ALPHAVIRUS CAPSID PROTEASE AND ITS INHIBITORS

Ph.D THESIS

by

BENAZIR FATMA

DEPARTMENT OF BIOTECHNOLOGY
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE
ROORKEE-247667(INDIA)
MAY, 2019

भारतीय प्रौद्योगिकी संस्थान रुड़की
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

ALPHAVIRUS CAPSID PROTEASE AND ITS INHIBITORS

A THESIS

Submitted in partial fulfilment of the
requirements for the award of the degree

of

DOCTOR OF PHILOSOPHY

in

BIOTECHNOLOGY

by

BENAZIR FATMA

DEPARTMENT OF BIOTECHNOLOGY
INDIAN INSTITUTE OF TECHNOLOGYROORKEE
ROORKEE-247667(INDIA)
May, 2019

भारतीय प्रौद्योगिकी संस्थान रुड़की
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

भारतीय प्रौद्योगिकी संस्थान रुड़की
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

©INDIAN INSTITUTE OF TECHNOLOGY ROORKEE, ROORKEE- 2019
ALL RIGHTS RESERVED

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE
ROORKEE

CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in the thesis entitled "**ALPHAVIRUS CAPSID PROTEASE AND ITS INHIBITORS**" in partial fulfilment of the requirements for the award of the degree of Doctor of Philosophy and submitted in the Department of Biotechnology of the Indian Institute of Technology Roorkee, Roorkee is an authentic record of my own work carried out during a period from July, 2013 to May, 2019 under the supervision of Dr. Shailly Tomar, Assistant Professor, Department of Biotechnology, Indian Institute of Technology Roorkee, Roorkee.

The matter presented in the thesis has not been submitted by me for the award of any other degree of this or any other institution.

BENAZIR FATMA

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

(Shailly Tomar)
Supervisor

Dated: ……………………...

The Ph.D. Viva-Voce Examination of **Ms. Benazir Fatma**, Research Scholar, has been held on ***16th September, 2019*** at Department of Biotechnology, IIT Roorkee, Roorkee, INDIA.

Chairman, SRC **Signature of External Examiner**

This is to certify that the student has made all the corrections in the thesis.

Signature of Supervisor **Head of Department**

ABSTRACT

Alphaviruses are distributed worldwide and has acquired infamy by being health threat to humans and livestock. Almost 20 different alphaviruses are categorized as medically important viruses as the infection accompanies arthritis, encephalitis, rashes and fever. Alphavirus infection is transmitted to humans majorly by the mosquito bite hence alphavirus is categorized among arbovirus. Ross River virus (RRV), Sindbis virus (SINV), Semliki Forest virus (SFV) and Venezuelan Equine encephalitis virus (VEEV) are exploited as viral vectors for gene delivery therefore serve as important bait in gene therapy. Aura virus (AURAV) is an alphavirus which displays similarity with SINV in terms of structure organization and gene sequence. The nucleotide sequence of promoter of AURAV bears striking similarity with the promoter of SINV. AURAV has not yet gained notoriety for causing infections in humans but it is serologically similar to pathogenic WEEV and SINV so it is quite relevant to study the virus. Chikungunya virus (CHIKV) has created menace time and again worldwide especially in tropical and temperate region. The re-emergence of CHIKV is a matter to be pondered upon especially for developing nations where basic sanitation facilities get compromised because of their economy crunch. The prime source of virus spread is by mosquito bite. CHIKV infection is accompanied by fever, severe joint pain, muscle pain, headache, and nausea. Alphavirus is categorized in the group IV Togaviridae family of viruses. Structurally alphaviruses are enveloped viruses which exhibits T=4 icosahedral symmetry. Diameter of alphavirus is approximately 70nm. A mature virion of alphavirus acquires lipid membrane from the host. Alphavirus exhibits protruded spikes which are generated by the assembly of trimer of heterodimer of the viral glycoprotein E1 and E2. The envelope E1 and E2 proteins localised in transmembrane has its C-terminal exposed to cytoplasm via which nucleocapsid interacts and budding of virus is promoted. The genome of alphavirus is a 49S single-stranded, positive-sense, RNA approximately 12 kb in length. Along with the genomic RNA, alphaviruses make 26S subgenomic RNAs which encode for the viral structural proteins. Translation of the subgenomic RNA generates structural polyprotein and capsid protein (CP) is part of it. The CP cleaves itself from the N-terminus of structural polyprotein owing to its cis-proteolytic activity which is confined in the C-terminus subdomain. Post cleavage of the CP from its polyprotein, the autoproteolytic activity halts as the C-terminus Trp (conserved) residue binds to the S1 specificity (active site) pocket blocking the active site of the serine CP protease. The cleaved structural polyprotein is processed by the host cell protease to generate E3, E2, 6K and E1

structural proteins. All these structural proteins along with CP play a vital role in the virus budding and encapsidation of the viral genome RNA.

The cleaved CP interacts with the genomic RNA through its N-terminal domain which is laden with basic amino acidresidues. This interaction is highly specific and mediates encapsidation of genomic RNA and nucleocapsid (NC) assembly. The N-terminal domain of CP is highly disordered and till date no crystal structure comprising the domain has been determined yet. There is alarming need to study the N-terminal domain of alphavirus CP for understanding RNA and capsid interactions at the atomic level. The NC is enclosed by the phospholipid bilayer comprising trimeric spikes constituted by the assembly of heterodimer of E1 and E2. The CP has a small hydrophobic pocket present in the C-terminal domain via which the cytoplasmic tail of E2 (cdE2) binds and promotes the budding of virus. Since CP plays a prominent role in the viral life cycle by being involved in viral RNA encapsidation, structural polyprotein processing and in virus budding, therefore is a potential target for antiviral drug development.

Chapter 1 is about literature reviewed on alphavirus. It briefs about the alphavirus life cycle, genome organisation of alphavirus, functional significance of non-structural and structural proteins of alphavirus. The chapter details about CP, its structural and functional role in viral assembly and budding. The chapter focuses on capsid protease as potential antiviral drug target and brief significance of repurposing of FDA approved drugs.

Chapter 2 comprises of three sections, where the first section details about the expression, purification and crystallization of the C-terminal truncated; protease active form of chikungunya virus capsid protein (CVCP active, 110-259 amino acid residues). Due to the unsuccessful crystallization attempts of CVCP active, the conserved C-terminal residue W261 was computationally removed from the coordinate file of the crystal structure of CVCP native (inactive proteinPDB ID: 5H23) for making the active site accessible for the docking of compounds from the library of FDA approved drugs. The second part details about the *in vitro* study of selected compounds which were best hits based on the binding energies against the active site of CVCP. The inhibitory constants (*K*i) of the compounds and their mode of inhibitions were determined using fluorogenic peptide substate having scissile bond W-S in a FRET- based protease assay. The 3rd section describes the antiviral effects of these CVCP inhibitors in the cell-based antiviral assays. From the library of FDA approved drugs, four compounds showed best hits against the active site of CVCP by molecular docking. These compounds inhibited the proteolytic activity of CVCP in the FRET based assay and showed different mode of inhibitions; two compounds showed competitive while the rest two showed distinct mixed and non-competitive inhibition and the calculated *K*i were in the micromolar range. In the cell-based antiviral assays, these compounds showed antiviral potential against CHIKV. Plaque reduction assay, q-RT-PCR, western blot analysis and immunofluorescence assay showed these compounds target the CHIKV life cycle at the later stage of cell infection.

Chapter 3 attempts to study the structural and functional perspectives of the N-terminal domain of alphavirus CP. In number of crystal structures, the N-terminus of the alphavirus CP is disordered [1], [2], [3], [4], [5], [6].This chapter is categorised into three sections; the first part aims to crystallize and determine the three-dimensional structure of the N- terminal domain of CVCP and AVCP. This section includes the cloning and expression of AVCP19 (19-265 amino acid residues) and CVCP20 (20-259 amino acid residues). The purification of CVCP20 was unsuccessful but AVCP19 was purified by immobilized metal affinity chromatography (IMAC) using phosphate buffer. Co-crystallization AVCP19 with 12-mer DNA was done which was again not sucessful. Also co-crystallisation of AVCP80 (80-267 amino acid residues) done with 12-mer DNA is in accordance to the literature surveyed which shows in SINVCP, 81-113 residues are involved in specific binding of CP with the genomic RNA and mediates the formation of NCP cores [7]. The crystals diffracted to 1.2 Å but neither electron density of the N-terminal domain was found nor was the density for 12-mer DNA detected. The second section includes the biophysical studies of interaction of nucleic acid with purified AVCP19. The CP-DNA interactions were studied by size exclusion chromatography and differential scanning calorimetry (DSC). Size exclusion chromatography confirmed formation of CP-DNA complex and DSC showed shift in the Tm of AVCP19 bound DNA which suggested the interaction provides stability to the domain. The third section describes an attempt to study the effect of DNA binding on the proteolytic activity of active CVCP and AVCP. The conserved C-terminal Trp residue of both AVCP19 and CVCP20 constructs were deliberately removed while making the constructs to study the *trans*- protease activity of these in presence of DNA oligos.

Chapter 4 concludes the thesis topic. The structure-assisted drug-repositioning approach has been used to identify potent antiviralmolecules for chikungunya therapeuticswhere the potential inhibitors of CHIKV CP protease were identified by virtual screening of compounds from the library of FDA approved drugs. The results of fluorogenic substarte peptide-based FRET assay and cell-based antiviral assays with the best hits compounds validated the above finding and led to the conclusion that alphavirus capsid protease could be a potential antiviral drug target. For the first time DNA-CPcomplex was purified by size-exclusion chromatography and stability of the complex was determined by DSC technique. The increased shift in the Tm of ‘DNA bound AVCP19’ showed the nucleic acid interaction with the N-terminal domain of alphavirus CP provides stability to the disordered region. Crystal structure of nucleic acid bound alphavirus CP can certainly help in studying the three-dimensional structure and molecular contact details of CP and nucleic acid.

भारतीय प्रौद्योगिकी संस्थान रुड़की
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

DEDICATED TO MY GRANDPARENTS AND PARENTS

भारतीय प्रौद्योगिकी संस्थान रुड़की
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

भारतीय प्रौद्योगिकी संस्थान रुड़की
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

ACKNOWLEDGMENTS

The current thesis might never have seen the light of the day without the constant effort, patience and valuable time put forth by my supervisor Dr. Shailly Tomar who painstakingly introduced me to the field of Molecular Virology, Biochemistry, Biophysics and Structural Biology. She nurtured in me the ability for independent work. Her magnanimous generosity and unflagging interest has helped me in all endeavors. I am extremely grateful to her by providing me with all the resources I required for my thesis work. Her untiring efforts of taking lab meetings,checking my weekly report and guiding me in troubleshooting the problems has literally helped me sail through my Ph.D. I owe her immensely for providing me oppurtunity to work in her well furnished lab and helping me in gauging progress at every stage. The guidance and support which I got from my mentor cum teacher is unparalled.

It takes me immense pleasure in expressing my sincere gratitude to Proffesor Pravindra Kumar for his timely suggestions in structural studies and providing facility of macrocrystallography unit. His sincere effort in offering suggestions in crystallography and bioinformatics is quite commendable. He has taught us crystallography very lucidly in our PhD coursework and kept everybody in high spirits by his inspiring scientific outlook.

I am extremely thankful to the members of my student research committee Dr. A. K. Sharma (chairman SRC), Professor R. P. Singh (internal expert) and Professor Ramesh Chandra (external expert) for their inquisitive questions and valuable suggestions during my semester's progress report presentations. Their profound effort of developing in me skill to answer the scientific queries has been extremely rewarding.

I heartily acknowledge the faculty members of the department for their selfless guidance and cooperation towards the students. My sincere thanks to the faculty members whose classes I took in my Ph.D coursework. I owe my sincere thanks to Dr. A. K. Sharma (Current), Professor Partha Roy (Former) and Professor Ramasare Prasad (Former) Head, Department of Biotechnology for ensuring and improvising constant basic facilities in the department. I heartily acknowledge the sincere efforts of the office staff members at Biotechnology Department of IIT Roorkee for managing and handling the official data of student.

My humble gratitude to my doting lab seniors Dr. Megha Aggarwal, Dr. Manju Narwal, Dr. Monu Batra, Dr. Rajesh Sharma, Dr. Ramanjit Kaur, Dr. Harvijay Singh and Dr. Rajat Mudgal for getting me acquainted to the lab culture during my intial phase of PhD. They all had been supportive with their rationale ideas and suggestions. I owe my senior Dr. Megha Aggarwal for teaching me protein-purification in my first year of PhD. My sincere thanks to Dr. Rajat Mudgal and Dr. Ramanjit Kaur for making me learn molecular cloning and helping me whenever I had any query or any glitch in my experiments. I got my hands- on experience in my FRET- based experiments from my seniors Dr. Harvijay Singh and Dr. Rajesh Sharma which was very generous of them. My heartfelt thanks to the post doc of our lab, Dr. Ravi Kumar for teaching me basic techniques of cell culture patiently. I would also like to acknowledge the research scholars of ‘structural biology& protein engineering lab and ‘protein biochemistry lab’. My sincere thanks to Dr. Bibekanand for patiently introducing me to the DSC technique. I would also like to acknowledge Dr. Pooja for her help in bioinformatics. I owe my sincere thanks to Neetu in helping me in solving crystal structure of AVCP80 tirelessly day and night. I am extremely grateful to my cordial and cheerful lab mates Dr. Anjali, Ms. Vedita, Mr. Akshay, Ms. Supreeti, Ms. Ruchi, Ms. Shweta, Mr. Sanket and Mr. Mandar. I am extremely thankful to my dear batch mates Dr. Surinder, Dr. Poonam, Dr. Anjali, Dr. Deepa, Dr. Nishu, Jyoti, Mukta, Lipi and Ankita for always keeping me cheerful with their gregarious and helpful attitude. My sincere thanks to Ms. Ruchi and Mr. Akshay for tirelessly proofreading my thesis.Finally my heartfelt thanks to my dearest Nanajaan, Naniammi, Dadajaan, Dadiammi, my parents, and my sister for always encouraging and supporting me in every sphere of life and for their selfless profound love for me. I am highly indebted to them. I am equally thankful to the other members of my family both paternal and maternal for always cheering me up and being affectionate and caring. I am extremely inspired by my uncle Dr. Javed Khan and aunt Dr. Shemaila. It is because of them I got motivation to pursue my career in Science and Research. I am grateful to them for their constant motivation. I am extremely thankful to my sister Aaisha and my cousin Faiz. I would like to acknowledge my husband Dr. Shafi Imran for his emotional and moral support. Finally I would like to acknowledge Ministry of Human Resource Development, India for the financial assistantship in my Ph.D.

Last but never the least my humble gratitude to almighty for blessing me and making me feel content with my life.

LIST OF PUBLICATIONS

Research Publications:

1. Sharma R, **Fatma B**, Saha A, Bajpai S, Sistla S, Dash PK, Parida M, Kumar P, Tomar S. Inhibition of chikungunya virus by picolinate that targets viral capsid protein. Virology, 2016 Nov 1;498:265-76.
2. **Fatma B**, Kumar R , Sharma R , Kesari P and Tomar S* Structure-based identification and evaluation of antivirals targeting Chikungunya virus capsid protease.(Communicated in ACS Infectious Diseases)

Book Chapter

Tomar S, Mudgal R, **Fatma B**. Flavivirus Protease: An Antiviral Target. In Viral Proteases and Their Inhibitors 2018 (pp. 137-161).

LIST OF CONFERENCES

Conference papers:

- **Fatma B.**, Aggarwal M., Tomar S. "Studies of N-Terminal domain of alphavirus capsid protein". RASBDD (Recent Advances in Structural Biology & Drug Discovery) held at IIT Roorkee, Roorkee, India (October 9-11, 2014) (Page 54).
- **Fatma B.**, Kumar R. , Sharma R. , Kesari P. and Tomar S.* "**Studies of Antiviral Molecules targeting Chikungunya virus specific serine protease"** Intervirocon 2018 held at PGIMER, Chandigarh.Awarded with a certificate as the second best poster presentation.

