community for developing in silico tools. We also express our gratitude towards Dr. Akshey Jain and Dr. Vandana Singla for providing us blood samples of healthy volunteers.

Compliance with ethical standards

Conflict of interest All authors declare that they have no conflict of interest.

References

1. Kuhn JH, Becker S, Ebihara H et al (2010) Proposal for a revised taxonomy of the family Filoviridae: classification, names of taxa and viruses, and virus abbreviations. Arch Virol 155(12):2083–2103
2. Lee JE, Fusco ML, Hessell AJ, Oswald WB, Burton DR, Saphire EO (2008) Structure of the Ebola virus glycoprotein bound to an antibody from a human survivor. Nature 454(7201):177–182
3. Sanchez A, Lukwiya M, Bausch D, Mahanty S, Sanchez AJ, Wagoner KD, Rollin PE (2004) Analysis of human peripheral blood samples from fatal and nonfatal cases of Ebola (Sudan) hemorrhagic fever: cellular responses, virus load, and nitric oxide levels. J Virol 78(19):10370–10377
4. Ilesanmi O, Alele FO (2016) Knowledge, attitude and perception of Ebola virus disease among secondary school students in Ondo State, Nigeria, October, 2014. PLoS Curr 8: ecurrents.outbreaks.c04b88cd5cd03cccb99e125657eecd76
5. Crisis update—November (2018) https://www.msf.org/drc-2018-ebola-outbreak-crisis-update Accessed 5 Jan 2019
6. Brauburger K, Boehmann Y, Tsuda Y, Hoenen T, Olejnik J, Schumann M, Ebihara H, Muhlberger E (2014) Analysis of the highly diverse gene borders in Ebola virus reveals a distinct mechanism of transcriptional regulation. J Virol 88(21):12558–12571
7. Sun Y, Guo Y, Lou Z (2012) A versatile building block: the structures and functions of negative-sense single-stranded RNA virus nucleocapsid proteins. Protein Cell 3(12):893–902
8. Huang Y, Xu L, Sun Y, Nabel GJ (2002) The assembly of Ebola virus nucleocapsid requires virion-associated proteins 35 and 24 and posttranslational modification of nucleoprotein. Mol Cell 10(2):307–316
9. Trunschke M, Conrad D, Enterlein S, Olejnik J, Brauburger K, Muhlberger E (2013) The L-VP35 and L-L interaction domains reside in the amino terminus of the Ebola virus L protein and are potential targets for antivirals. Virology 441(2):135–145
10. Su Z, Wu C, Shi L et al (2018) Electron cryo-microscopy structure of Ebola virus nucleoprotein reveals a mechanism for nucleocapsid-like assembly. Cell 172(5):966–978
11. Kirchdoerfer RN, Abelson DM, Li S, Wood MR, Saphire EO (2015) Assembly of the Ebola virus nucleoprotein from a chaperoned VP35 complex. Cell Rep 12(1):140–149
12. Dong S, Yang P, Li G, Liu B, Wang W, Liu X, Xia B, Yang C, Lou Z, Guo Y, Rao Z (2015) Insight into the Ebola virus nucleocapsid assembly mechanism: crystal structure of Ebola virus nucleoprotein core domain at 1.8 Å resolution. Protein Cell 6(5):351–362
13. Wilson JA, Bray M, Bakken R, Hart MK (2001) Vaccine potential of Ebola virus VP24, VP30, VP35, and VP40 proteins. Virology 286(2):384–390
14. Sakabe S, Sullivan BM, Hartnett JN et al (2018) Analysis of $CD8^+$ T cell response during the 2013–2016 Ebola epidemic in West Africa. Proc Natl Acad Sci USA 115(32):E7578–E7586
15. Herbert AS, Kuehne AI, Barth JF et al (2013) Venezuelan equine encephalitis virus replicon particle vaccine protects nonhuman primates from intramuscular and aerosol challenge with ebolavirus. J Virol 87(9):4952–4964
16. Swenson DL, Wang D, Luo M, Warfield KL, Woraratanadharm J, Holman DH, Dong JY, Pratt WD (2008) Vaccine to confer to nonhuman primates complete protection against multistrain Ebola and Marburg virus infections. Clin Vccine Immunol 15(3):460–467
17. Martins KA, Jahrling PB, Bavari S, Kuhn JH (2016) Ebola virus disease candidate vaccines under evaluation in clinical trials. Expert Rev Vaccines 15(9):1101–1112
18. Lohia N, Baranwal M (2014) Conserved peptide containing overlapping CD4 + and CD8 + T cell epitopes in H1N1 influenza virus: An immunoinformatics approach. Viral Immunol 27(5):225–234
19. Sirskyj D, Diaz-Mitoma F, Golshani A, Kumar A, Azizi A (2011) Innovative bioinformatic approaches for developing peptide-based vaccines against hypervariable viruses. Immunol Cell Biol 89(1):81–89
20. Hossain MU, Keya CA, Das KC, Hashem A, Omar TM, Khan MA, Rakib-Uz-Zaman SM, Salimullah M (2018) An immunopharmacoinformatics approach in development of vaccine and drug candidates for West Nile Virus. Front Chem 6(246). https://doi.org/10.3389/fchem.2018.00246
21. Ali MT, Islam MO (2015) A highly conserved GEQYQQLR epitope has been identified in the nucleoprotein of Ebola virus by using an in silico approach. Adv Bioinformatics 2015:278197. https://doi.org/10.1155/2015/278197
22. Lohia N, Baranwal M (2017) Immune response of highly conserved influenza A virus matrix 1 peptides: matrix 1 peptides for influenza vaccine. Microbiol Immunol 61(6):225–231
23. Afley P, Dohre SK, Prasad GB, Kumar S (2015) Prediction of T cell epitopes of Brucella abortus and evaluation of their protective role in mice. Appl Microbiol Biotechnol 99(18):7625–7637
24. Bounds CE, Terry FE, Moise L, Hannaman D, Martin WD, De Groot AS, Suschak JJ, Dupuy LC, Schmaljohn CS (2017) An immunoinformatics-derived DNA vaccine encoding human class II T cell epitopes of Ebola virus, Sudan virus, and Venezuelan equine encephalitis virus is immunogenic in HLA transgenic mice. Hum Vaccin Immunother 13(12):2824–2836
25. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32(5):1792–1797
26. Miotto O, Heiny A, Tan TW, August JT, Brusic V (2008) Identification of human-to-human transmissibility factors in PB2 proteins of influenza A by large scale mutual information analysis. BMC Bioinformatics 9(1):18
27. Dhiman G, Lohia N, Jain S, Baranwal M (2016) Metadherin peptides containing $CD4^+$ and $CD8^+$ T cell epitopes as a therapeutic vaccine candidate against cancer. Microbiol Immunol 60(9):646–652
28. Rammensee H, Bachmann J, Emmerich NP, Bachor OA, Stevanovic S (1999) SYFPEITHI: database for MHC ligands and peptide motifs. Immunogenetics 50(3–4):213–219
29. Larsen MV, Lundegaard C, Lamberth K, Buus S, Lund O, Nielsen M (2007) Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. BMC Bioinformatics 8:424
30. Nielsen M, Lundegaard C, Lund O (2007) Prediction of HLA class II binding affinity using SMMalign, a novel stabilization matrix alignment method. BMC Bioinformatics 8:238
31. Bhasin M, Raghava GP (2004) SVM based method for predicting HLA-DRB1*0401 binding peptides in an antigen sequence. Bioinformatics 20(3):421–423
32. Singh H, Raghava GP (2001) ProPred: prediction of HLA-DR binding sites. Bioinformatics 17(12):1236–1237

