

**EPITOPE MAPPING OF HCV STRUCTURAL PROTEINS
FOR POTENTIAL VACCINE DESIGN; AN *INSILICO*
APPROACH**



Submitted By

Sania Munir ID: S2016254005

Supervised By

Dr. Muhammad Sohail Afzal

**Department of Life Sciences
School of Science
University of Management and Technology,
Lahore, Pakistan
2018**

EPITOPE MAPPING OF HCV STRUCTURAL PROTEINS FOR POTENTIAL VACCINE DESIGN; AN *INSILICO* APPROACH

Submitted to University of Management and Technology Lahore

In partial fulfillment of the requirements

for the award of degree of

**MS
BIOTECHNOLOGY**

BY

SANIA MUNIR

ID

S	2	0	1	6	2	5	4	0	0	5
---	---	---	---	---	---	---	---	---	---	---

Session: 2016- 2018

**Department of Life Sciences
School of Science
University of Management and Technology,
Lahore, Pakistan**



بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

DECLARATION

I SANIA MUNIR, Student of MS. BIOTECHNOLOGY, ID S2016254005 aware of and understand the university's policy on plagiarism and I certify that this thesis titled **“Epitope Mapping of HCV Structural Proteins for Potential Vaccine Design; An *insilico* Approach”** is my own work, except where indicated by referencing, and the work presented in it has not been submitted in support of another degree or qualification from this or any other university or institute of learning.

The experimental work is almost entirely my own work; the collaborative contributions have been indicated clearly and acknowledged. Due references have been provided on all supporting literatures and resources.

Dated: _____

SANIA MUNIR

Thesis Similarity Report

RESEARCH COMPLETION CERTIFICATE

This is to certify that the thesis entitled “**Epitope Mapping of HCV Structural Proteins for Potential Vaccine Design; An *insilico* Approach**” submitted by **Sania Munir, ID: S2016254005** has been accepted towards the partial fulfillment of the requirement for MS degree in **Biotechnology**. The quantum and quality of the work contained in this thesis is adequate for the award of degree.

Supervisor

Dr. M. Sohail Afzal

Assistant Professor

Department of Life Sciences,

UMT, Lahore

External Examiner

Chairman

Dr. Nouman Rasool

Assistant Professor

Department of Life Sciences,

UMT, Lahore

Dean

Dr. Muhammad Azhar Iqbal

Professor

School of Science,

UMT, Lahore

DEDICATION

I dedicate
my work to
my beloved parents Mr. and Mrs. Muhammad Munir
and my husband Syed Khayyam Hassan Gillani
for their endless support during my thesis compilation.

ACKNOWLEDGMENTS

All the praises are for the Allah Almighty who bestowed me with the ability and potential to complete this Thesis Report. Words are very few to express my gigantic meek obligations to my affectionate Parents for their prayers and strong fortitude to enabling me to achieve this job.

My deep sense of gratitude to my project supervisor Dr. Muhammad Sohail Afzal (Assistant Professor, Department of Life Sciences, University of Management and Technology, Lahore) for his unending support and guidance.

My heartfelt thanks to Dr. Nouman Rasool (Chairperson Department of Life Sciences, University of Management and Technology, Lahore) for his ultimate advices and directions in becoming a disciplined worker.

I'm really thankful to Dr. Muhammad Ali (Assistant Professor, Department of Life Sciences, University of Management and Technology, Lahore) for his sincere guidance during the compilation of my thesis work.

Sania Munir

TABLE OF CONTENTS

Title	Page no.
List of Tables	IX
List of Figures	X
List of Abbreviations	XI
Abstract	XII
Chapter 1: Introduction and Review of Literature.....	
1.1: Hepatitis.....	1
1.2: Family Classification.....	1
1.3: A brief history, discover and Progress of HCV.....	3
1.4: HCV prevalence.....	5
1.4.1: Global Level	5
1.4.2: HCV prevalence in Pakistan.....	7
1.5: HCV genotypes.....	8
1.6: Liver disease Progression.....	10
1.7: Structure of HCV Virions.....	10
1.8: Core.....	12
1.8.1: What is Capsid?	12
1.8.2: Core biogenesis and Properties.....	12
1.9: Glycoproteins (E1, E2).....	13
1.10: HCV Epidemiology; A great mystery.....	14
1.11: Bioinformatics; An <i>in silico</i> tool.....	15
1.12 Vaccine for HCV.....	16
Chapter 2: Materials and Methods.....	18
2.1: Sequence Retrieval	18
2.1.1: NCBI.....	18
2.1.2: BLASTP.....	19
2.1.3: MSA.....	20
2.1.3.1: Clustal Omega.....	20
2.1.4: Jalview.....	21

2.2: Prediction of T- Cell Binding Epitopes.....	22
2.3: Immungenicity Prediction of MHC-I.....	22
2.3.1: ANN.....	22
2.3.2: HMM.....	22
2.3.3: SVM.....	23
2.3.4: Peptide Selection Criteria.....	23
2.4: Antigenicity Analysis.....	23
2.5: Structure Visualization.....	24
Chapter 3: Results.....	25
3.1: HCV Core Result.....	25
3.1.1: Consensus Sequence.....	25
3.1.2: IEDB.....	25
3.2: HCV E1 Result.....	31
3.2.1: Consensus Sequence	31
3.2.2: IEDB.....	31
3.3: HCV E2 Result	36
3.3.1: Consensus Sequence.....	36
3.3.2: IEDB.....	36
Chapter 4: Discussion.....	41
Chapter 5: References.....	44