Seminars/ Courses/ Workshop attended:

- Participated in the GIAN course Methods and Techniques in Integrated Structural Biology: Towards Structure Based Drug Development held at IIT Roorkee, Roorkee, India (Jan 15-21, 2018)
- Participated in the GIAN course Recent Advancements in Biophysical Techniques and Virology held at IIT Roorkee, Roorkee, India (April 15-21, 2018)

CONTENTS

Page number:

CANDIDATE'S DECLARATION
ABSTRACT I
ACKNOWLEDGEMENT VII
LIST OF PUBLICATIONS AND CONFERENCES XI
CONTENTS XV
LIST OF FIGURES XXI
LIST OF TABLES XXIII
ABREVIATIONS XXIV

CHAPTER 1 Review of literature **1**
1.1 Introduction 1
1.2 Alphavirus life cycle 6
1.3 Genome organization of alphavirus 7
1.4 Non-structural proteins 9
1.5 Structural proteins 12
1.6 Alphavirus capsid structure 14
1.7 The *trans*- activity of alphavirus CP protease 16
1.8 Capsid protein as essential antiviral drug target 17
1.9 Repurposing drugs against CHIKV 17

CHAPTER 2 Structure-based screening of CVCP protease inhibitors for drug repurposing **18**

2.1 Chapter summary 18
2.2 Introduction 19
2.3 Materials and methods 21
2.3.1 Expression of CVCP active 21

2.3.2 Purification of CVCP active 22
2.3.3 Crystallisation trial of CVCP active 22
2.3.4 Structure-based screening and molecular docking 23
2.3.5 Sequence and structure analysis 23
2.3.6 CVCP serine protease assay 24
2.3.7 CVCP serine protease inhibition assay 24
2.3.8 Cells, virus and compounds 25
2.3.9 Cytotoxicity studies 25
2.3.10 Assessment of antiviral activity 25
2.3.11 Plaque based reduction assay 26
2.3.12 Inhibitory potency 26
2.3.13 Effect of inhibitors on viral RNA synthesis 27
2.3.14 Western immunoblot analysis 27
2.3.15 Immuno-fluorescence assay 28
2.3.16 Quantitative real-time PCR (qRT-PCR) 28
2.4 Results 29
2.4.1 Expression of CVCP active (110-259 amino acid residues) 29
2.4.2 Purification of CVCP active by Ni-NTA column 30
2.4.3 Reverse Ni -NTA of CVCP active 30
2.4.4 Crystallisation trial of CVCP active 31
2.4.5 Molecular Docking 32
2.4.6 FRET based CVCP inhibition assay 39
2.4.7 Cytotoxicity studies 41
2.4.8 Assessment of antiviral activity by plaque assay 41
2.4.9 Inhibitory potency 43
2.4.10 Effects of inhibitors on viral RNAsynthesis 44
2.4.11Western immunoblot analysis 45
2.4.12 Immunofluorescence assay 46
2.4.13 Quantitative real-time PCR 47
2.4 Conclusion 48

CHAPTER 3 Structural and functional perspective of the N-terminal CP domain of AVCP and CVCP 51

3.1 Chapter Summary 51

3.2 Introduction 51

3.3 Materials and methods 53

3.3.1 Cloning of CVCP 20 (20-259) 53

3.3.2 Expression of CVCP20 54

3.3.3 Purification of CVCP20 54

3.3.3 Cloning of AVCP19 (19-265) 55

3.3.4 Expression and solubility of AVCP19 56

3.3.5 Purification of AVCP19 57

3.3.6 Co- crystallisation trial of AVCP 19-12 mer DNA 57

3.3.7 Purification of AVCP80 58

3.3.8 Co- crystallisation trial of AVCP 80- 12 mer DNA 59

3.3.9 Size exclusion chromatography to study AVCP 19-(48 mer) DNA interactions 59

3.3.10 Differential scanning calorimetry to study AVCP19-(48mer)DNA interactions 59

3.3.11 *Trans*- protease activity of AVCP19 60

3.3.12 *Trans*- protease activity of AVCP19 bound 48-mer DNA 61

3.4 Results 61

3.4.1 Cloning of CVCP 20 61

3.4.2 Expression and purification of CVCP 20 62

3.4.3 Cloning of AVCP 19 63

3.4.4 Expression and purification of AVCP 19 64

3.4.5 Co- crystallisation trial of AVCP 19 with 12- mer DNA 65

3.4.6 Purification of AVCP 80 65

3.4.7 Co- crystallisation trial of AVCP 80 with 12 mer- DNA 66

3.4.8 Size exclusion chromatography to study AVCP-19 DNA interactions 70

3.4.9 Differential scanning calorimetry to study the shange in the stability of N-teminal domain of AVCP 19 when bound to DNA (48 mer) 73

3.4.10 Trans –protease activity of AVCP19 74
3.4.11 Trans–protease activity of AVCP19 bound 48mer DNA 75
3.5 Conclusion 76

CHAPTER 4 Discussion (Alphavirus capsid- an essential antiviral drug target) 78
4.1 Structural and functional significance of N-terminal domain of alphavirus CP 78
4.2 Structure based screening and evaluation of C-terminal domain of CVCP protease inhibitors for drug repurposing 80

REFERENCES 83

LIST OF FIGURES

Figure. No.	Particulars	Page No.
1.2.1	Graphical representation of alphavirus transmission cycle	6
1.2.2	Illustrative represntation of alphavirus life cycle	8
1.3.1	Genome organization of alphavirus	9
1.4.1	Non-structural polyprotein processing	11
1.5.1	Processing of alphavirus structural polyprotein	12
1.5.2	Representation of functional domains of SINV capsid protein	13
1.6.1	Cartoon view representation of crystal structure of AVCP (PDB ID: 4AGK).	15
2.2.1	Substrate residues of the CP protease from the different members of the genus alphavirus	20
2.4.1.1	SDS PAGE analysis of expression of CVCP active	29
2.4.2.1	SDS PAGE analysis of purification of CVCP active	30
2.4.2.2	SDS PAGE profile of reverse Ni-NTA showing cleaved CVCP active	31
2.4.4.1	SDS PAGE profile of concentrated CVCP active	32
2.4.5.1	Binding analysis of substrate into AVCP and CVCP protease active site	33
2.4.5.2	Molecular interactions of EAC, DRV, AP4 and PSU with CVCP active site shown by LigPlot representation	36
2.4.6.1	Double reciprocal plots of CVCP protease inhibition by the LOPAC® screened compounds	40
2.4.8.1	Assessment of the compound cytotoxicity by MTT assay (a) and Inhibition of CHIKV with (b) AP4, (c) EAC and (d) PSU compounds in Vero cells by plaque assay	42
2.4.9.1	Evaluation of anti-CHIKV activity of (a) AP4, (b) EAC and (c) PSU compound in Dose-dependent manner	43
2.4.10.1	Agarose gel electrophoresis profile of RT-PCRproducts of 12 hpi.	44

2.4.11.1	**Western blot band quantification raw data**	**45**
2.4.11.2	**Analysis of CHIKV CP expression in Vero cells**	**46**
2.4.12.1	**Effect of AP4, EAC and PSU treatment on CHIKV infection was assessed by Immuno-fluorescence assay**	**47**
2.4.13.1	**Reduction of CHIKV RNA in Vero cells by AP4, PSU and EAC in post treatment assay**	**48**
3.4.1.1	**Cloning confirmation of CVCP20 by PCR amplification and RE digestion by *Nde*I & XhoI PCR**	**62**
3.4.2.1	**SDS PAGE (15%) profile of expression and purification of CVCP20**	**63**
3.4.3.1	**Cloning of AVCP 19 (residues 19-265)**	**64**
3.4.4.1	**15 % SDS PAGE of the purification profile of AVCP19**	**65**
3.4.6.1	**SDS PAGE profile of purification of AVCP80 and reverse Ni-NTA of AVCP 80**	**66**
3.4.7.1	**Crystal image of AVCP 80-12mer DNA and Diffraction pattern of crystal at 1.5 Å**	**67**
3.4.7.2	**Structure of AVCP 80 -12 mer (DNA) complex**	**68**
3.4.7.3	**Cartoon view representation of AVCP 80 superimposed on AVCP 110**	**68**
3.4.8.1	**Gel filtration chromatography of AVCP19**	**70**
3.4.8.2	**Gel filtration chromatography of AVCP19 - 48 mer DNA showing the eluion volumes peaks**	**71**
3.4.8.3	**SDS PAGE profile of eluted samples of gel filtration of AVCP 19**	**72**
3.4.8.4	**Agarose gel profile of eluted samples of gel filtration of AVCP 19**	**72**
3.4.9.1	**Thermogram of AVCP19 & AVCP19-48mer DNA using differential Scanning calorimetry**	**73**
3.4.10.1	**3.4.10.1 Kinetic analysis of AVCP 19**	**75**
3.4.10.2	**Kinetic analysis of AVCP 19 bound 48-mer DNA**	**76**

LIST OF TABLES

1.1	Compilation of crystal structures of different members of alphavirus family	16
2.1	Analysis of molecular interactions of substrate (WS) binding into the active site of AVCP and CVCP protease	34
2.2	List of best hits obtained from screening of LOPAC^{1280} library	35
2.3	Hydrogen bond interactions analysis of inhibitors in the active site of AVCP and CVCP protease	37
2.4	Analysis of hydrophobic interactions of inhibitory molecules with the active site of AVCP and CVCP protease	38
2.5	Amino acid residues of AVCP and CVCP which frequent in hydrophobic interactions with the drugs	39
3.1	DSC parameters used in conducting the experiment to detect the stability of AVCP19	60
3.2	Number of samples and their concentrations used in the DSC experiment	60
3.3	Refinement statistics of resolved crystal structure of AVCP80	69

LIST OF ABBREVIATIONS

˚C	Degree Celsius
AP4	P1', P4' Di(adenosine-5') tetraphosphate ammonium salt
AURAV	Aura virus
AVCP	Aura virus capsid protein
CBB	Coomassie Brilliant Blue
CHIKV	Chikungunya virus
CPE	Cytopathic effect
CVCP	Chikungunya virus capsid protein
DNA	Deoxyribonucleic Acid
DNA	Deoxyribonucleic acid
DRV	Darunavir
DSC	Differential Scanning Calorimetry
EAC	Eptifibatide acetate
EC_{50}	Half maximal effective concentration
EDTA	Ethylenediaminetetraacetic acid
ER	Endoplasmic Reticulum
H	Hour
S	Second
H-bonds	Hydrogen bonds
IMAC	Immobilized Metal Ion Affinity Chromatography
Kb	Kilobase
kDa	Kilodalton
K_i	Inhibition constant
K_m	Michaelis-Menten constant
LB broth	Luria Bertani broth
µM	Micromolar

MD simulation	Molecular dynamics simulation
Min	Minutes
ml	Millilitre
M	Molar
mM	Millimolar
MOLREP	Macromolecular refinement
MNTD	Maximum non toxic dose
mRNA	Messenger RNA
MS	Mass Spectrometry
MW	molecular weight
N	stoichiometric value
Ni-NTA	Nickel – Nitrilotriaceticacid Agarose
ORF	Open reading frame
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PEG	Polyethylene Glycol
PEG	Polyethyleneglycol
PSU	Paromomycin Sulphate
RMSD	Root mean square deviation
RMSF	Root means square fluctuation
RNA	Ribonucleic Acid
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SINV	Sindbis Virus
Tm	Melting temperature
UTE	Untranslated Region
Vi	Initial velocity
WEEV	Western Equine Encephalitis Virus
µl	Microlitre

CHAPTER 1

Introduction and Review of literature

1.1 Introduction

The genus alphavirus belongs to the family Togaviridae which is subcategorized in the ‘group IV’ of the ‘classification of viruses’ by Baltimore [8]. The genus alphavirus comprises of 40 recognized members which are distinctly represented as enveloped positive single stranded RNA viruses. Majority of alphavirus are constituted as arbovirus as they are transmitted by arthropod vectors majorly by mosquitoes in vertebrates especially in humans, rodents, birds, horses, fish etc.[9]. Southern elephant seal virus, a newly isolated virus is supposedly transmitted by the louse *Lepidophthirus macrorhini* in Macquarie island elephant seal [10]. Salmon pancreatic disease virus infects salmon and trout and the vector responsible is speculated to be louse *Lepeophtheirus salmonus* [11]. Fort Morgan virus and Bijou Bridge strain virus are transmitted by swallow bugs *Oeciacus vicarious* in nestling Cliff swallows and symbiotic House sparrows [12].

Alphaviruses have acquired notoriety worldwide and their outbreak has created havoc in five out of six continents Africa, Asia, America, Australia and Europe. Old World viruses and New World viruses are further distinguished owing to the differences in their phylogenetic and phenotypic characteristics where the former is majorly responsible for arthralgia, fever and rash while the latter is notorious to cause encephalitis and febrile illness. On the contrary of this, SINV belongs to the Old World viruses but shares familiar genetic characteristics with the New world viruses whereas Una and Mayaro viruses belong to New World category but genetically and antigenically they are more closely related to the Old World [13]. The Old World alphaviruses require nsP2 protein for the down regulation in the host cellular transcription whereas capsid is responsible for ceasing cellular transcription in the host infected by New World alphaviruses [14], [15]. The mortality rate by the New world alphaviruses infection is comparatively higher and thus they are assigned in the category of potential B bio agents [16]. Due to the assorted degrees of antigenic cross reactivity exhibited by the means of alphavirus glycoproteins E1 and E2, the 26 members of alphavirus is categorized into six antigenic complexes [17]. There are 26 members of alphavirus which pose health threats to humans and livestock. These pathogens amass to possess an important area of concern in medical biology. CHIKV, Mayaro virus (MAV), SINV, RRV , WEEV, EEEV and O’ nyong’nyong virus (ONNV) are few names which are known to cause range of diseases, all of these exhibit common symptoms like febrile illness, acute joint pain, rashes etc in the affected patients. Western Equine encephalitis (WEEV) virus is a mosquito borne (*Culex and Aedes*) species alphavirus, causes fatal encephalitis in humans and livestock which usually leads to indefinite neurological sequelae in the affected young children and likelihood of death in the affected elderly because of the intensified neurological complication. Survivors escaping from less severe neurological complications suffer from permanent recurring headaches, fatigues and irritability. WEEV is geographically distributed towards the western parts of Canada, United States and in Southern region of Argentina. Till now no vaccine and effective treatment is available against WEEV or any alphavirus [18], [19]. SINV infection characterised by febrile illness, rash and arthritis is transmitted by *Aede*s and *Culex* species of mosquitoes. The virus had gain notoriety in causing epidemic in Fennoscandia in the initial 1990s and has been a major concern for its widespread in the regions of Europe, Asia, Africa and Australia [20], [21]. No vaccine or preventive medical treatment is available against the virus [22]. The WEEV complex (antigenic complex) comprises of viruses that cause arthritis (SINV) along with the encephalitis causing viruses Fort Morgan Virus (FMV), High lands J virus (HJV) and WEEV [23]. WEEV is a New world virus but it exhibits serological similarity with its contrary part Sindbis virus (SINV) from Old World clade. The emergence of WEEV attributes to the recombination between an Eastern Equine Encephalitis virus (EEEV) and SINV [22], [23]. Chikungunya virus (CHIKV) harbours on its vector mosquito *Aedes aegypti* and *Aedes albopictus* and is transmitted to the vertebrate hosts via sylvatic and urban cycle[24]. CHIKV infection was first reported in Tanzania in 1962 and is geographically distributed over Asia, Africa and lately (2000 onwards) in Europe and Americas.CHIKV is an Old World alphavirus and is classified into three genotypes Asian, East Central South African and West African on the basis of its geographical distribution [25]. CHIKV infection is accompanied by febrile illness and polyarthralgia symptoms [26]. CHIKV infection has comparatively adverse effect on neonates than the elderly [27]. The CHIKV infected neonates tend to suffer from neurological disorders which include seizure-disorders, encephalitis and severe encephalopathy [28], [29], [30], [31]. There has been no anti-CHIKV drugs available in market nor there is any vaccine against it [32]. Aura virus (AURAV) is an alphavirus which was firstly reported and isolated in early 1960s from *Aedes serratus* and *Culex* species in Brazil and few years later from *Aedes serratus* in Argentina [33]. Neither there has been any case of AURAV infection reported in humans, nor any vertebrate host for AURAV has been reported [34], [35]. AURAV lies in WEEV complex [34]. AURAV shares serological similarities with WEEV and SINV [36]. There are definite nucleotide sequence similarity between the RNA of AURAV and

SINV. The promoter sequence of AURAV bears identical similarity with its SINV analogues. The recurrent sequence in the 3' non translated region (NTRs) of AURAV is identical with their SINV counterparts. In spite of the nucleotide sequence similarity these two viruses share, they tend to diverge by exhibiting 73% amino acid sequence resemblance in the non-structural proteins and 62% equivalent amino acid sequences in the structural proteins. The extensive comparative study of the nucleotide sequence data of AURAV, SINV and WEEV has led to the inference that AURAV is the prototype of SINV-like viruses in the NEW World.The AURAV in contrast with the other alphaviruses is apt in encapsidation of subgenomic RNA along with the genomic RNA for the production of virus particles. The Aura virus particles exhibit T=4 structures which resemble with other alphaviruses [36].

The structural organisation details of the alphaviruses have been evaluated by crystallography and cryo-electron microscopy studies revealing the detailed organization of virion. The 5' capped alphavirus genomic RNA with polyA tail at 3' end is of approximately 11.7 kb size. The genomic RNA is encompassed by 240 numbers of capsid protein (encapsidation) [37], [38]. In alphaviruses, a mature virion is a well characterized enveloped particle (acquired from the cell membrane of host) with protruding 80 trimeric spikes (each trimer is composed of heterodimers of viral glycoprotein E2 and E1), exhibiting icosahedral symmetry(T=4) [39]. The alphavirus genome has~7 kb long ORF at the 5' end encoding the nonstructural polyprotein and a~4 kb long ORF at the 3' end encoding for the structural polyprotein [1]. The alphavirus capsid protein (CP) is a versatile protein. It plays a prominent role in the viral structural polyprotein processing, genomic RNA encapsidation, budding and in the downregulation of host transcription [40], [41], [42], [7]. The CP has two functional domains; the N-terminal domain & the C-terminal domain [5], [37], [43].

The N-terminal domain is highly disordered and comprises of basic charged amino acid residues mainly Lysine and Arginine, which possibly aids in binding with the genomic RNA and its encapsidation [37], [44]. The C-terminal domain exhibits cis-autoproteolytic activity which allows the cleavage of CP from the structural polyprotein. The cleavage occurs between the conserved scissile bond W-S, which is essential for the structural polyprotein processing [43]. After cleavage, the C-terminal W residue of CP stays boundin the active site of the protease domain and halts the proteolytic activity. A conserved hydrophobic pocket is present in CP which facilitates the binding of the cytoplasic tail of E2 glycoprotein,which is an essential interactionthat initiates the process of virus budding [37], [41], [45],[46], [47], [48]. The multidisciplinary role of alphavirus CP in the virus life cycle makes it a vulnerable antiviral drug target.

In the current study, potential inhibitory molecules against CVCP has been identified by virtual screening from the library of FDA approved drugs and 4 compounds from the screened compound library exhibiting best hits in terms of binding energy against the target active site of CVCP were selected for testing of anti-protease activity of CVCP using the substrate specific fluorogenic peptide based FRET based assay. After validatingthe anti-protease activity of the four selected compounds and performing enzyme inhibition kinetics, the compounds were subjected to *in vitro* cell based antiviral assays to study the antiviral effect of CVCP inhibitors on CHIKV replication in infected cells.