33. Saha S, Raghava GP (2006) Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. Proteins 65(1):40–48
34. Saha S, Raghava GPS (2006) AlgPred: prediction of allergenic proteins and mapping of IgE epitopes. Nucleic Acids Res 34:W202–W209
35. Kurcinski M, Jamroz M, Blaszczyk M, Kolinski A, Kmiecik S (2015) CABS-dock web server for the flexible docking of peptides to proteins without prior knowledge of the binding site. Nucleic Acids Res 43(W1):W419–W424
36. Lohia N, Baranwal M (2018) Highly conserved hemagglutinin peptides of H1N1 influenza virus elicit immune response. 3 Biotech 8(12):492. https://doi.org/10.1007/s13205-018-1509-3
37. Arango MT, Perricone C, Kivity S, Cipriano E, Ceccarelli F, Valesini G, Shoenfeld Y (2017) HLA-DRB1 the notorious gene in the mosaic of autoimmunity. Immunol Res 65(1):82–98
38. Nikbin B, Nicknam MH, Hadinedoushan H, Ansaripour B, Moradi B, Yekaninejad M, Aminikhah M, Ranjbar MM, Amirzargar A (2017) Human leukocyte antigen (HLA) class I and II polymorphism in Iranian healthy population from Yazd Province. Iran J Allergy Asthma Immunol 16(1):1–13
39. Wieczorek M, Abualrous ET, Sticht J, Álvaro-Benito M, Stolzenberg S, Noé F, Freund C (2017) Major histocompatibility complex (MHC) class I and MHC class II proteins: conformational plasticity in antigen presentation. Front Immunol 8(292). https://doi.org/10.3389/fimmu.2017.00292
40. Blaszczyk M, Kurcinski M, Kouza M, Wieteska L, Debinski A, Kolinski A, Kmiecik S (2016) Modeling of protein-peptide interactions using the CABS-dock web server for binding site search and flexible docking. Methods 93:72–83
41. Li W, Joshi MD, Singhania S, Ramsey KH, Murtl AK (2014) Peptide vaccine: progress and challenges. Vaccines 2:515–536
42. Francis JN, Bunce CJ, Horlock C, Watson JM, Warrington SJ, Georges B, Brown CB (2015) A novel peptide-based pan-influenza A vaccine: a double blind, randomised clinical trial of immunogenicity and safety. Vaccine 33(2):396–402
43. Zom GG, Welters MJP, Loof NM et al (2016) TLR2 ligand-synthetic long peptide conjugates effectively stimulate tumor-draining lymph node T cells of cervical cancer patients. Oncotarget 7(41):67087–67100
44. Agallou M, Athanasiou E, Koutsoni O, Dotsika E, Karagouni E (2014) Experimental validation of multi-epitope peptides including promising MHC Class I- and II-restricted epitopes of four known Leishmania infantum proteins. Front Immunol 5(268). https://doi.org/10.3389/fimmu.2014.00268
45. Vani J, Shaila MS, Chandra NR, Nayak R (2006) A combined immuno-informatics and structure-based modeling approach for prediction of T cell epitopes of secretory proteins of Mycobacterium tuberculosis. Microbes Infect 8(3):738–746
46. Lohia N, Baranwal M (2015) Identification of conserved peptides comprising multiple T cell epitopes of matrix 1 protein in H1N1 influenza virus. Viral Immunol 28(10):570–579
47. Dikhit MR, Kumar A, Das S, Dehury B, Rout AK, Jamal F, Sahoo GC, Topno RK, Pandey K, Das VNR, Bimal S, Das P (2017) Identification of Potential MHC Class-II-restricted epitopes derived from Leishmania donovani antigens by reverse vaccinology and evaluation of Their CD4 + T-cell responsiveness against visceral leishmaniasis. Front Immunol 8(1763). https://doi.org/10.3389/fimmu.2017.01763
48. Yassin GM, Amin MA, Attia AS (2016) Immunoinformatics identifies a lactoferrin binding protein a peptide as a promising vaccine with a global protective prospective against moraxella catarrhalis. J Infect Dis 213(12):1938–1945