LIST OF TABLES

Sr. No.	Title	Page No.
3.1	MHC-I Predicted T-cell Epitope sequences of HCV core protein on the basis of ANN and SMM Methods	26
3.2	MHC-I Predicted T-cell Epitope sequences of HCV Core protein with Antigenicity Score	28
3.3	MHC-I Predicted T-cell Epitope sequences of HCV E1 protein on the basis of ANN and SMM Methods	32
3.4	MHC-I Predicted T-cell Epitope sequences of HCV E1 protein with Antigenicity Score	33
3.5	MHC-I Predicted T-cell Epitope sequences of HCV E2 on the basis of ANN and SMM Methods	37
3.6	MHC-I Predicted T-cell Epitope sequences of HCV E2 protein with Antigenicity Score	38

LIST OF FIGURES

Sr No.	Title	Page No.
1.1	Classification of <i>Flaviviridae</i> family	2
1..2	HCV polyprotein structure	2
1.3	Discovery of HCV; First clone of HCV isolated by cDNA library screening of chronic infected chimpanzee plasma	4
1.4	A summary of HCV outbreaks	5
1.5	HCV infection prevalence globally and distribution of genotypes in recent days	7
1.6	HCV prevalence in different regions of Pakistan	8
1.7	HCV genotypes distribution in Pakistan	9
1.8	HCV; Liver disease progression	10
1.9	Structure of HCV virion	11
1.10	Overview of HCV life cycle	11
1.11	A schematic diagram for the prediction of epitopes for vaccine development	17
2.1	An overview of materials and methods	18
2.2	Blastp window showing default settings	20
2.3	A view of Clustal Omega result file	21
2.4	Vaxijen server 2.0 window with desired settings	24
3.1	Consensus sequence generated by Jalview software	25
3.2	3D model of hepatitis C virus generated by I-Tasser software	29
3.3	Epitope RLGVRATRK at position 43- 51 representing its 3D structure with the names of amino acids. Loop showing epitope location on the surface	30
3.4	Epitope GPRLGVRAT at position 41-49 presenting 3D model produced by software Swissprot, epitope is locating on the surface, facing hydrophilic potion.	30
3.5	Consensus sequence of HCV E1 made by Jalview software	31
3.6	3-dimensional model of gpE1	34
3.7	GpE1 epitope MMNWSPTTA model representing its amino acids to be faced at external side	35
3.8	Consensus sequence of HCV E2 made by Jalview software	35
3.9	Consensus sequence of HCV E2 made by Jalview software	36
3.10	3-dimensional model for gpE2	39
3.11	GpE2 epitope MFVGGFEHR model showing its peptide position	40
3.12	GpE2 epitope HTIGGAAAR model showing its amino acids present at outer surface	40

LIST OF ABBREVIATIONS

ANN	Artificial Neural Networks
BLAST	Basic Local Alignment Search Tool
CTL	Cytotoxic T lymphocytes
E1	Envelop Protein 1
E2	Envelop Protein 2
gpE1	Glycoprotein envelop protein 1
gpE2	Glycoprotein envelop protein 2
HMM	Hidden Markov Models
HLA	Human Leucocyte Antigen
HIV	Human Immune Deficiency virus
HBV	Hepatitis B virus
HAV	Hepatitis A virus
HCV	Hepatitis C virus
HCC	Hepatocellular Carcinoma
IC	Inhibitory Concentration
IEDB	Immune Epitope Database
MSA	Multiple Sequence Alignment
NANBH	Non-A non-B Hepatitis
NCBI	National Center for Biotechnology Information
PCR	Polymerase Chain Reaction
RdRp	RNA dependent RNA polymerase
SR-B1	Scavenger receptor class B type 1
SVM	Support vector Machines
VLDL	Very low density lipoprotein
WHO	World Health Organization

ABSTRACT

Hepatitis C virus (HCV) is a leading cause of morbidity and mortality globally. Recent studies have shown an increased burden (2.8%) of HCV infection, which is nearly equal to 185 million infections globally. Highest prevalence rated countries are located in Asia and Africa. Pakistan has the high HCV seroprevalence rates between 2.4% and 6.5% in general population. Due to error prone nature, HCV has the ability to escape immune system surveillance. High variation in genome is the primary hurdle in vaccine development against the virus. After the infection host, Cytotoxic T- lymphocytes contribute their role in the eradication of infection. For successful eradication of HCV a vaccine development is a need of the hour. HCV Structural proteins are one of the potential immunogenic viral proteins. The study is aimed to analyze the core, E1 and E2 protein as a potential vaccine candidate. In this study the sequences of HCV structural proteins were taken and comparison was made by BLASTP to find out maximum potential candidates for a vaccine development. All the sequences were aligned by using CLUSTAL OMEGA software and consensus sequence was developed by using JALVIEW software. Immune Epitope Database (IEDB) was used to predict epitope sequences. The predicted epitopes were analyzed to determine their antigenicity through Vaxijen server version 2.0. I- Tasser software further used for the visualization of predicted epitopes positions and immunogenicity.

Key Words: HCV; Core protein; Immunity; Epitope; Vaccine; *Insilico*