The literature survey shows that the crystal structures of SINV CP inactive, VEEV CP inactive, AVCP active and CVCP native have been determined and have revealed a chymotrypsin-like protease fold in the C-terminal CP domain of alphaviruses [4], [5], [45], [49]. In this study, virtual screening of potential inhibitory compounds was done by docking of molecules in the active site of the enzymatically active form of AVCP using the PDB ID 4AGK coordinates. The compounds with potential inhibitory activity were selected based on the binding energy of compounds into the active site of AVCP.

Till date there is no crystal structure of the active form of CVCP and the aim of the study was to identify chikungunya virus capsid protease inhibitors. Therefore, for the structure-based identification of CVCP inhibitors, the docking of the above identified and selected compounds in the active site of CVCP was done. To gain access to the active site of CVCP, the conserved C-terminal W261 residue from the coordinate file of the crystal structure of CVCP native (inactive proteinPDB ID: 5H23) was *in silico*removed to make the active site accessible for docking of small molecules. These four identified molecules also fitted well in the active site of CVCP. Thus, these compounds were tested for anti-protease activity. The substrate specific FRET based assay for CVCP activity has been previously developed for high-throughput screening of inhibitors in our laboratory and reported [5]. The assay facilitated this study in conducting the inhibition assay against CVCP for testing the selected four compounds. These compounds were further tested in the *in vitro*cell-based antiviral assays, prior to which MTT based cytotoxicity assay was conducted to determine the MNTD (maximum non toxic dose of these compounds) of each compound. Based on the MTT assay the concentration of the compounds used in the cell culture for antiviral assays were within the MNTD limit. The CVCP protease inhibitors targeted CHIKV at the post entry stage in the cell-based antiviral assays which suggests that these inhibitors are potential antivirals against CHIKV.

In this study an attempt was also made to determine the three-dimensional structure of the N-terminal domain of alphavirus capsid. The N-terminal domain of alphavirus capsid is disordered. No crystal structures of N-terminal domain of alphavirus CP is reported till now [4], [5], [45], [49], [50]. This is due to the stretch of positive residues in the N- terminal domain and the stability in the disordered region is attained by the electrostatic interactions with the genomic RNA [51].The subdomain 1(SD-1; residues range 1-37) of N-terminal domain of VEEV CP is involved in nucleocapsid assembly and in the virus assembly. The mutation (deletion/substitution by similar sequence of other alphaviruses) in the sequence of SD-1 has a detrimental effect on the release of infectious virus [52]. The SD-2 (residues range 38-55) of N-terminal domain of SINV CP assists the CP dimerization which stabilises the core assembly *in vitro* [53], [54]. The SD-2 (residues range: 38-51) of N-terminal domain of VEEV CP plays the role of nuclear export signal (NES) which mediates the inhibition of nuclear pore trafficking [55].The SD-3 (residues range: 52-110) of N-terminal domain of VEEV CP has basic positive charged residues in abundance. SD-3 inhibits NC assembly until charge neutralization is attained by RNA interactions. SD3 is reported to mediate most of the CP-RNA binding [6]. The SD4 (residues range: 111-126) of the N-terminal domain of VEEV CP is actively involved in the viral genome packaging [6]. The SD4 of N-terminal domain is highly conserved among all the members of alphavirus [51]. The mutations in the SD4 cause nonspecific NCP formation in SFV and SINV [56], [54]. In SINV CP, 81-113 amino acid residues specifically binds to the encapsidation signal of the genomic RNA which is shown to provide stability in the oligomerisation of alphavirus CP and in formation of NC cores [57]. It has been reported that the CP of members of New World alphavirus causes host transcriptional shutoff which exhibit cytopathic effects, an essential characteristic of an alphavirus infection [58]. In the N-terminal domain of VEEV CP, 33-68 residues are reported to block multiple nuclear import pathways [59]. To decipher the three-dimensional structure of the N-terminal domain of alphavirus, cloning of AVCP19 (19-265 residues) and CVCP20 (20-259 amino acid residues) were done. The purification of CVCP 20 was unsuccessful but AVCP19 was purified by immobilized metal affinity chromatography (IMAC) iusing the phosphate buffer. Co-crystallization attempts of AVCP19 with 12mer DNA were made. No crystal of AVCP19 DNA complex was obtained. Another construct of AVCP (AVCP 80; 80-267 residues) was used for co-crystallisation with 12 mer DNA.The crystal were obtained that diffracted to 1.2 Å but no electron density of the N-terminal domain was observed,also density of DNA was absent. The CP-DNA interactions were studied by size exclusion chromatography and differential scanning calorimetry (DSC). The size exclusion chromatography showed the 12-mer DNA interactions with the N-terminal domain of AVCP 19 and DSC showed the shift in the Tm of the AVCP 19 bound DNA. The terminal Trp residue of both AVCP19 and CVCP 20 were deliberately removed while making the constructs to study the *trans*proteolytic activity of CP having N-terminal domain and also to detect the change in the proteolytic activity in DNA bound N-terminal domain of CP.

1.2 Alphavirus life cycle

The genus alphavirus comprises of majority of medically important viruses which include WEEV, VEEV, EEEV, RRV, CHIKV, SINV, BFV etc. All these medically important alphaviruses are arbovirus as these are transmitted by mosquitoes. In the infected vector (mosquito), the alphavirus replicates and spreads via the salivary discharge [60], [61], [62]. The transmission cycle of alphavirus is in compliance with the order; mosquito – vertebrate. The alphavirus transmission cycle is elucidated by the graphical representation in Fig 1.2.1.

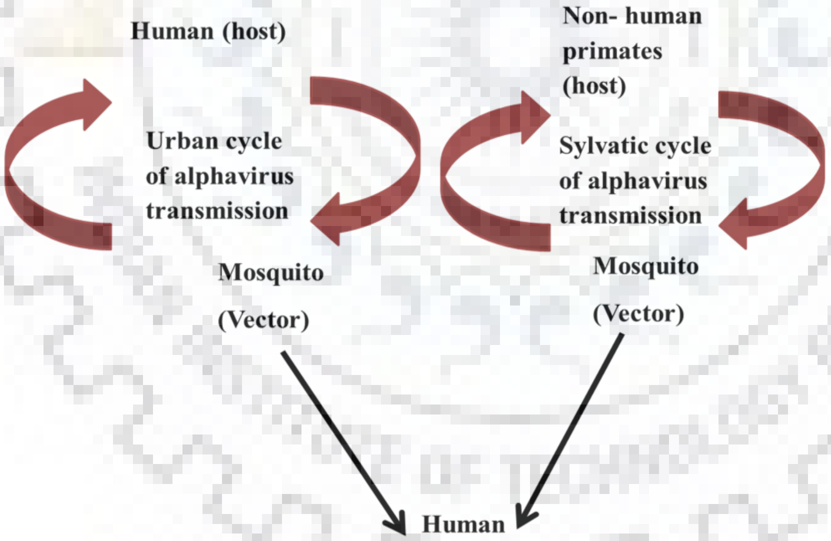

Human (host)
Urban cycle of alphavirus transmission
Mosquito (Vector)
Non- human primates (host)
Sylvatic cycle of alphavirus transmission
Mosquito (Vector)
Human

Figure 1.2.1 : Graphical representation of alphavirus transmission cycle [18]**.**

The alphavirus life cycle begins by attachment of virus to the cell membrane of the host with the help of certain receptors. Alphavirus uses wide range of host cell receptors; heparin, integrin, heparin sulphate, C-type lectins such as DC-SIGN and L-SIGN, class I major histocompatibility antigen, laminin for its entry inside the host cell [63], [64], [65]. An illustrative representation of life cycle of alphavirus is shown in Fig. 1.2.2. The E2 glycoprotein of alphavirus is required for the attachment of virus with the plasma membrane of the host cell. Inside the endosome of host cell, acidification occurs in the vicinity of vacuolar ATPase which leads to the virion fusion with the endosomal membrane. This is superseded by the virus entry inside the host cell via receptor facilitated endocytosis [4], [66]. The acidic environment inside the endosome triggers the conformational changes in E1-E2 glycoproteins which causes detachment of heterodimer E1-E2 and lodging of E1 fusion peptide inside the endosomal membrane [67], [68]. Prior to the acidic environment in the endosome, the E1 peptide is masked behind the heterodimer of E1-E2 inaccessible to endosomal membrane. These series of events facilitate release of NC core into the cytoplasm of the host cell [66]. The translation of non-structural polyprotein (P1234) is from the two-thirds of the 5' capped genome- length mRNA. The nsp123 and nsp4 are generated by the cleavage of P1234 between nsp3 and nsp4 by virus specific protease nsP2. The P123 is further processed to produce non-structural proteins nsp1, nsp2 and nsp3. Thesegenerated nsp1, nsp2, nsP3 and nsp4 along with few unspecified factors constitute replication complex. The subgenomic 26S RNA upon translation generates structural polyprotein. The structural polyprotein is processed to produce structural proteins CP, E3, E2, 6K and E1. The CP of alphaviruses is a serine protease and it comprises of the chymotrypsin-like fold that exhibits autoproteolytic activity [4], [5], [45], [49]. Upon translation of the structural polyprotein, the CP cleaves itself from the polyprotein at the scissile bond W-S. After clevage of CP, the N- terminus of the structural polyprotein possesses signal sequence which leads to its translocation in the endoplasmic reticulum and Golgi apparatus for its further processing and post translational- modifications. When the PE2-E1complex translocates towards the trans Golgi apparatus, the host protease Furin cleaves the PE2 to generate E2 and E3 [37]. The encapsidation of genomic RNA is attributed by the cleaved CP which leads to the development of nucleocapsid assembly and nucleocapsid core [69].The cytoplasmic domain of E2 (cdE2) binds to the hydrophobic pocket of CP (near the C-terminus of CP) which triggers the mature virion to bud out from the surface of host cell [70].

1.3 Genome organization of alphavirus

The genome of alphavirus is ~11.7 kb long. It is comprised of 49S positive-sense single-stranded RNA. The genomic RNA is capped with 7-methyguanosine at the 5' end and is polyadenylated at the 3'end. This 49S RNA has two open reading frames (ORF's) which generates non-structural polyprotein (nsP1-nsP2-nsP3-nsP4) and structural polyprotein (C-E3-

E2-6K-E1). The ORF present near the 5' end is two-third of the genome and encodes for the non-structural polyprotein.

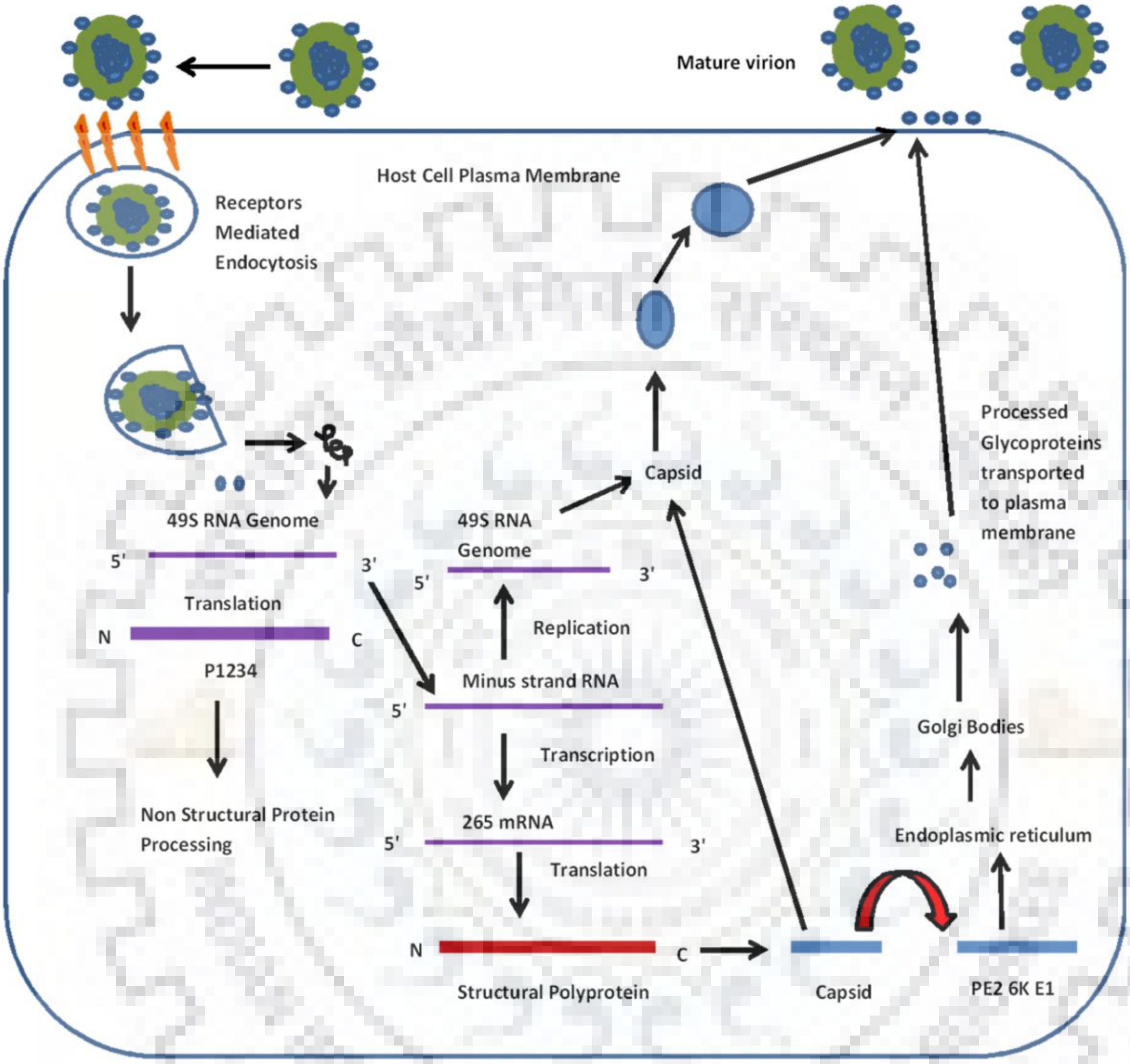

Mature virion
Host Cell Plasma Membrane
Receptors
Mediated
Endocytosis
Capsid
Processed
Glycoproteins
transported
to plasma
membrane
49S RNA Genome
5'
3'
Translation
N
C
P1234
Non Structural Protein
Processing
49S RNA
Genome
5'
3'
Replication
Minus strand RNA
5'
Transcription
265 mRNA
5'
3'
Translation
N
C
Structural Polyprotein
Capsid
PE2 6K E1
Golgi Bodies
Endoplasmic reticulum

Figure 1.2.2: Illustrative representation of alphavirus life cycle. The entry of alphavirus in the host cell is facilitated via the host receptors. The acidic medium in the endosome triggers conformational changes in the glycoproteins E1 and E2 which releases the viral genomic RNA in the cytoplasm. The genomic RNA upon translation generates non-structural polyprotein which is further processed by nsP2 protease to produce four individual non-structural proteins. The 26S subgenomic RNA upon translation produces structural polyprotein CP-E3-E2-6K-E1. The CP present at the N- terminus of structural polyprotein is cleaved off. The CP interacts with the viral genomic RNA, encapsidates it and forms nucleocapsid cores. The CPs of assembled nucleocapsid cores interact with the cytoplasmic tails of the viral glycoproteins and this interaction facilitates the budding of newly formed virus particles from the infected host cell [68].

The ORF encoding structural polyprotein is present near the 3'end and constitutes one-third of the genome [37], [71], [72], [73]. The 26S subgenomic RNA is ~4.2 kb long and translates to produce structural polyprotein [44]. There are 4 conserved domains in the viral genomic RNA collectively known as conserved sequence elements (CSE's). These CSE's has crucial role in the replication of alphaviruses by interacting with the viral or host proteins. The viral genomic RNA has CSE1 confined to the 5' end. The CSE1 is 44 nucleotides long and serves as promoter for the synthesis of positive sense strand RNA [74]. The CSE2 is comprised of 51 nucleotides and serves as promoter in the synthesis of negative-sense strand RNA [75]. The CSE3 is comprised of 24 nucleotides and is involved in the subgenomic RNA transcription [76]. Near the 3' end of viral genomic RNA is 19 nucleotide long CSE4 which serves as the core promoter in the synthesis of minus-strand RNA [38], [77], [78], [79]. The schematic organisation of genome is shown in the Fig.1.3.1.

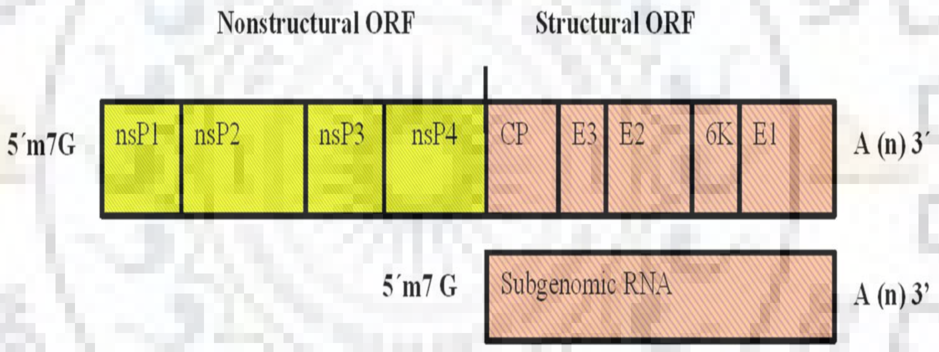


Figure 1.3.1: Genome organization of alphavirus [37]. The alphavirus genome is ~11.7kb long 49S positive-sense single-stranded RNA with the 5'end capped by 7-methylguanosine and 3' end polyadenylated. The ORF present near the 5' end encodes non-structural polyprotein which is processed to produce non- structural proteins. The 3'end is the ORF that encodes structural polyprotein which undergoes further processing to produce structural proteins [37], [71].