49. Sundar K, Boesen A, Coico R (2007) Computational prediction and identification of HLA-A2.1-specific Ebola virus CTL epitopes. Virology 360(2):257–263
50. Dutta DK, Rhodes K, Wood SC (2015) In silico prediction of Ebola Zaire GP(1,2) immuno-dominant epitopes for the Balb/c mouse. BMC Immunol 16:59. https://doi.org/10.1186/s12865-015-0126-8
51. Dikhit MR, Kumar S, Vijaymahantesh et al (2015) Computational elucidation of potential antigenic CTL epitopes in Ebola virus. Infection Genetics and Evolution 36:369–375
52. Ruibal P, Oestereich L, Lüdtke A et al (2016) Unique human immune signature of Ebola virus disease in Guinea. Nature 533(7601):100–104
53. Theaker SM, Rius C, Greenshields-Watson A, Lloyd A, Trimby A, Fuller A, Miles JJ, Cole DK, Peakman M, Sewell AK, Dolton G (2016) T-cell libraries allow simple parallel generation of multiple peptide-specific human T-cell clones. J Immunol Methods 430:43–50
54. Gupta M, Greer P, Mahanty S, Shieh WJ, Zaki SR, Ahmed R, Rollin PE (2005) CD8-Mediated Protection against Ebola Virus Infection Is Perforin Dependent. J Immunol 174(7):4198–4202
55. Simmons G, Lee A, Rennekamp AJ, Fan X, Bates P, Shen H (2004) Identification of murine T-cell epitopes in Ebola virus nucleoprotein. Virology 318(1):224–230
56. Wilson JA, Hart MK (2001) Protection from Ebola virus mediated by cytotoxic T lymphocytes specific for the viral nucleoprotein. J Virol 75(6):2660–2664
57. Fritsche PJ, Helbling A, Ballmer-Weber BK (2010) Vaccine hypersensitivity—update and overview. Swiss Med Wkly 140(17–18):238–246
58. Cohen AD, Shoenfeld Y (1996) Vaccine-induced autoimmunity. J Autoimmun 9(6):699–703
59. Meyboom RH, Fucik H, Edwards IR (1995) Thrombocytopenia reported in association with hepatitis B and A vaccines. Lancet 345(8965):1638
60. Topaloglu H, Berker M, Kansu T, Saatci U, Renda Y (1992) Optic neuritis and myelitis after booster tetanus toxoid vaccination. Lancet 339(8786):178–179
61. Patja A, Mäkinen-Kiljunen S, Davidkin I, Paunio M, Peltola H (2001) Allergic reactions to measles-mumps-rubella vaccination. Pediatrics 107(2):E27
62. Businco L (1994) Measles, mumps, rubella immunization in egg-allergic children: a long-lasting debate. Ann Allergy 72(1):1–3
63. Atsmon J, Kate-Ilovitz E, Shaikevich D, Singer Y, Volokhov I, Haim KY, Ben-Yedidia T (2012) Safety and immunogenicity of multimeric-001–a novel universal influenza vaccine. J Clin Immunol 32(3):595–603
64. Hou Y, Guo Y, Wu C, Shen N, Jiang Y, Wang J (2012) Prediction and identification of T cell epitopes in the H5N1 influenza virus nucleoprotein in chicken. PLoS One 7(6):e39344
65. Tawar RG, Duquerroy S, Vonrhein C et al (2009) Crystal structure of a nucleocapsid-like nucleoprotein-RNA complex of respiratory syncytial virus. Science 326(5957):1279–1283
66. Yabukarski F, Lawrence P, Tarbouriech N, Bourhis JM, Delaforge E, Jensen MR, Ruigrok RW, Blackledge M, Volchkov V, Jamin M (2014) Structure of Nipah virus unassembled nucleoprotein in complex with its viral chaperone. Nat Struct Mol Biol 21(9):754–759
67. Alayyoubi M, Leser GP, Kors CA, Lamb RA (2015) Structure of the paramyxovirus parainfluenza virus 5 nucleoprotein-RNA complex. Proc Natl Acad Sci USA 112(14):E1792–E1799
68. He Y, Li J, Mao W et al (2018) HLA common and well-documented (CWD) alleles in China. HLA https://doi.org/10.1111/tan.13358
69. Robinson J, Halliwell JA, Hayhurst JD, Flicek P, Parham P, Marsh SG (2015) The IPD and IMGT/HLA database: allele variant databases. Nucleic Acids Res 43:423–431

70. Kalyanaraman N (2018) In silico prediction of potential vaccine candidates on capsid protein of human bocavirus 1. Mol Immunol 93:193–205
71. Watanabe S, Noda T, Kawaoka Y (2006) Functional mapping of the nucleoprotein of Ebola virus. J Virol 80(8):3743–3751
72. Albertini AA, Wernimont AK, Muziol T, Ravelli RB, Clapier CR, Schoehn G, Weissenhorn W, Ruigrok RW (2006) Crystal structure of the rabies virus nucleoprotein-RNA complex. Science 313(5785):360–363
73. Noda T, Hagiwara K, Sagara H, Kawaoka Y (2010) Characterization of the Ebola virus nucleoprotein-RNA complex. J Gen Virol 91(6):1478–1483
74. Medaglini D (2018) Correlates of vaccine-induced protective immunity against Ebola virus disease. Semin Immunol. https://doi.org/10.1016/j.smim.2018.07.003
75. Henao-Restrepo AM, Camacho A, Longini IM et al (2017) Efficacy and effectiveness of an rVSV-vectored vaccine in preventing Ebola virus disease: final results from the Guinea ring vaccination, open-label, cluster-randomised trial (Ebola Ça Suffit!). Lancet 389(10068):505–518
76. Halperin SA, Arribas JR, Rupp R, Andrews CP, Chu L, Das R, Simon JK, Onorato MT, Liu K, Martin J, Helmond FA (2017) Six-month safety data of recombinant vesicular stomatitis virus-zaire ebola virus envelope glycoprotein vaccine in a phase 3 double-blind, placebo-controlled randomized study in healthy adults. J infect Dis 215(12):1789–1798
77. Ledgerwood JE, DeZure AD, Stanley DA et al (2017) Chimpanzee Adenovirus Vector Ebola Vaccine. N Engl J Med 376(10):928–938
78. Ewer K, Rampling T, Venkatraman N et al (2016) A monovalent chimpanzee adenovirus ebola vaccine boosted with MVA. N Engl J Med 374(17):1635–1646