1.4 Non-structural proteins

The two-third of the genome constituting ORF present at the 5' end encodes non-structural polyprotein which is processed to produce 4 essential non-structural proteins nsP1, nsP2, nsP3 and nsP4 that necessitates the replication of viral RNA.Post alphavirus infection, the released viral genomic RNA translates to generate non-structural polyprotein which gets processed by the proteinase activity of nsp2 protease [80]. The translated non-structural polyprotein can be either complete P1234 or in the vicinity of opal codon (UGA), translation of P123 occurs first followed by the translation of P1234. In the latter case, P123 is processed to produce nsp1, nsp2 and nsp3 while nsp4 is generated from the second translated nsp1234 precursor. The nsp2 exhibits *cis*-protease activity while processing P1234 into P123 and nsp4. By altering the cleavage sites between nsP1 and nsP2 by site specific mutagenesis, the processing at nsP2/nsP3 site was prevented and led to the accumulation of P123 *in vitro*. This suggests the cleavage between nsP1 and nsP2 is essential for the initiation of polyprotein processing [81]. The *trans*-protease activity of nsp2 is exhibited in the release of nsp1 from P123 [82]. Afterwards, the *trans*- protease activity of nsP2 at the proteolytic cleavage site between nsP2 and nsP3 releases nsP2 and nsP3 [83], [82]. These non-structural proteins together with some unspecified factors comprises viral replication complex. The functions of non-structural proteins of alphavirus are briefly described below:

NSP1: The nsP1 is a ~ 60 kDa protein which plays a major role in the alphavirus replication. The alphavirus replication is mediated by the N-terminal domain of nsP1.The nsp1 possesses essential methyltransferases (MTase) and guanylyltransferase (GTase) enzymatic activities required for the capping of viral genome and subgenomic RNA [84], [85], [86], [87]. The nsp1 is reported to be inassociation with the membrane of host cell and endocytic organelles,thus affixing the replication complex to host cell membrane [87], [88]. The nsp1 is reported to function in the synthesis of minus-stand RNA [89].

NSP2: The nsP2 is a ~ 90 kDa protein evincing multiple functions. The N-terminal domain of nsP2 has a helicase activity accompanied by nucleoside triphosphatase activity (NTPase). The C-terminal domain of nsP2 possesses protease activity [43]. The crystal structure of nsp2 protease resembles the cysteine protease fold [90], [91]. Inhibitors of nsP2 viral protease of CHIKV have been identified by *in silico* molecular docking approach whereby peptide analogs of the substrate were observed to bind the active site [92]. Several other novel inhibitors against nsP2 protease of CHIKV have been screened by the virtue of *in silico* predictions, validated by enzyme inhibition assays and cell-based virus inhibition approaches [93]**.**

NSP3: The nsP3 is ~ 60 kDa protein and comprises of three important domains: the macrodomain present in N-terminal region, zinc binding domain in central and a hypervariable domain (HVD) present in the C-terminus. Macromain is involved in binding with ADP- ribose and its derivatives [94]. These macrodomains hydrolyse MARylated acidic residues. Mutating these residues impedes the viral replication [95]. The HVD exhibits ADP-ribose 1' phosphate phosphatase and RNA binding activity and plays an important role in the synthesis of subgenomic RNA and minus strand RNA [94], [96], [97]. The zinc binding domain is exclusively present in alphaviruses and is absent in the closely related member Rubella virus of the same family Togaviridae. The function of this domain is still evasive however a recent study on this domain by conducting mutation in the conserved residues of this domain speculates its role in the replication of viral genome and expression of structural protein of alphavirus [98]. Mutation in the protein has shown to cause defects in the initiation of minus-strand synthesis along with the defects in the synthesis of subgenomic RNA [99].

NSP4: The nsP4 is ~ 70 kDa core viral polymerase protein having GDD motifs which exhibits terminal adenylyltransferase activity (TATase) and RNA-dependent RNA polymerase (RdRp) activity involved in RNA replication [100], [101], [102]. The schematic diagram elucidating processing of non-structural polyprotein is shown in Fig.1.4.1

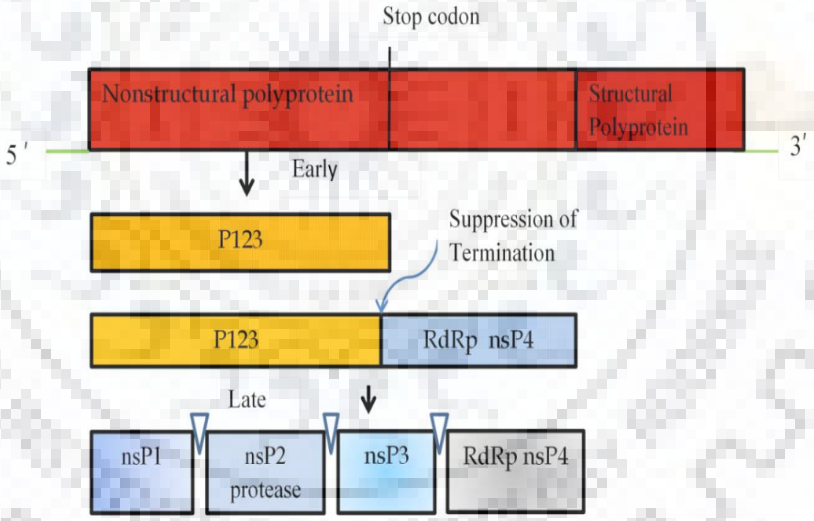

Stop codon
Nonstructural polyprotein
Structural Polyprotein
5′
3′
Early
P123
Suppression of Termination
P123
RdRp nsP4
Late
nsP1
nsP2 protease
nsP3
RdRp nsP4

Figure 1.4.1: Non-structural polyprotein processing. The nsP1234 polyprotein is processed by virus specific nsP2 cysteine protease to produce nsP1 capping enzyme, nsP2 protease, nsP3 and RNA dependent RNA polymerase (RdRp) nsP4[38].

1.5 Structural Proteins

The structural polyprotein is translated from the 26S subgenomic RNA. The ORF encoding the structural polyprotein is present at the 3'end of the viral genomic RNA and constitutes one-third of the genome. Once the structural polyprotein is translated, the CP cleaves itself from the polyprotein owing to its autoproteolytic activity. The cleaved polyprotein has now signal peptide in the N-terminus which directs its translocation to the endoplasmic reticulum and Golgi apparatus wherein the polyprotein undergoes post translation modification and further processing to generate other structural proteins by host protease like furin and signal peptidase [103], [104]. The schematic representation of structural polyprotein processing is shown in Fig.1.5.1

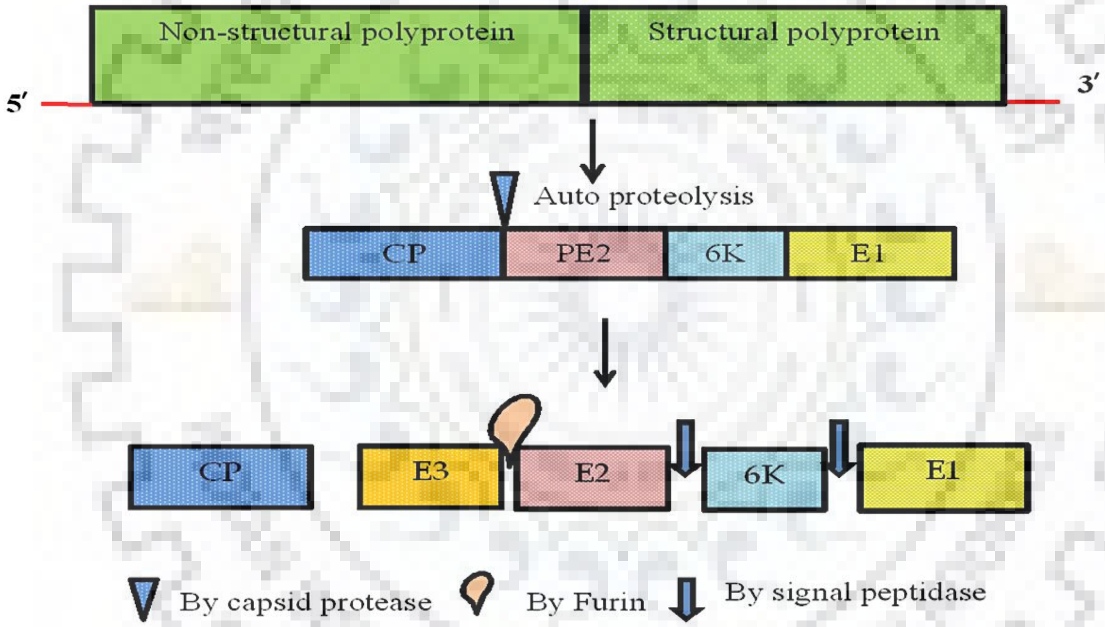


Figure 1.5.1: Processing of alphavirus structural polyprotein [105]**.** Structural polyprotein is translated from the transcribed 26S subgenomic RNA .The structural polyprotein is processed by the viral capsid protease and host proteases to generate functional structural proteins in the order C-E3-E2-6K-E1 [40], [70].

The CP of alphavirus is present at the N-terminus of structural polyprotein. The CP cleaves itself from the translated polyprotein by the virtue of its autoproteolytic activity. The CP of alphavirus is broadly categorised into two functional domains which include N-terminal and C-terminal domain. The N-terminal domain of CP has RNA binding locus as the domain is rich in basic amino acid residues.The C-terminal domain of CP has a serine protease which exhibits cis-proteolytic activity [37], [5]. The CP protease domain also possesses a small hydrophobic pocket through which it interacts with the cytoplasmic tail of the viral glycoprotein E2. This interaction has been proposed to be the key process in the budding of virus particles [70], [49], [106]. The functional domains of alphavirus CP are represented in Fig.1.5.2.

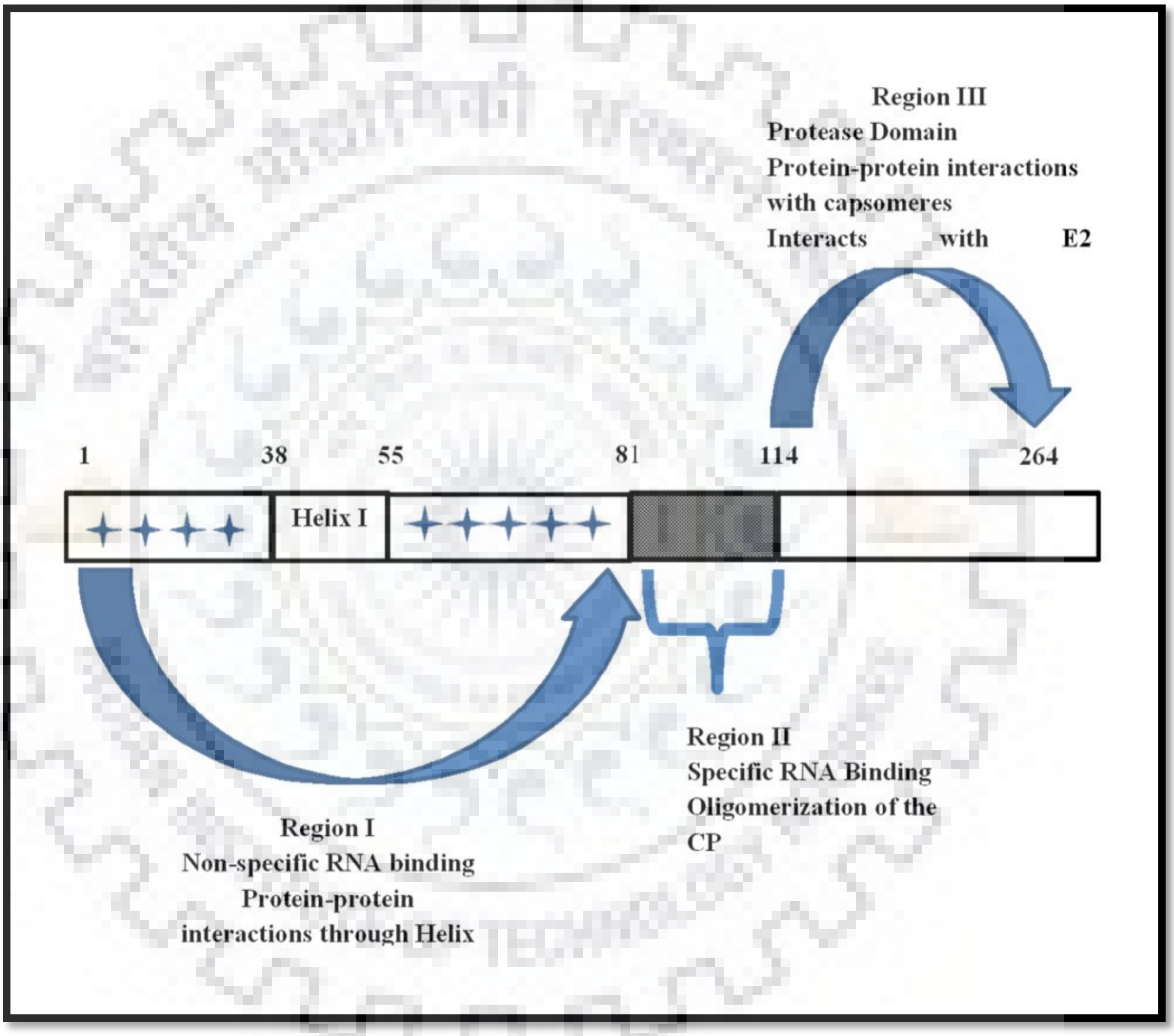

Region III
Protease Domain
Protein-protein interactions with capsomeres
Interacts with E2
1
38
55
81
114
264
Helix I
Region II
Specific RNA Binding
Oligomerization of the CP
Region I
Non-specific RNA binding
Protein-protein interactions through Helix

Figure 1.5.2: Representation of the functional domains of SINV capsid protein[97]

Prior to the translocation of the cleaved structural polyprotein in the ER and Golgi apparatus, the conserved cysteine residues of precursor E2 (pE2) undergoes palmitoylation. The palmotylated cytoplasmic domain of E2 (cdE2) allows it to anchor to the inner side of the membrane which mediates the cdE2-capsid interactions [46], [41], [108], [105]. Before transiting to the plasma membrane through the trans-Golgi,the heterodimer of pE2-P1 gets cleaved by furin to generate E2 and E3 [109], [105]. The passage of E3 to the plasma membrane is essential for the budding as it aids in stabilization of heterodimer [110]. E2 is an essential viral protein which facilitates virus entry by interacting with the host cell receptors thus the cleavage of pE2 is a necessity [111]. The self-assembly of heterodimers of E1-E2 makes trimeric spikes [37], [112], [113]. The envelope of a mature virion is interspersed with eighty trimeric spikes in an orderly symmetry (T=4 icosahedral symmetry) [114]. The studies carried out by cryo-electron microscopy has shown the presence of N-linked glycosylation sites on E2 through which it binds heparin sulphate on the surface of mammalian cell [115], [116].

The 6K is a membrane associated protein comprising of approximately 60 amino acid residues [117]. In SFV, 6K facilitates the lodging of E1 into the membrane of ER by acting as stop transfer signal [118]. 6K are speculated to function as viroporin as they regulate membrane permeability in the host cell and help in budding of the virus [103]. In each virion approximately 30 copies of 6K are incorporated. The palmitoylated conserved cysteine residues of 6K is essential for the generation of infectious particle. The mutated cysteine residues in 6K exhibit abnormality in the virion particles with diminished infectivity [119], [120].

E1 is required for the genesis of icosahedral structure of the alphavirus. E1 and pE2 assemble to make heterodimer which necessitates virus assembly. The E1-pE2 undergoes furin cleavage to generate E1-E2 heterodimer in the course of their transportation to the cell surface or in virus assembly. In a mature virion, a triplet of E1-E2 heterodimer form spikes (trimer) in T=4 icosahedral symmetry. These spikes protrude out from the virus surface thus giving it a typical icosahedral structure [121]. Transframe protein (TF) is produced by the frameshift which occurs during the 6K protein translation. TF is speculated to play an important role in virus assembly [122].

1.6 Alphavirus capsid structure

The N-terminal domain of CP of alphavirus is highly disordered due to the presence of abundant basic charged amino acid residues [1]. The C-terminal of CP has protease domain which exhibits *cis*-proteolytic activity. The CP gets cleaved from the newly translated structural polyprotein owing to its auto proteolytic activity. After getting detached from structural polyprotein, the terminal conserved tryptophan residue of CP binds to the active site of CP and halts further proteolytic activity renders the serine protease inactive [4], [123]. Crystal structures of the inactive CP (after cleavage form of CP) of AURAV, CHIKV, SINV, SFV and VEEV are reported which exhibit fold similar to the structure of chymotrypsin serine protease

[49], [106], [43], [4]. The only crystal structure of active form of CP has been determined for AURAV [5]. The catalytic triad underlies between the two subdomains of CP which is comprised of three conserved amino acid residues Ser, His and Asp distantly located in the sequence. These residues come nearby in the tertiary structure of CP to constitute active form of the CP [5]. A small hydrophobic pocket is present at the reverse side of the active site of CP in which cdE2 binds to promote virus budding and it is also postulated to be involved capsid-capsid interactions [106], [43], [124].