Identification of peptides containing epitopes of Ebola virus eliciting immune response

ORIGINALITY REPORT

9%	2%	9%	0%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	"Marburg- and Ebolaviruses", Springer Science and Business Media LLC, 2017 Publication	1%
2	"Encyclopedia of Signaling Molecules", Springer Nature, 2018 Publication	1%
3	Ruth E. Soria-Guerra, Ricardo Nieto-Gomez, Dania O. Govea-Alonso, Sergio Rosales-Mendoza. "An overview of bioinformatics tools for epitope prediction: Implications on vaccine development", Journal of Biomedical Informatics, 2015 Publication	<1%
4	"Novel Technologies for Vaccine Development", Springer Nature, 2014 Publication	<1%
5	Christopher F. Basler. "Innate immune evasion by filoviruses", Virology, 2015 Publication	<1%

6	link.springer.com Internet Source	<1%
7	Murray, Michael J.. "Ebola Virus Disease : A Review of Its Past and Present", Anesthesia & Analgesia, 2015. Publication	<1%
8	Sahil Jain, Manoj Baranwal. "Conserved peptide vaccine candidates containing multiple Ebola nucleoprotein epitopes display interactions with diverse HLA molecules", Medical Microbiology and Immunology, 2019 Publication	<1%
9	"Immunoinformatics", Springer Nature, 2014 Publication	<1%
10	Dong-Shan Yu, Tian-Hao Weng, Xiao-Xin Wu, Frederick X.C. Wang, Xiang-Yun Lu, Hai-Bo Wu, Nan-Ping Wu, Lan-Juan Li, Hang-Ping Yao. "The lifecycle of the Ebola virus in host cells", Oncotarget, 2017 Publication	<1%
11	Roman Kogay, Christian Schönbach. "Epitope Predictions", Elsevier BV, 2018 Publication	<1%
12	Jeffrey E. Lee, Marnie L. Fusco, Ann J. Hessell, Wendelien B. Oswald, Dennis R. Burton, Erica Ollmann Saphire. "Structure of the Ebola virus	<1%

glycoprotein bound to an antibody from a human survivor", Nature, 2008
Publication

13	Sette, Alessandro, Sinu Paul, Kerrie Vaughan, and Bjoern Peters. "The Use of the Immune Epitope Database to Study Autoimmune Epitope Data Related to Alopecia Areata", The Journal of Investigative Dermatology Symposium, 2015. Publication	<1%
14	Jackson Emanuel, Andrea Marzi, Heinz Feldmann. "Filoviruses: Ecology, Molecular Biology, and Evolution", Elsevier BV, 2018 Publication	<1%
15	Lohia, Neha, and Manoj Baranwal. "Identification of Conserved Peptides Comprising Multiple T Cell Epitopes of Matrix 1 Protein in H1N1 Influenza Virus", Viral Immunology, 2015. Publication	<1%
16	"Encyclopedia of Cancer", Springer Nature, 2017 Publication	<1%
17	dayfres.blogspot.com Internet Source	<1%
18	academic.oup.com Internet Source	<1%

19	www.frontiersin.org Internet Source	<1%
20	"Food Allergens", Springer Science and Business Media LLC, 2017 Publication	<1%
21	Shahram Nazarian, Seyed Latif Mousavi Gargari, Iraj Rasooli, Jafar Amani, Samane Bagheri, Masoome Alerasool. "An in silico chimeric multi subunit vaccine targeting virulence factors of enterotoxigenic Escherichia coli (ETEC) with its bacterial inbuilt adjuvant", Journal of Microbiological Methods, 2012 Publication	<1%
22	journals.plos.org Internet Source	<1%
23	H. Singh, G. P. S. Raghava. "ProPred: prediction of HLA-DR binding sites", Bioinformatics, 2001 Publication	<1%
24	in-silico-pharmacology.springeropen.com Internet Source	<1%
25	Lyn M. O'Brien, Margaret G. Stokes, Stephen G. Lonsdale, David R. Maslowski et al. "Vaccination with recombinant adenoviruses expressing Ebola virus glycoprotein elicits protection in the interferon alpha/beta receptor	<1%

knock-out mouse", Virology, 2014
Publication

26	Xingdong Yang. "An introduction to epitope prediction methods and software", Reviews in Medical Virology, 03/2009 Publication	<1%
27	Zhang, Z.W.. "Screening and identification of B cell epitopes of structural proteins of foot-and-mouth disease virus serotype Asia1", Veterinary Microbiology, 20100106 Publication	<1%
28	J. Figueredo. "Prediction of Cellular Immune Responses against CFTR in Patients with Cystic Fibrosis after Gene Therapy", American Journal of Respiratory Cell and Molecular Biology, 12/07/2006 Publication	<1%
29	Li C. Chong, Asif M. Khan. "Vaccine Target Discovery", Elsevier BV, 2019 Publication	<1%
30	Handbook on Immunosenescence, 2009. Publication	<1%
31	www.ncbi.nlm.nih.gov Internet Source	<1%
32	www.jove.com Internet Source	<1%

33	Kazi Ada, Candy Chuah, Abu Bakar Abdul Majeed, Chiuan Herng Leow, Boon Huat Lim, Chiuan Yee Leow. "Current progress of immunoinformatics approach harnessed for cellular- and antibody-dependent vaccine design", Pathogens and Global Health, 2018 Publication	<1%
34	Ondondo, Beatrice O.. "The influence of delivery vectors on HIV vaccine efficacy", Frontiers in Microbiology, 2014. Publication	<1%
35	Yasser EL-Manzalawy. "Predicting linear B-cell epitopes using string kernels", Journal of Molecular Recognition, 07/2008 Publication	<1%
36	www.nature.com Internet Source	<1%
37	Nakayama, Eri, and Masayuki Saijo. "Animal models for Ebola and Marburg virus infections", Frontiers in Microbiology, 2013. Publication	<1%
38	science.sciencemag.org Internet Source	<1%
39	www.pnas.org Internet Source	<1%
	V Brusic. "Bioinformatics for characterisation of	