None of the crystal structures of CP of any alphavirus has revealed the density and the 3D structureof the N-terminal region. Crystal structures of the protease domain of CP exhibit two ß-barrel sub-domains represented by a Greek key motif which is similar to the other serine proteases (Fig. 1.6.1). In between the two ß-barrel sub-domains lies the catalytic triad. The residues of the catalytic triad like other serine proteases constitute histidine, aspartate and serine.The conserved serine residue of the catalytic triad upon mutation banishes the protease activity of the CP [125].

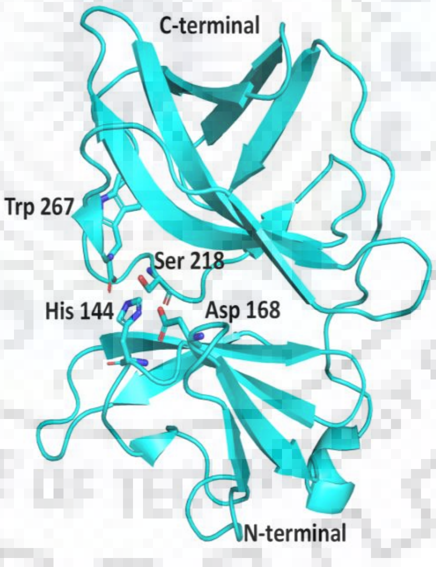


Figure 1.6.1. Cartoon view representation of the crystal structure of AVCP (PDB ID: 4AGK). The protease domain of CP constitutes two ß-barrel domains and the catalytic triad is interspersed between the N-terminal and C- terminal subdomains.The catalytic triad residues are shown as sticks. The terminal Trp 276 residue is shown bound to the active (After the *cis* cleavage, the conserved C-terminal Trp residue binds to the active site and halts the trans-protease activity of CP [45].

1.7 The *trans*-activity of alphavirus CP protease

After the *cis*-proteolytic activity of alphavirus CP, the enzyme is rendered inactive because of the binding of the terminal Trp residue of CP in the active site [4]. Potential of *trans* activity was reported for CP protease in SFV by removing the terminal Trp residue of CP, thus making it catalytically active [126]. The *trans* activity was determined in the CP of AURAV and CHIKV by truncating the conserved terminal Trp residue and using fluorogenic peptide which bears the exact sequence present in the scissile bond. By the virtue of FRET-based assay, catalytic efficiency of protease was determined which exhibited high efficiency1.11× 10^3 $M^{-1}s^{-1}$ to cleave the substrate [5].

Table 1.1: Compilation of crystal structures of different members of alphavirus family

Alphavirus	PDB ID
Sindbis Virus (SINV)	1SVP, 1 WYK, 2SNW, 1KXB,1KXC,1KXD, 1KXE
Venezuelan Equine Encephalitis Virus (VEEV)	1EP5,1EP6
Aura Virus (AURAV)	4AGK, 4AGJ, 4UON, 5G4B
Chikungunya Virus	5H23
Semliki Forest Virus (SFV)	1VCP, 1VCQ

1.8 Capsid protein as essential antiviral drug target

The CP of alphaviruses plays an important role in the viral life cycle. The N-terminal domain of CP facilitates encapsidation of genomic RNA, keeping the genetic content of virus intact. The C-terminal serine protease domain of CP attributes in the structural polyprotein processing and in the budding of newly formed virus particles. All these key features of CP make it a vulnerable antiviral drug target.

1.9 Repurposing drugs against CHIKV

CHIKV re-emergence on a global scale especially in developing nations is quite alarming because of lack of licensed anti CHIKV drug/vaccines. Although CHIKV infection is accompanied by low rate of mortality but its rate of morbidity is high. To curb the CHIKV menace there is need of developing cost effective anti-CHIKV drugs in short span of time. Drug repurposing/repositioning from the library of FDA approved drugs for combating CHIKV could be an ideal rescue. This will facilitate the repurposed drugs for CHIKV to surpass ADME test (Absorption, distribution, metabolism and excretion tests) in clinical trials thus saving cost and time. Drug repurposing exploits many approaches e.g. molecular docking where computational data (structure based) is exploited and binding position between ligand (drugs) and remedial target is anticipated. The speculated drugs can be further validated by *in vitro* based high throughput biochemical enzyme inhibition assays along with the cell-based antiviral assays [127], [128], [129], [130], [131].

CHAPTER 2

Structure-based identification and evaluation of chikungunya virus capsid protease inhibitors as antivirals

2.1. Chapter summary

The *cis* autoproteolytic activity of CP of alphavirus occurs between the conserved C-terminal Trp residue (P_1 residue) and N- terminal Ser residue ($P_{1'}$) of E3 glycoprotein. Post this *cis* proteolytic activity, the C-terminal Trp residue binds to the active site of CP protease and halts the further proteolytic activity. The binding of Trp residue in the active site, prevents the binding of next or *trans* substrate to the active site of CP protease thereby rendering the protease inactive to undergo next proteolysis [66]. However it was found that by removing the last Trp residue from the CP makes the protease active again and the *trans* activity of CP protease is generated. Kinetic studies of the CP protease of AVCP and CVCP active using fluorogenic peptide specific FRET based assay has depicted the presence of *trans* activity and has also shown efficient catalytic efficiency [5].

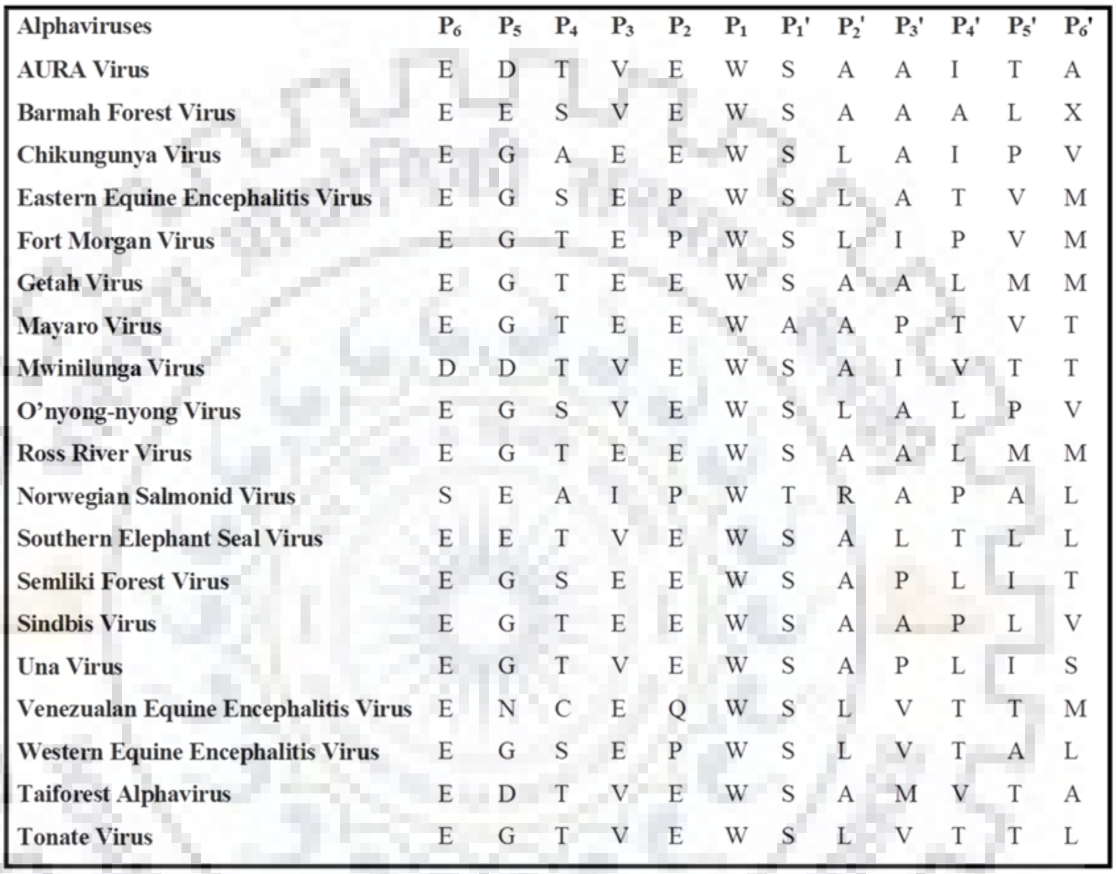
The chapter 2 attempts to crystallise CHIKV CP protease in its active state by removing the C-terminal tryptophan residue. The chapter also aims to investigate the potential inhibitors of CHIKV capsid protease (CVCP protease) by virtually screening the compounds from the library of FDA approved drugs by molecular docking.The current study has exploited the peptide specific FRET based inhibition assay to validate the inhibitory potential of the 4 identified compounds AP4, EAC, PSU and DRV. The K_i calculated for AP4, EAC, DRV and PSU were 133 ±27.45µM, 135± 76.25µM, 147.4± 10.93µM and 353.0±17µM respectively. Interestingly, these molecules not only inhibited the protease activity of purified CHIKV capsid in FRET based protease assay, but also suppressed CHIKV replication *in vitro* cell based assays. Plaque reduction assay revealed that the identified compounds significantly suppressed virus replication in the post treatment. The half maximal effective concentration (EC_{50}) of the compounds were in the following ascending order EAC > AP4> PSU and were 4.01 µM, 10.66 µM and 22.91 µM respectively. Chikungunya viral RNA levels also diminished in inhibitor-treated cells which were analysed by q-RT-PCR. Additionally, immunofluorescence assay using E2 glycoprotein antibody confirmed the anti-CHIKV potential of these compounds. In conclusion, structure-assisted drug-repositioning strategy identified novel capsid protease inhibitors that are potential anti-chikungunya agents.

2.2 Introduction

The origin of CHIKV dates back to the year 1952 AD when the first case related to it was reported in the Makonde plateau in Africa and thus, the term is derived from the local language 'Makonde' which has literal meaning "stooped posture", diagnosed as one of the most common arthritic symptom caused by the virus . Since then time and again the virus has been creating menace in tropical countries and has also spread towards north and temperate areas [132]. CHIKV is a mosquito borne virus transmitted majorly by *Aedes agyptii* and *Aedes albopictus* species [132], [133]. Patients infected by CHIKV show severe symptoms like fever, severe joint pain, muscle pain, headache, skin rash and nausea. Currently no antiviral treatment has been attained against it [134], [135]. CHIKV has been mentioned among the list of 'neglected tropical diseases' as it hovers around tropical and subtropical conditions where the majority of the population is poverty stricken which leads to compromise in the basic sanitation amenities and one of the basic measure to control the cause gets limited [136], [137]. CHIKV belongs to the genus alphavirus in the family Togaviridae.

The structural polyprotein is translated from the 26S subgenomic viral RNA and its intermediates are aligned in the order of CP-E3-E2-6K-E1 [37], [14]. The capsid (CP) upon translation gets chopped off from the amino terminus of the nascent structural polyprotein owing to the protease activity of the capsid confined to its C-terminus. The detached E3-E2-6K-E1 is then translocated to the endoplasmic reticulum (ER) where the host signalase cleaves the polyprotein at N- and C-terminal end of the 6K- peptide into E3-E2, 6K and E1[43], [138].

Alphavirus CP is a versatile protein which actively participates in the virus budding and assembly of virion [67], [108], [41]. Recently, the crystal structure of CVCP with PDB ID 5H23 has been determined at 2.2 Å by Sharma et al., 2018 [45]. CVCP crystal structure along with previously determined crystal structures of alphavirus CP revealed that the conserved C-terminal Trp261 remains bound at P1 position near the active site pocket, which makes the capsid protease inactive and brings the trans-protease activity to halt [4], [5], [49], [106], [124]. Deletion of the C-terminal Trp261 residue makes the active site accessible for the incoming substrate/ligands and leads to re-activation of the serine protease activity of CVCP [5]. Additionally, crystal structure of the C-terminal truncated Aura virus capsid protease (AVCP active) reveals the accessibility of the active site to substrate peptides [5]. The budding of virions from the plasma membrane involves the interaction between the NC and glycoproteins and it has been found that the C-terminal domain of CP has hydrophobic pocket where cytoplasmic domain cdE2 binds [67], [139]. The C-terminal domain of the CP possesses chymotrypsin-like serine proteinase fold and serine protease enzymatic activity [4], [5], [49]. After cis-cleavage the C-terminal residue Trp 261 remains bound to the S1 substrate specificity pocket, which brings the protease activity of CP to halt [5], [43], [140]. The *in vitro trans*-protease activity of the CHIKV CP (CVCP) has been determined using peptide/substrate specific FRET based assay [5].



Alphaviruses	P_6	P_5	P_4	P_3	P_2	P_1	P_1'	P_2'	P_3'	P_4'	P_5'	P_6'
AURA Virus	E	D	T	V	E	W	S	A	A	I	T	A
Barmah Forest Virus	E	E	S	V	E	W	S	A	A	A	L	X
Chikungunya Virus	E	G	A	E	E	W	S	L	A	I	P	V
Eastern Equine Encephalitis Virus	E	G	S	E	P	W	S	L	A	T	V	M
Fort Morgan Virus	E	G	T	E	P	W	S	L	I	P	V	M
Getah Virus	E	G	T	E	E	W	S	A	A	L	M	M
Mayaro Virus	E	G	T	E	E	W	A	A	P	T	V	T
Mwinilunga Virus	D	D	T	V	E	W	S	A	I	V	T	T
O'nyong-nyong Virus	E	G	S	V	E	W	S	L	A	L	P	V
Ross River Virus	E	G	T	E	E	W	S	A	A	L	M	M
Norwegian Salmonid Virus	S	E	A	I	P	W	T	R	A	P	A	L
Southern Elephant Seal Virus	E	E	T	V	E	W	S	A	L	T	L	L
Semliki Forest Virus	E	G	S	E	E	W	S	A	P	L	I	T
Sindbis Virus	E	G	T	E	E	W	S	A	A	P	L	V
Una Virus	E	G	T	V	E	W	S	A	P	L	I	S
Venezualan Equine Encephalitis Virus	E	N	C	E	Q	W	S	L	V	T	T	M
Western Equine Encephalitis Virus	E	G	S	E	P	W	S	L	V	T	A	L
Taiforest Alphavirus	E	D	T	V	E	W	S	A	M	V	T	A
Tonate Virus	E	G	T	V	E	W	S	L	V	T	T	L

Figure 2.2.1: Substrate residues of the CP protease from the different members of the genus alphavirus. Substrate sequence of the capsid protease belonging to different alphaviruses depicting the amino acid residues P6-P6' encompassing the scissile bond residues tryptophan and serine (tryptophan and alanine in Mayaro Virus; tryptophan and threonine in Norwegian Salmonid Virus).

In 2004 and 2013, the resurgence of CHIKV in the 22 nations has posed grave threat to the health and economy especially in the developing nations. Currently no vaccine or FDA approved anti CHIKV drug is available. The CHIKV epidemic and related dire economic repercussions can be controlled by repurposing the FDA approved drugs against it. For a compound to pave its way in the market as a drug, it has to undergo various pre-clinical trials which include cytotoxicity testing of the compound, bioavailability, monetary expenditure in the compound synthesis etc. followed by final clinical trials. To overcome these hurdles, it is necessary to find the alternate functions of the already approved FDA drugs instead of designing new compounds against the prevalent diseases. With the advent of bioinformatics and chemiinformatics, repurposing of drugs targeting viral enzymes has become easier. Screening of inhibitors through *in-silico* techniques has shown successful results in *in vitro* biochemical assays and cell-based assays [92], [93], [141], [142]. Recent study has reported the antiviral property of piperazine against CHIKV while piperazine derivatives are already commercially available in market as anti -depressants and anti-histamine [47]. Similar studies related to drug repurposing has been carried against Ebola virus wherein Sinfungin (antibiotic and anti-protozoan) and Indinavir (HIV protease inhibitor) has been repurposed as anti- Ebola drugs by *in silico* approach [143].The current study has predicted few FDA approved compounds as anti CHIKV. The compounds were screened from the LOPAC® library Sigma Aldrich having 1280 pharmacologically active compounds by docking them with the active site of CP protease of AVCP and CVCP. The LOPAC compounds are either already marked drugs or are being used in clinical trials, therefore chances of these molecules surpassing the ADME tests (Absorption, distribution, metabolism, excretion) and blood brain barrier is fairly high thus saving cost and time. In this study, for the first timeby *in silico* crystal structure-based identification andevaluation- of small molecules targeting the protease activity of alphavirus CP has been reported.

2.3 Materials and methods

2.3.1 Expression of CVCP active

The CVCP active (residues106-259) expression protocol was similar to the previously reported methodology for the same [5]. Briefly,transformation of chemically competent *Escherichia coli* strain Rosetta (DE3) cells was done using the recombinant pET28c-CVCP active (110-259 amino acid residues) plasmid preceded by a His-tag, which was cleavable by the TEV Protease present at the N-terminus of recombinant protein. A distinct colony representing the transformed cells was picked and inoculated in 10 ml Luria Bertani broth (LB) supplemented with 50 µg/ml kanamycin and 35 µg/ml chloramphenicol, grown at 37 ºC, overnight at a constant agitation of 200 rpm. Large scale harvesting of cells was done in 1L LB medium along with the 50 µg/ ml kanamycin and 35 µg/ml chloramphenicol. The bulk culture was grown at 37 ºC till the optical density of the culture attained 0.8 at 600 nm (OD_{600}). The culture was induced by 0.4 mM isopropyl-ß-thiogalactopyranoside (IPTG) and was allowed to grow at 18ºC overnight. The overnight grown culture was harvested and was checked for the expression and solubility of CVCP active on 15% sodium dodecyl sulphate (SDS) polyacrylamide gel.

2.3.2 Purification of CVCP active

The purification of CVCP active (residues106-259) was similar to the previously reported active CVCP (residues106-259) [5]. The harvested cell pellet was re-suspended in 30 ml of ice cold lysis buffer (50 mM Tris, 100 mM NaCl, 10 mM imidazole; pH 7.6).The re-suspended pellet was subjected to French press (Constant Systems Ltd, Daventry, England) at 40 KPSI (2700 bar) for the disruption of cells. To segregate insoluble debris from soluble viral recombinant protein along with the bacterial proteins, the solution was centrifuged at 12,000g for 1h at 4°C. CVCP active was separated and isolated from the bacterial proteins by loading the supernatant onto a column containing Ni- NTA beads (Nickel- Nitrilotriacetic acid), pre-equilibrated with the binding buffer (50 mM Tris, 100 mM NaCl; pH 7.6). By exploiting the technique Immobilized Metal Ion Affinity Chromatography (IMAC), CVCP active was purified by running a gradient of imidazole (25 mM to 250 mM). Eluted fractions containing the His-tagged protein were pooled with TEV protease in 1:25 ratio (TEV:CVCP active) and was subjected to dialysis against the buffer having 50mM Tris,20mM NaCl; pH 7.6 overnight at 4ºC. The dialysed sample was reloaded onto the column containing Ni-NTA beads pre-equilibrated with 50mM Tris and 20mM NaCl; pH 7.6 (reverse Ni-NTA). The flow-through had CVCP active withoutHis-tag. The fractions of each step of purification were analysed on 15 % SDS PAGE. The flow-through of reverse Ni- NTA when run on SDS PAGE showed ~ 98 % purity of band of CVCP active when stained by Coomassie brilliant blue (CBB).The eluted protein was concentrated up to 2 mg/ml by Amicon Ultra-15 concentrator (Millipore) having 3 kDa cut-off.

2.3.3 Crystallisation trial of CVCP active

The purified CVCP active was concentrated up to 7mg/ml at 4°C usingAmicon® Ultra-15 centrifugal filteration unit having 3 kDa cut off at 4000 rpm. The concentrated protein in the buffer 50 mM Tris-HCl pH 7.6 and 20 mM NaCl was subjected to crystallisation trial by employing the sitting drop vapor diffusion technique. Various screen conditions of Hampton Research including Peg ion, Crystal screen, Index and Morpheus were used as reservoir buffer. The drop (1X protein: 1X reservoir buffer) used was 2 µl against the 50 µl reservoir buffer. The trays were kept unperturbed for 1 month in the vibration free chamber.

2.3.4 Structure-based Screening and Molecular Docking

The crystal structure of AVCP in apo, complex, active and inactive state and CVCP in inactive (native) state are available in the protein data bank (PDB). In the absence of crystal structure of CVCP active, the conserved C-terminal W261 residue from the coordinate file of the crystal structure of CVCP inactive protein was removed to make the active site accessible for the docking of small molecules. The crystal structure of active Aura virus capsid protease (AVCPΔ2) was used for *in silico* screening of LOPAC1280 library. The grid parameters were optimized by docking of the capsid protease substrate peptide (WS). AVCPΔ2 and CVCP active (residues 106-259) were prepared for molecular docking using AutoDock Tools-1.5.6. Hydrogen atoms and Kollman charges (Kollman charges=11.02) were added to the protein. Atoms of proteins were assigned AD4 type and the protein was saved in .pdbqt file. The name of the ligands and their molecular structure are given in Table 1. The ions were removed from ligands and they were prepared for docking by addition of hydrogen atoms and Gasteiger charges. After assignment of number of rotatable bonds, ligands were saved as .pdbqt file format. The atomic potential grid map was calculated by AutoGrid 4 with a spacing of 0.375Å and the Grid box dimensions of 40Å x 40 Å x 40Å. The center point co-ordinates was set X= -11.497, Y = 3.384 and Z =32.738. Lamarckian genetic algorithm in combination of grid based energy evaluation method was used for docking. The number of total GA runs was set to 50. Other docking parameters were set as default.

2.3.5 Sequence and structure analysis

Multiple sequence analysis was done using Clustal Omega and structural comparisons were done using 350 COOT [144]. The docking images were made using PyMol [145] and LigPlot [146].

2.3.6 CVCP serine protease assay

The protease activity of purified CVCP protein was determined using peptide specific FRET based assay as described by Aggarwal et al., 2015 [44]. The FRET assay substrate was a fluorogenic peptide having DABCYL (quencher) and EDANS (fluorophore) as FRET pairs present at N- and C- terminus, respectively. The fluorogenic peptide sequence DABCYL-Gly-Ala-Glu-Glu-Trp↓Ser-Leu-Ala-Ile-Glu-EDANS (Biolinkk New Delhi, India) has the scissile bond between Trp and Ser which is highly conserved among alphaviruses. EDANS, the fluorophore conjugate in the peptide has excitation maxima at 340 nm and emission maxima at 490 nm, while the conjugated DABCYL is a non-fluorescent quencher which absorbs in the range of 470 - 520 nm. The proteolytic assay was conducted in a 96- well black round bottom non polystyrene plates at 25 ºC in 100 µL reaction mixture containing 20 mM HEPES pH 7.0 and 2 µM purified protein with a fixed concentration of substrate, and fluorescence was analysed at excitation and emission wavelength of 340 nm and 490 nm respectively using a multimode plate reader Cytation 3 (BioTek Instruments, Inc.). The negative control was reaction without the enzyme and all the reactions were conducted in triplicate. A curve was plotted using relative fluorescence units RFU (Fluorescence measured of enzymatic reaction at fixed period of time minus the fluorescence measured of the control) at different intervals of time. The resultant slope of the curve is designated as the fluorescence extinction coefficient (FEC) which determines the amount of product formed or the amount of cleaved peptide formed in the given period of time (RFU/min). Additionally conversion factor is obtained by plotting a curve using fluorescence produced by substrate at varying concentration. The resultant slope of the curve was taken as the conversion factor (RFU/µM). For the quantification of fluorescence values, determination of initial velocity of enzyme is required which was calculated by dividing the conversion factor from the fluorescence extinction coefficient (RFU/min) / (RFU /µM) to get µM concentration of peptide cleaved per min (µM/min).

2.3.7 CVCP serine protease inhibition assay

To label the docking results, the peptide specific FRET based inhibition assay against the capsid protease of CVCP was performed. The compounds with showed maximum binding energy against the active site of CVCP were purchased from Sigma Aldrich. Stocks of all the compounds except DRV (soluble in 100% DMSO) were made in the Milli-Q water. The working concentration of DRV had <1% (v/v) DMSO. The *in vitro* CVCP inhibition assay was conducted in the 96- well black round bottom non polystyrene treated plate. The purified CVCP active and inhibitor were pre incubated for 30 min at 25 ºC in the 20 mM HEPES pH 7.0 (reaction buffer), then the substrate was added in a 100µL reaction and fluorescence readings were taken. For the comprehensive study of kinetic mechanism of CVCP inhibition, the reaction range consisted of varying concentrations of substrate and inhibitor. The inhibition of enzyme was marked by observing the relative decrease in fluorescence unit of enzyme-inhibitor reaction for a fixed period of time (compared with control having 0 µM inhibitor). The decrease in fluorescence implies the reluctance in dissociation of the FRET pair because of the plausible binding of inhibitor at the active site of CVCP. By the help of Lineweaver-Burk plot, mode of inhibition was determined for each inhibitor by plotting reciprocal values of initial velocity and substrate concentration by linear regression. The kinetic inhibition constant K_i was calculated by plotting initial velocities versus varied concentrations of substrate by employing non-linear regression analysis using GraphPad prism 5 software.

2.3.8 Cells, virus and compounds

Vero cells (NCCS, Pune, India) were maintained in Dulbecco's-modified essential media (DMEM; HiMedia, India) supplemented with 10 % heat-inactivated fetal bovine serum (FBS; Gibco, USA) and 1X penicillin-streptomycin solution and incubated at 37 ºC with 5% CO_2. A clinical isolate of CHIKV [93] was used for the present study. The CHIKV isolate was propagated into Vero cells and virus titer was quantified by plaque forming unit per milliliter (pfu/ml). AP4, EAC, PSU and DRV were dissolved in water except DRV. DRV was dissolved in 100 % dimethyl sulfoxide (DMSO). The stock concentration of compounds was 10 mM. The working concentrations for each compound were prepared in cell culture medium from their respective stocks.

2.3.9 Cytotoxicity studies

To measure the cytotoxicity level of AP4, EAC, PSU and DRV in Vero cells, a colorimetric assay (MTT assay) was done before testing the antiviral activity of the compounds. Briefly, Vero cells were seeded into a 96-well plate at a seeding density of 2×10^5 cells/well and incubated at 37 °C with 5% CO_2. Different concentration (0.05 mM to 5 mM) of each compoundwith control (without compounds) was diluted in culture media and added to the cells in the respective wells in triplicate. After 48 h, the medium was gently aspirated and 5mg/mL MTT (3-(4, 5-Dimethylthiazol- 2-yl)-2, 5- Diphenyltetrazolium Bromide) solution was added to each well, and incubated for 4 h at 37 °C. The formazan crystals were dissolved with DMSO after removing media. The absorption values were measured at 570 nm using multi-mode plate reader Cytation 3 (BioTek Instruments, Inc.). The percentage cell viability was then calculated against the control well with cells not treated with any compound.

2.3.10 Assessment of antiviral activity

Prior to infection, Vero cells were seeded in a 24-well plate with 2×10^5 cells/well seeding density. AP4, EAC and PSU were added to wells in duplicate at final concentrations of 1mM, 0.5mM, 0.1mM and 0mM while DRV was added in 100 µM, 50 µM, 10 µM and 0 µM concentrations. For the pretreatment assay, cell monolayer were treated with the compounds for

4 h at 37°C followed by washing with PBS and infection with CHIKV at multiplicity of infection (MOI) of 1. After 1.15 h virus adsorption, cells were incubated in DMEM with 2 % FBS at 37°C with 5% CO_2. In the simultaneous treatment assay, Vero cells were infected with CHIKV at an MOI of 1 and were simultaneously treated with different concentrations of compoundsfor 1.15 h. The treated cells were washed with PBS and incubated in maintenance medium for 24 h at 37°C with 5% CO_2. In the post treatment assay, Vero cells were infected with CHIKV at an MOI of 1 for 1.15 h. After virus adsorption, cells were treated with the different concentrations of compounds in DMEM with 2% FBS and incubated at 37 °C with 5% CO_2 for 24 h. After 24 h, supernatants were harvested for plaque assays and cells were harvested for viral RNA quantification for all three treatments. All experiments were repeated thrice.

2.3.11 Plaque based reduction assay

Plaque assay was performed for the detection of antiviral activity of compounds. Briefly, Vero cells were seeded at desired cell density in 24-well tissue culture plate. At optimum confluency (>90%), 10-fold serial dilutions of harvested supernatant was used for infection of Vero cells. After 1.15 h virus adsorption, cells were overlaid with 1% carboxymethyl cellulose (CMC) in MEM with 2% FBS and incubated at 37ºC with 5% CO_2 for 48 h. Overlay medium was discarded and the cells were fixed with 10% formaldehyde. The fixed cells were stained with 1% crystal violet solution for plaque visualization. The plaques were counted and quantified in pfu/ml. The percentage of inhibition based on plaque assay was determined for each compound at different concentrations.

2.3.12 Inhibitory potency

Dose-response curves were generated by measuring pfu/ml in plaque assay for a range of compound concentrations. Six different concentrations (0 μM-100 μM) for each compound were used to generate inhibition curves for calculating their half maximal effective concentration (EC_{50}). Briefly, Vero cell were seeded in 24-well plates at desired cell density. At optimum cell confluency, cell monolayer were infected with CHIKV at 1 MOI and incubated for 1.15 h at 37 °C with 5% CO_2. After adsorption, varying concentration of each compound in DMEM with 2% FBS was added to the monolayer cells and incubated for 24 h. The supernatant were harvested and reduction in virus titer was determined by plaque assay as described above. EC_{50} of each compound were calculated by plotting the graph of percentage inhibition vs. compound concentration.

2.3.13 Effect of inhibitors on viral RNA synthesis

To detect the effect of inhibitors on the level of viral RNA synthesis, a minus-strand-specific reverse transcriptase PCR (RT-PCR) assay was done in infected cells as described by Mayuri et al., 2008 [147]. Briefly, Vero cell were seeded in 6-well plates at desired cell density. At optimum cell confluency, cell monolayer were infected with CHIKV at 10 MOI and incubated for 1.15 h at 37 °C with 5% CO_2. After adsorption, 100 µM concentration of each compound in DMEM supplemented with 2% FBS was added to the monolayer cells and were incubated for 6h and 12h. The cells were harvested for detection of minus-strand RNA synthesis. Total RNA was extracted from the infected cells using Tri-reagent (MRC) according to the manufacturer's instructions and treated with RQ1 RNase-free DNase I (Promega) as per manufacturer's instructions (Promega, USA). The quality and quantity of the extracted RNA was checked by NanoDrop spectrophotometer and were used in a two-step minus-strand-specific RT-PCR. The minus-strand specific cDNA synthesis was performed using AccuScript High Fidelity cDNA synthesis Kit (Agilent). For each reaction per sample, 100 ng of RNA and CHIKV E1 plus-sense primer bearing sequence (5'-GGAAATAACATCACTGTAACTGCCTATGCAAACG-3') were used. The minus-strand specific cDNA was serially diluted tenfold in water and used in the two-step RT-PCR. The RT-PCR was done with herculase II fusion DNA Polymerase using CHIKV E1 plus-sense and minus-sense primers (5'-GCATAGCACCACGATTAGAATC-3') [93]. The difference in the level of RNA synthesis was analysed by checking the intensity of RT-PCR products in 1.5% Tris-Acetate-EDTA agarose gel (AGE) using ethidium bromide stain on the Gel Doc™ EZ Image system (Bio-Rad) using Image Lab Software 5.1.

2.3.14 Western Blot analysis: To validate the inhibitory potential of AP4, EAC and PSU on the protease activity of CP in CHIKV infected cells.

The CVCP expression to indirectly measure the amount of CP released as a result of autoproteolytic activity of CP was examined by performing Western immunoblot on the ce;; lysates of infected Vero cells treated with AP4, EAC or PSU. Before infection, Vero cells were pre-treated with 100 µM AP4, EAC or PSU separately and incubated for 6 h. After 6 h incubation, the cells were infected with MOI 1, CHIKV 181/25 virus for 1.5 h. Post CHIKV infection after 1.5 h, washing was done thrice with 1X PBS to remove the excess virus and 100 µM compounds were added separately in the DMEM media supplemented with 2% FBS in the respective wells for 12 h. After 12 hpi, the supernatant was removed and cells were lysed using lysis buffer (RIPA [50mM Tris HCl, PH 7.4, 1% Triton X-100, 0.5% Sodium deoxylcholate, 0.1% SDS, 1mM EDTA,10mM NaF) and samples were separated on a 10% SDS-PAGE.The seperated proteins were electrophoretically transferred to the nitrocellulose membrane and were probed with SINV-specific rabbit polyclonal antibody anti-CP (1:3000). Mouse anti-actin monoclonal antibody (1:3000) (Chemicon) was used as an internal control. Infrared-labelled (IRDye 680 and IRDye 800) goat anti-rabbit and goat anti-mouse secondary antibodies (Li-Cor) were used for the detection of proteins. Finally, the blots were scanned with the Odyssey infrared imager (Li-Cor, Lincoln, NE), and protein bands were quantified using Odyssey software version 3.

2.3.15 Immuno-fluorescence assay

Vero cells were seeded in 24-well plates at a density of 2×10^5 cells/well and were infected with CHIKV at a MOI of 1 for 1.15 h. After virus infection, each compound (AP4, EAC, and PSU) at a final concentration of 100μM were added in duplicate and incubated at 37°C with 5% CO_2 for 24 h. The supernatant was discarded and cells were washed with PBS. Then cells were fixed with ice cold 1:1 solution of methanol and acetone at 25°C for 15 min, followed by permeabilization with 0.1% TritonX-100 in PBS for 10 min at 25°C. After washing, cells were incubated with 1:100 dilution of anti alphavirus mouse monoclonal antibody (Santa Cruz Biotechnology, USA, #sc-58088) for 1 h at 25°C. Cells were washed with PBS and incubated with 1:500 anti-Mouse IgG (whole molecule)– fluorescein isothiocyanate (FITC) antibody produced in goat (Sigma, #F9006) for 1 h at 25°C. For staining of nucleus, cells were treated with 4′, 6-diamidino-2-phenylindole (DAPI) and using inverted fluorescence microscope (Leica) images were captured.

2.3.16 Quantitative real time PCR (qRT-PCR)

Total RNA from the harvested cells was isolated using Trizol reagent in accordance with the manufacturer’s instructions. The quality and quantity of the extracted RNA was checked by nanovue NanoDrop spectrophotometer and DNase I treatment was conducted as per manufacturer’s instructions (Promega, USA). First strand cDNA synthesis was carried out by using commercially available AccuScript High Fidelity cDNA synthesis Kit (Agilent) as per manufacturer’s protocol with equal amount (100 ng) of RNA for all the reactions. This cDNA was used subsequently to amplify E1 gene along with beta-actin as internal control using KAPA SYBR fast universal qPCR kit on QuantStudio™ 5 System (Applied Biosystems, Carlsbad, CA) using standardized protocol in triplicate. The details of primer used for amplification of E1 and beta-actin were described in Singh et al., 2018 [93]. ΔΔCt method was used to calculate relative values for each compound and the fold change was calculated relative to CHIKV infected control

2.4 Results

2.4.1 Expression of CVCP active (110-259 residues)

The expression of CVCP active in the *Rosetta* cells was optimum when the culture exhibiting optical density 0.8 at 600nm (OD600) was induced by 0.4 mM IPTG and was allowed to grow for 16 h at 18º C. The overexpressed band of CVCP active of approximately 16 kDa is shown in Fig 2.4.1.1. Optimum expression of soluble CVCP active ~16 kDa is seen in the lane in which the supernatant of induced culture was run.