40	allergens, allergenicity and allergic crossreactivity", Trends in Immunology, 2003 Publication	<1%
41	L. Guilherme, K. C. Faé, F. Higa, L. Chaves et al. "Towards a Vaccine Against Rheumatic Fever", Clinical and Developmental Immunology, 2006 Publication	<1%
42	clinicalmolecularallergy.biomedcentral.com Internet Source	<1%
43	www.news-medical.net Internet Source	<1%
44	Jeffrey E Lee. "*Ebolavirus* glycoprotein structure and mechanism of entry", Future Virology, 11/2009 Publication	<1%
45	www.oncotarget.com Internet Source	<1%
46	Helen McSparron, Martin J. Blythe, Christianna Zygouri, Irini A. Doytchinova, Darren R. Flower. "JenPep: A Novel Computational Information Resource for Immunobiology and Vaccinology", Journal of Chemical Information and Computer Sciences, 2003 Publication	<1%
47	P. Baldi. "PEPITO: improved discontinuous B-	

cell epitope prediction using multiple distance thresholds and half sphere exposure", Bioinformatics, 05/07/2008
Publication
<1%

48 Michael Jan, Shu Meng, Natalie C. Chen, Jietang Mai, Hong Wang, Xiao-Feng Yang. "Inflammatory and Autoimmune Reactions in Atherosclerosis and Vaccine Design Informatics", Journal of Biomedicine and Biotechnology, 2010
Publication
<1%

49 Salman Sadullah Usmani, Rajesh Kumar, Sherry Bhalla, Vinod Kumar, Gajendra P.S. Raghava. "In Silico Tools and Databases for Designing Peptide-Based Vaccine and Drugs", Elsevier BV, 2018
Publication
<1%

50 Brandsma, J.L.. "Reversal of papilloma growth in rabbits therapeutically vaccinated against E6 with naked DNA and/or vesicular stomatitis virus vectors", Vaccine, 20101206
Publication
<1%

51 Mire, Chad E., Thomas W. Geisbert, Heinz Feldmann, and Andrea Marzi. "Ebola virus vaccines – reality or fiction?", Expert Review of Vaccines, 2016.
Publication
<1%

52	www.redbooks.ibm.com Internet Source	<1%
53	Navid Nezafat, Younes Ghasemi, Gholamreza Javadi, Mohammad Javad Khoshnoud, Eskandar Omidinia. "A novel multi-epitope peptide vaccine against cancer: An in silico approach", Journal of Theoretical Biology, 2014 Publication	<1%
54	Prabdial-Sing, Nishi, Adrian J Puren, and Sheila M Bowyer. "Sequence-based in silico analysis of well studied Hepatitis C Virus epitopes and their variants in other genotypes (particularly genotype 5a) against South African human leukocyte antigen backgrounds", BMC Immunology, 2012. Publication	<1%
55	eprints.lib.hokudai.ac.jp Internet Source	<1%
56	Giorgio Minotti, Pierantonio Menna, Emanuela Salvatorelli, Gaetano Cairo, Luca Gianni. "Anthracyclines: Molecular Advances and Pharmacologic Developments in Antitumor Activity and Cardiotoxicity", Pharmacological Reviews, 2004 Publication	<1%
57	"Gene Vaccines", Springer Nature, 2012 Publication	<1%

58	E. M. B. Thiemann, L. Andersson, R. Lillis, P. Withers et al. "The Mars Topside Ionosphere Response to the X8.2 Solar Flare of 10 September 2017", Geophysical Research Letters, 2018 Publication	<1%
59	patents.justia.com Internet Source	<1%
60	onlinelibrary.wiley.com Internet Source	<1%
61	www.bmj.com Internet Source	<1%
62	Carmen Diaz-Amigo, Bert Popping. "Accuracy of ELISA Detection Methods for Gluten and Reference Materials: A Realistic Assessment", Journal of Agricultural and Food Chemistry, 2013 Publication	<1%
63	curis.ku.dk Internet Source	<1%
64	media.proquest.com Internet Source	<1%
65	Thomas W. Geisbert, Lisa E. Hensley. "Ebola virus: new insights into disease aetiopathology and possible therapeutic interventions", Expert	<1%

Reviews in Molecular Medicine, 2004

Publication

Exclude quotes	On	Exclude matches	< 8 words
Exclude bibliography	On		