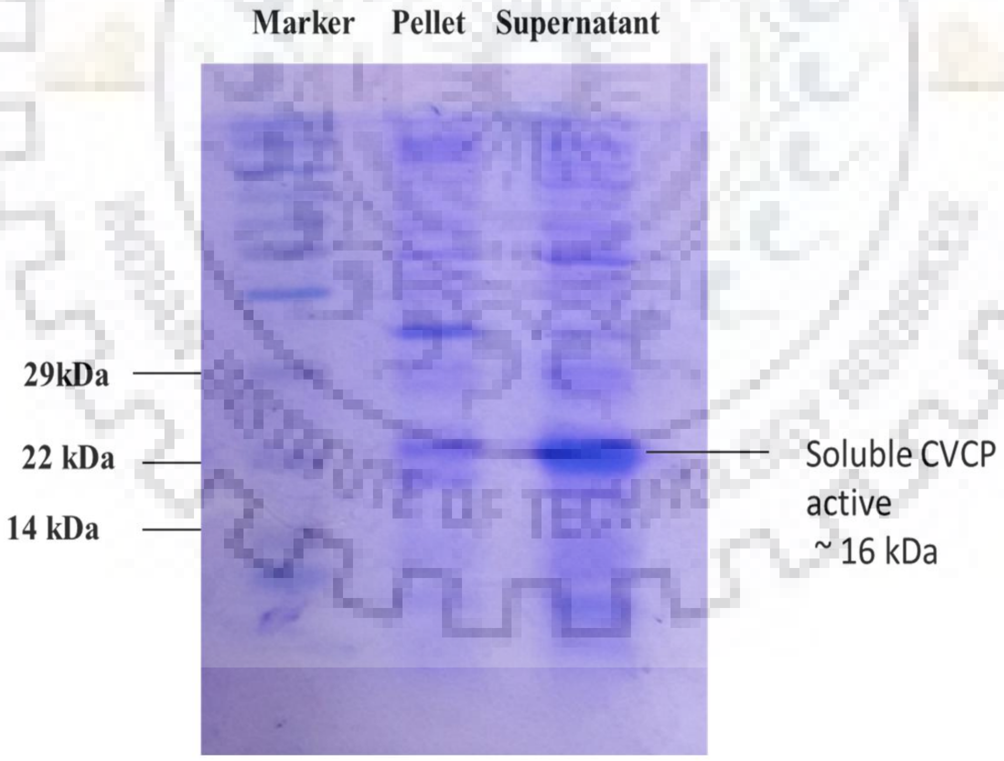
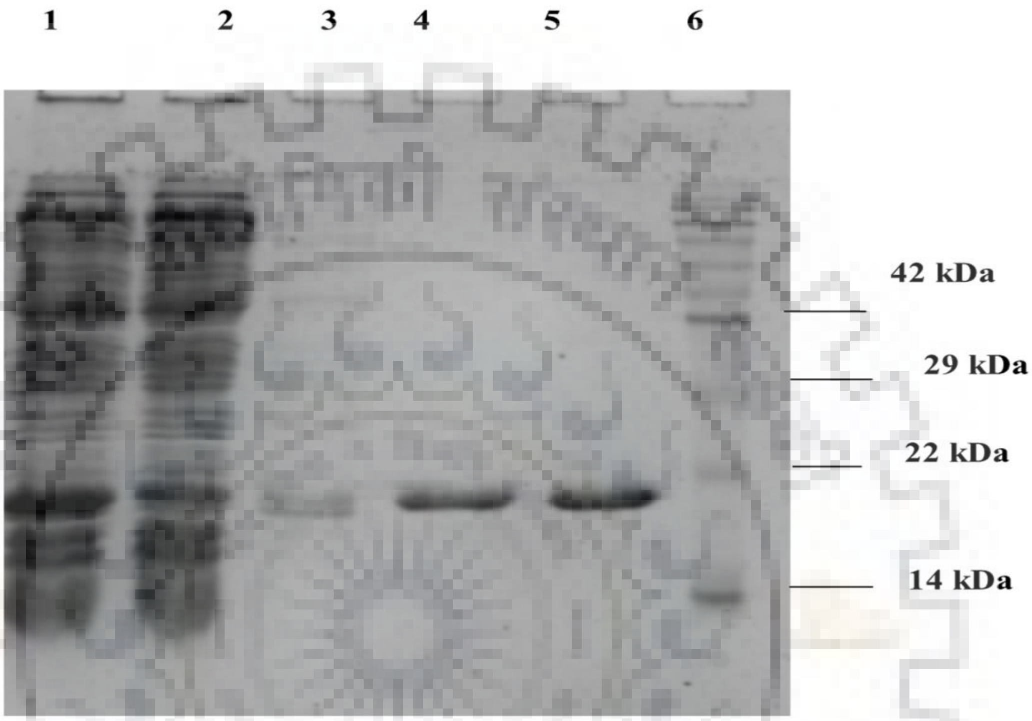


Figure 2.4.1.1.SDS PAGE analysis of expression of CVCP active.

2.4.2 Purification of CVCP active by Ni-NTA column

The soluble CVCP active in the supernatant was purified using Ni-NTA chromatography. Figure 2.4.2.1 shows SDS PAGE (15%) analysis of purified CVCP active of approx. 16 kDa, in the wash 2 and elution fraction (lane 4 and 5).



2.4.2.1.SDS PAGE analysis of purification of CVCP active. Lane 1: Supernatant, Lane 2: Flowthrough, Lane3: Wash 1 (25 mM imidazole, 0.5 M NaCl), Lane 4: Wash 2 (50mM imidazole, 0.5M NaCl), Lane 5- Elution (250mM imidazole), Lane 6: Pre stained Protein ladder of Thermo Scientific (175 kDA-10.5 kDa). Wash 2 and Elution has purified His tagged CVCP active ~16 kDa.

2.4.3 Reverse Ni -NTA of CVCP active

The His tag from the CVCP active was removed by the TEV Protease and purified cleaved CVCP active was obtained in the flowthrough of Reverse Ni- NTA. SDS PAGE (15%) shows the cleaved CVCP active ~16 kDa in Figure 2.4.3.1. The concentration of purified protein was measured via UV absorbance spectroscopy. The molar extinction of CVCP active was calculated online from Protparam tool (https://web.expasy.org/protparam/). The molar extinction coefficient of CVCP active is 18575.

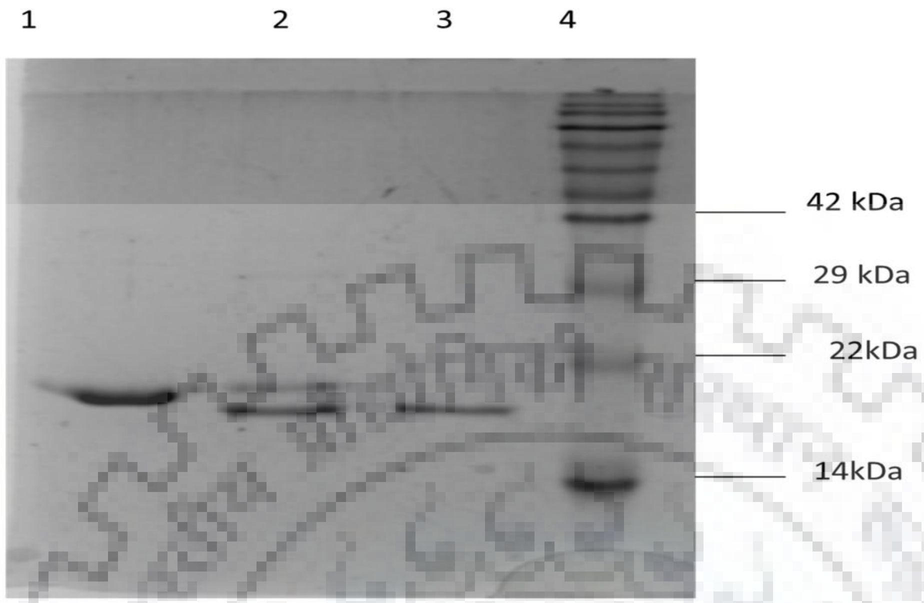


Figure 2.4.3.1. SDS PAGE profile of reverse Ni-NTA showing cleaved CVCP active ~16 kDa. (Lane1 : His tagged CVCP active (uncleaved) , Lane2 : cleaved sample after dialysis (before Reverse Ni), Lane 3- Flowthrough (cleaved CVCP active), Lane 4: Protein ladder (Prestained 175 kDa to 10.5 kDa thermoscientific).

The cleaved purified CVCP active was obtained in the flowthrough of the Reverse Ni- NTA. It was collected and concentrated for crystallization, and for enzyme actiivty.

2.4.4 Crystallisation of CVCP active

The main objective of this study was to determine the crystal structure of the active form of CVCP and use the active site conformation for virtual screening of compound libraries to identify potential CHIKV CP protease inhibitors. To determine the three dimensionalstructure and gain insights into the structural architecture of the active site and substrate binding site in the active form of CVCP, crystallisation trials were done using purified and concentrated protein. For crystallization trials protein was concentrated to ~7 mg/ml. Figure 2.4.4.1 shows the concentrated CVCP active protein. Crystal hits of ~ 7 mg/ml CVCP active were obtained by vapour diffusion method at 20 ºC in 28% PEG 3350, 0.1 M sodium citrate, pH 6.8 . The reservoir buffer had 100 µl volumes against the 2 µl drop (1µl reservoir buffer: 1µl protein).

100 µM octyl β-D glucopyranoside (BOG) in the drop was used as an additive for improving the size and quality of crystals. Diffractable crystals of CVCP active could not be obtained in the tentative predictable conditions.

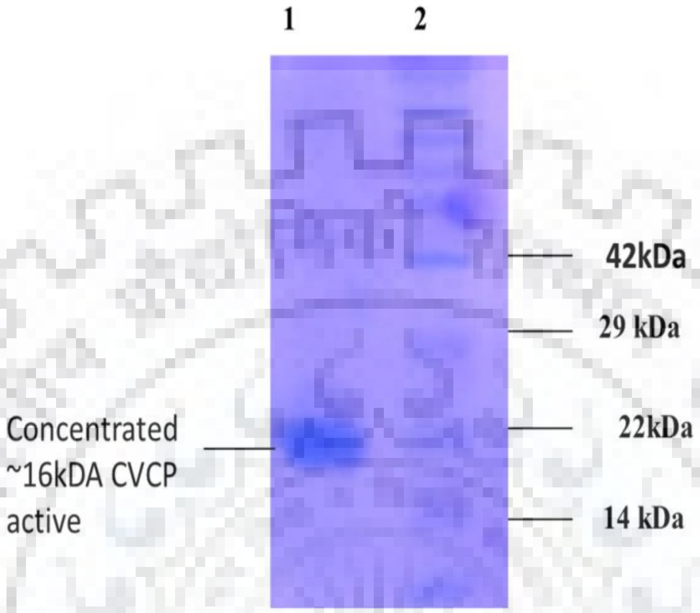


Figure 2.4.4.1.SDS PAGE profile of concentrated CVCP active.

2.4.5 Molecular Docking

In the absence of the crystal structure of the active form of CVCP, the pdb coordinates of the recently determined crystal structure of the inactive CHIKV CP were used. The virtual screening of potential alphavirus CP (AVCP) inhibitors from the LOPAC library of 1280 FDA approved drugs led to the identification of four compounds: AP4, EAC, DRV and PSU that displayed optimal binding affinity with the active site of AVCP. These selected compounds were docked into the active site of.CVCP in which the last residue was removed computationally and binding of compounds was compared to the substrate peptide (WS).Docking studies were performed using substrate/peptide (WS), EAC, AP4, PSU and DRV as ligands and the binding interactions of these molecules into the active site AVCP and CVCP protease were analysed in details. Insights into the binding of substrate with AVCP and CVCP protease active site are shown in Fig. 2.4.5.1 (a, b, c, d) and Table 2.1. Interestingly, the binding energy of EAC (-7.53 kcal/mol), AP4 (-6.24 kcal/mol), DRV (-6.38 kcal/mol) into the AVCP active site was found to be more than that of the substrate (-5.01 kcal/mol). On the other hand, the binding energy of EAC (-7.52 kcal/mol), AP4 (-6.41 kcal/mol), DRV (-6.49 kcal/mol) into the CVCP active site was also found to be more than that of substrate (-6.08 kcal/mol) (Table 2.2). Although the binding energy obtained for PSU for both AVCP and CVCP active proteins was less than the binding energy of the dipeptide WS substrate. However, testing of inhibitory activity of PSU against alphavirus protease using FRET based assayshowed inhibition. Using LigPlot the molecular interactions of substrate (WS), EAC, DRV, AP4 and PSU with the active site residues of CVCP were analyzed (Fig. 2.4.5.1). The analysis of hydrogen bond interactionsof inhibitors in the active site of AVCP and CVCP protease was summarized in Table 2.3. In Table 2.4, the analysis of hydrophobic interactions of inhibitory molecules with the active site of AVCP and CVCP protease was given.

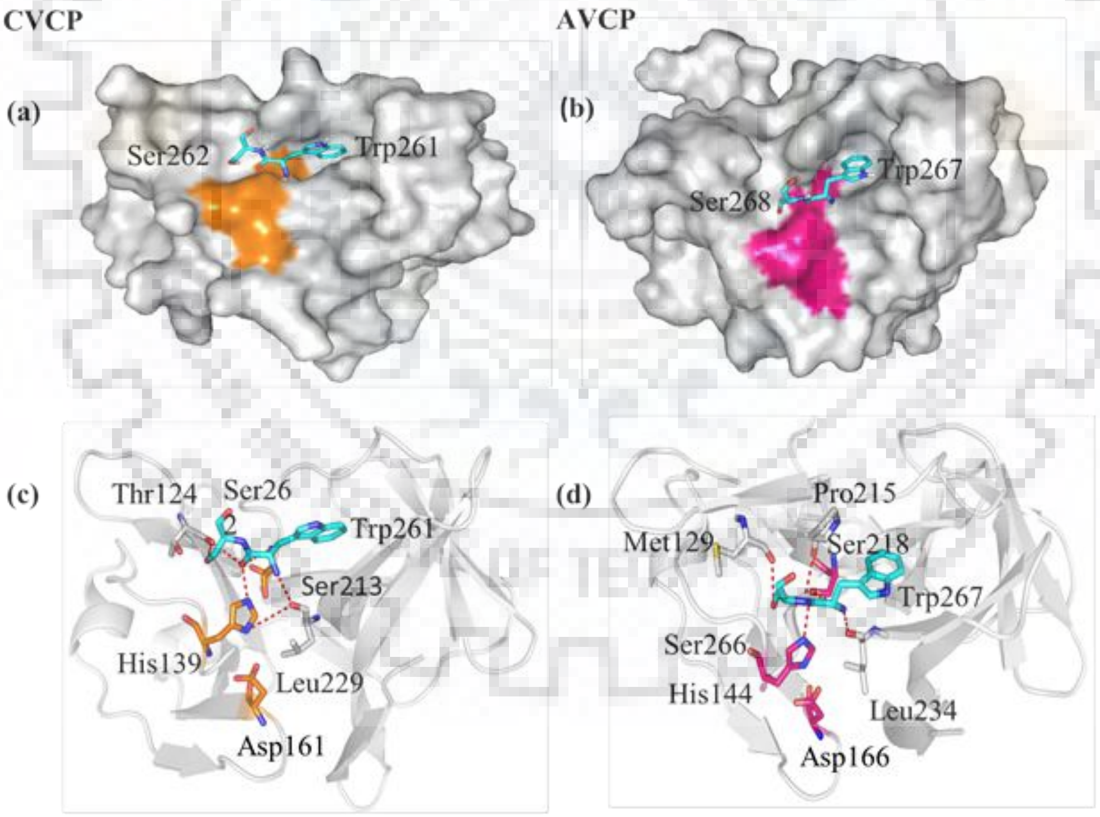


Figure 2.4.5.1. Binding analysis of substrate into AVCP and CVCP protease active site- (a) Surface view representation of CVCP protease and binding of substrate (stick) into protease active site, (b) Surface view representation of AVCP protease and binding of substrate (stick) into protease active site, (c) Polar interactions of CVCP protease active site and adjoining

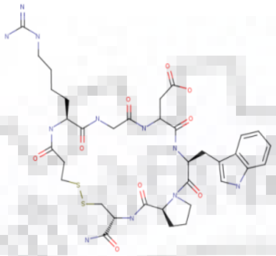
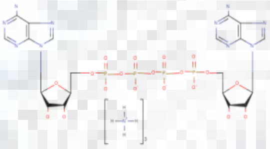
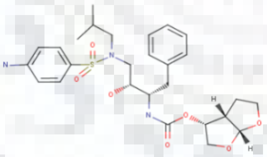
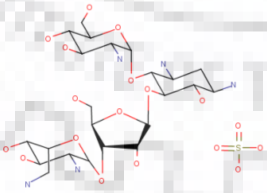
residues with substrate and (d) Polar interactions of AVCP protease active site and adjoining residues with substrate.

Table2.1 Analysis of molecular interactions of substrate (WS) binding into the active site of AVCP and CVCP protease.

AVCP		**CVCP**	
N; Trp265	O; Leu234	N; Trp261	O; Leu229
N; Ser266	O; Pro215	O; Trp261	O; Thr124
O; Trp265	NE2; His144	O; Trp261	NE2; His139
O; Ser265	O; Met129	O; Trp261	N; Ser213
N; Trp265	OG; Ser218	OG; Ser62	O; Thr124

Interestingly, the binding energy of EAC (-7.53 kcal/mol), AP4 (-6.24 kcal/mol), DRV (-6.38 kcal/mol) into the AVCP active site was found to be more than that of the substrate (-5.01 kcal/mol). On the other hand, the binding energy of EAC (-7.52 kcal/mol), AP4 (-6.41 kcal/mol), DRV (-6.49 kcal/mol) into the CVCP active site was also found to be more than that of substrate (-6.08 kcal/mol) (Table 2.1).

Table 2.2 List of best hits obtained from screening of LOPAC^{1280} library

Compound ID	Compound Name	Cat. No.	Structure	Selectivity	Description	Binding energy (kcal/mol)	
						AVCP	CVCP
645	Eptifibatide acetate	SM11042		GPIIb/IIIa	Inhibitor for fibrinogen binding to platelet membrane integrin glycoprotein GPIIb/IIIa receptor	-7.53	-7.52
41	P^1, P^4-Di (adenosine-5') tetraphosphate triammonium	D1262		Inhibitor	Competitively inhibits ADP-induced platelet aggregation	-6.24	-6.41
340	Darunavir	SM10937		HIV protease	HIV protease inhibitor with broad spectrum activity	-6.38	-6.49
1169	Paromomycin sulphate	P9297		Protein synthesis	Oligsaccharide type antibiotic; interferes with protein synthesis	-2.16	-1.68

Using LigPlot the molecular interactions of substrate (WS), EAC, DRV, AP4 and PSU with the active site residues of CVCP were analyzed (Figure 2.4.5.2). The analysis of hydrogen bond interactions of inhibitors in the active site of AVCP and CVCP protease is summarized in Table 2.3. In Table 2.4, the analysis of hydrophobic interactions of inhibitory molecules with the active site of AVCP and CVCP protease is given.

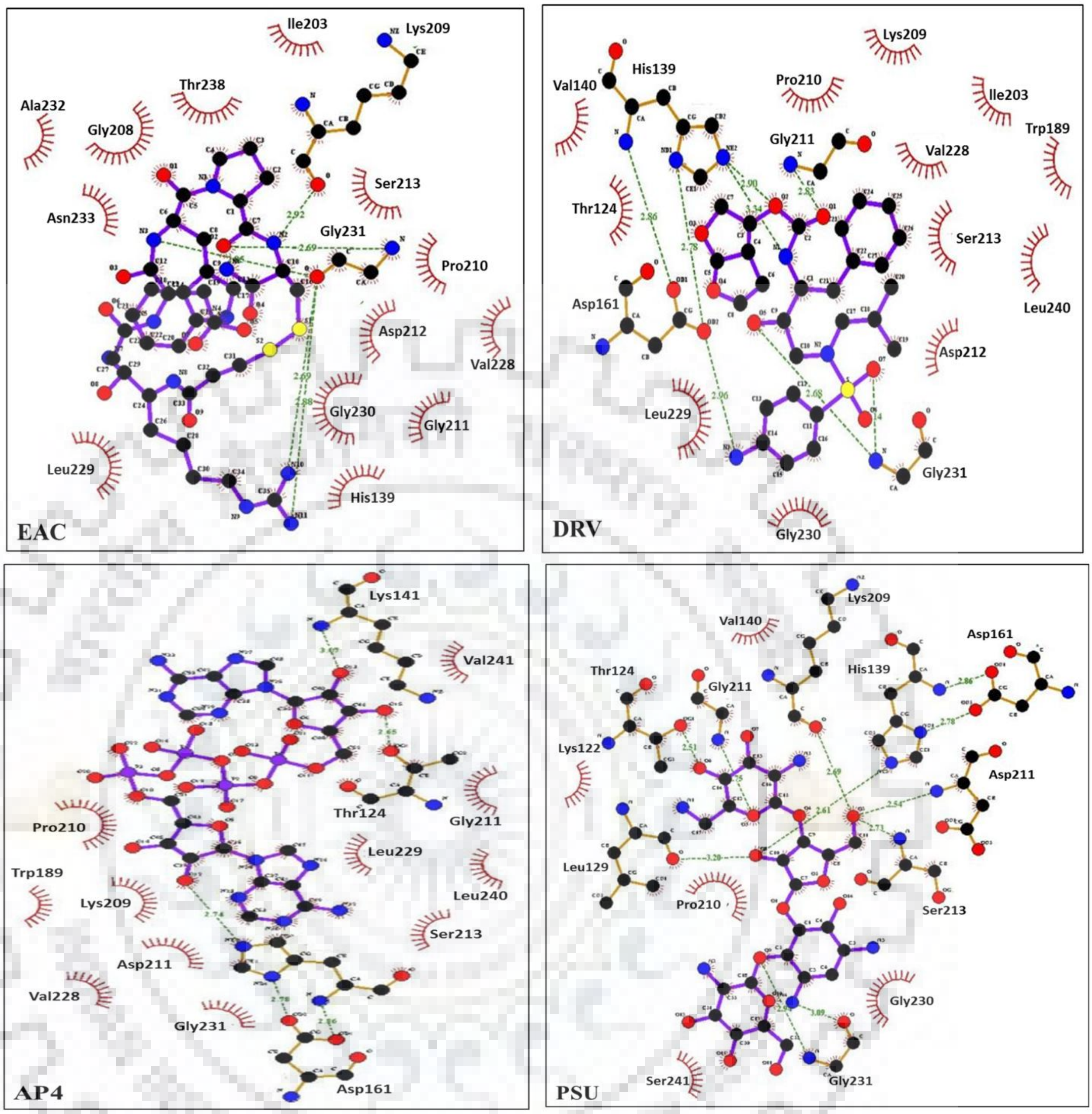

EAC
Ile203
Lys209
Thr238
Ala232
Gly208
Ser213
Asn233
Gly231
Pro210
Asp212
Val228
Gly230
Gly211
Leu229
His139
DRV
Val140
His139
Lys209
Pro210
Ile203
Trp189
Gly211
Val228
Thr124
Ser213
Asp161
Leu240
Asp212
Leu229
Gly231
Gly230
AP4
Lys141
Val241
Thr124
Gly211
Pro210
Leu229
Trp189
Lys209
Leu240
Ser213
Asp211
Val228
Gly231
Asp161
PSU
Lys209
Val140
Asp161
Thr124
Gly211
His139
Lys122
Asp211
Leu129
Pro210
Ser213
Gly230
Ser241
Gly231

Figure 2.4.5.2. Molecular interactions of EAC, DRV, AP4 and PSU with CVCP active site shown by LigPlot representation. The residues near the active site of CVCP involved in hydrophobic interactions with EAC, DRV, AP4 and PSU are shown in red color. The polar interactions of EAC, DRV, AP4 and PSU with the active site residues and nearby residues are shown with green dash lines.

Table 2.3: Hydrogen bond interactions analysis of inhibitors in the active site of AVCP and CVCP protease

Compound Name	**Hydrogen bond interactions**			
	AVCP		**CVCP**	
Eptifibatide acetate (EAC)	EAC	AVCP	EAC	CVCP
	N6	O; His261	N2	O; Lys209
	N	OC; Leu234	O2	N; Gly231
			N10	O; Gly231
			N11	O; Gly231
P^1,P^4-Di(adenosine-5') tetraphosphatetriammonium (AP4)	AP4	AVCP	AP4A	CVCP
	O23	N2; Lys127	O15	OG1; Thr124
	N34	O; Pro215	O13	N; Lys141
			O12	NE1; His139
Darunavir (DRV)	DRV	AVCP	DENV	CVCP
	O1	NE2;His144	N3	OD2; Asp161
	O1	OG; Ser218	O2	NE2; His139
	N1	O; Pro215	N1	NE2; His139
	O5	O; Leu234	O1	N; Gly211
	O7	N; Gly236	O7	N; Gly231
	N3	OD2; Asp166	O5	N; Gly231
Paromomycinsulphate (PSU)	PSU	AVCP	PSU	CVCP
	O10	N; Gly236	O3	O; Lys209
	N4	O; Gly136	O1	NE2; His139
			O3	N;Asp212
			O3	N; Ser213
			O9	N; Gly231
			O1	O; Leu229
			O6	OG1; Thr124
			O5	N; Gly211

Leu234, Gly236, Gly231, His 139 and Threonine124 residues present in the active site of AVCP and CVCP actively participate in the H-interactions as these are the common interactions as observed in the table 2.3. The common H-intercations with the compounds are by the polar residues, charged residues and hydrophobic residues present in active site of AVCP and CVCP. The residues residing in the active site of AVCP and CVCP which show non common interactions with the compounds are of major significance and should be potentially considered as active targets in the rational drug designing.

Table 2.4: Analysis of hydrophobic interactions of inhibitory molecules with the active site of AVCP and CVCP protease

Compound Name	**Hydrophobic interactions**	
	AVCP	**CVCP**
Eptifibatide acetate (EAC)	Lys258, His144, Trp194, Gly235, Ala237, Asp166, Pro215, Gly211, Thr243, Ile208, Leu245, Gly213, Gly214, Gly236	His139, Ile203, Gly205, Pro210, Gly211, Asp212, Ser213, Val228, Leu229, Gly230, Ala232, Asn233, Thr238
P1,P4-Di (adenosine-5') tetraphosphate triammonium (AP4A)	Met129, His144, Trp194, Asp217, Gly213, Gly214, Ser218, Gly235, Gly236, Ser246	Val140, Trp189, Lys209, Pro210, Gly211, Asp212, Ser213, Val228, Leu229, Gly231, Leu240,
Darunavir (DRV)	Met129, Val145, Trp194, Ile208, Gly213, Gly214, Gly235	Thr124, Val140, Trp189, Ile203, Pro210, Asp212, Ser213, Val228, Leu240, Leu229, Gly230
Paromomycin sulfate (PSU)	Met129, His144, Val145, Gly214, Pro215, Ser218, Leu234, Gly235	Lys122, Val140, Pro210, Gly230, Ser241

Among the hydrophic residues in the active site of AVCP and CVCP mainly glycine actively participates in the hydrophic interactions with the inhibitors. Apart from hydrophobic residues charged residues especially serine is responsible for hydrophobic interactions with the inhibitors followed by the polar and amphipathic residues. The hydrophobic residue Ala 237 of active site of AVCP and Asn233 polar residue of CVCP active exhibit non- common hydrophobic interactions with the EAC which can be exploited in the rational drug designing against AURAV and CHIKV respectively.

Table 2.5 Amino acid residues of active site of AVCP and CVCP which show common hydrophobic interactions with the drugs

Amino acid residues of AVCP	**Drugs**	**Amino residues of CVCP**	**Drugs**
His144	EAC, AP4 and PSU	Ileu 203	EAC and DRV
Trp194	EAC, AP4 and DRV	Pro 210	EAC, AP4 and PSU
Gly236	EAC and AP4	Gly 211	EAC and AP4
Pro215	EAC and PSU	Asp212	EAC, AP4 and DRV
Ile 208	EAC and DRV	Ser213	EAC, AP4 and DRV
Gly213	EAC, AP4 and DRV	Gly230	DRV and PSU
Gly 214	EAC, AP4, DRV and PSU	Val140	DRV and PSU
Gly235	EAC, AP4, DRV and PSU	Val228	AP4 and DRV

2.4.6 FRET based CVCP inhibition assay

With the main objective to block the auto-proteolytic cleavage of the CP from the N-terminus of the structural polyprotein, which will lead to defect in the further processing of the structural glycoproteins and inhibit virus replication, the *in silico* identified molecules were tested for anti-protease activity using the FRET based assay. First, these molecules were preliminary tested in a protease inhibition experiment which revealed significant reduction in fluorescence validating their inhibitory action against CVCP. Further, the FEC and the initial reaction rates of CVCP protease and in presence of inhibitory compounds were determined at varied substrate concentration to calculate the kinetic parameters.The fluorescence signal was reduced with the ascending concentration of compounds, which confirmed the inhibition of CVCP protease. Double reciprocal plot of CVCP protease with the compounds showed competitive inhibition for AP4 and DRV, mixed inhibition for EAC and non-competitive inhibition for PSU (Fig. 2.4.6.1 a to d). The inhibitor constant K_i of the compounds was calculated by fitting the data into the Michaelis-Menten equation. The K_i values calculated for compounds AP4, EAC, DRV and PSU were 133.6±27.45µM, 135±76.25µM, 147.4±10.93µM and 353.0±17.14µM respectively.

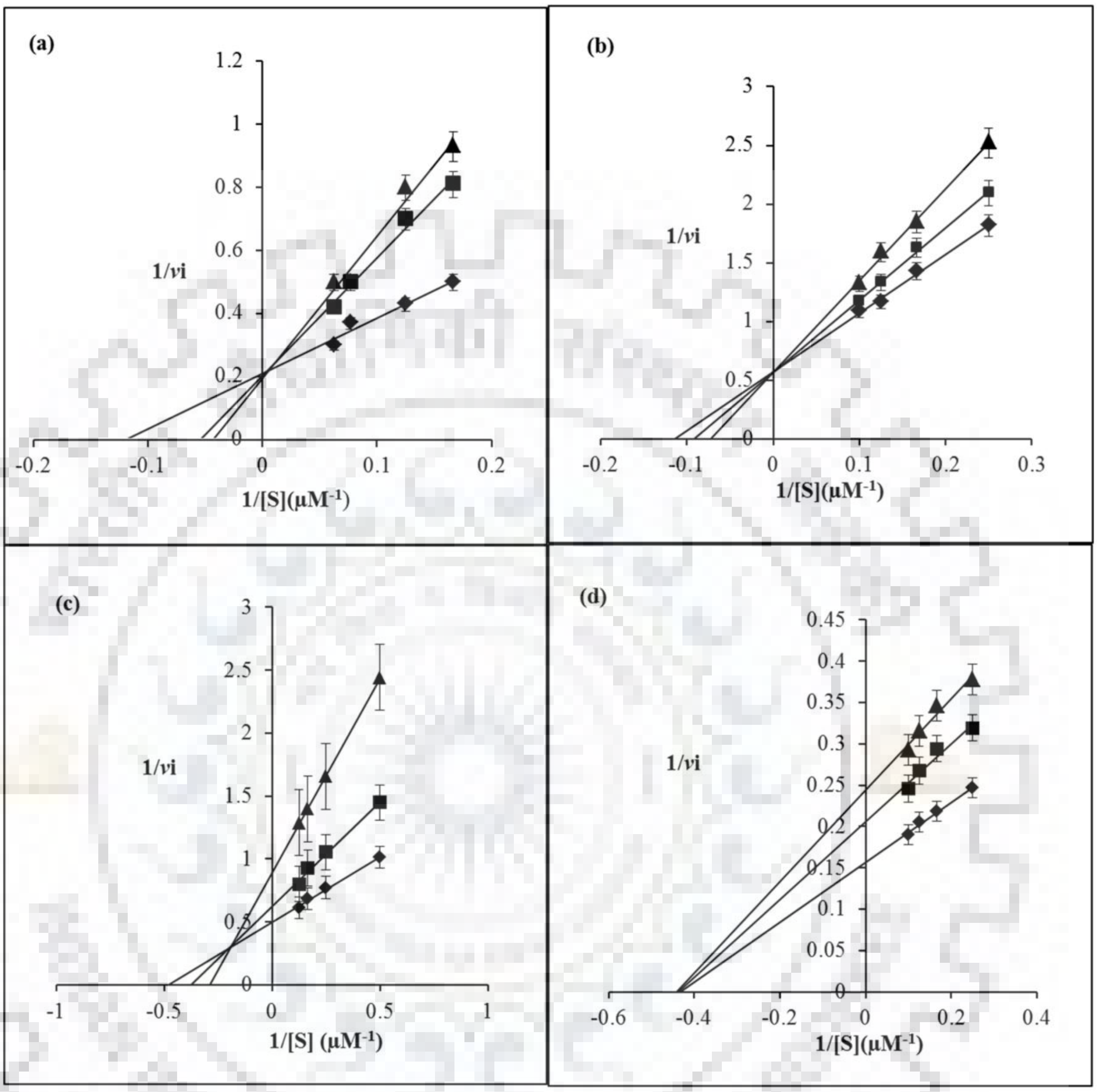


Fig. 2.4.6.1 Double reciprocal plots of CVCP protease inhibition by the LOPAC® screened compounds. (a) CVCP protease- AP4 inhibition kinetics; Lineweaver Burk plot showing the rate of CVCP protease- AP4 complex vs increase in substrate concentration. The mode of inhibition depicted is competitive inhibition(‘i’= ♦- 0 µM, ■- 100 µM, ▲-200 µM), (b) CVCP protease- DRV inhibition kinetics; Lineweaver Burk plot showing the rate of CVCP protease- DRV complex vs increase in substrate concentration. The mode of inhibition depicted is competitive inhibition (‘i’= ♦- 0 µM, ■- 33 µM, ▲- 80 µM), (c) CVCP protease- EAC inhibition kinetics; Lineweaver Burk plot showing the rate of CVCP protease-EAC complex vs increase in substrate concentration. The mode of inhibition depicted is mixed inhibition(‘i’= ♦- 0µM, ■- 100 µM, ▲-200 µM), (d) CVCP protease- PSU inhibition kinetics; Lineweaver Burk plot showing the rate of CVCP protease- PSU complex vs increase in substrate concentration. The mode of inhibition depicted is non-competitive inhibition (‘i’= ♦- 0 µM, ■- 100 µM, ▲-

200 µM). The experiments were performed in triplicates and the data points show the mean of two experimental values. Error bars represents standard deviation from two different experiments.

The *K*i calculated of all the four drugs are in micromolar concentration. The *K*i determines the extent of the binding affinity of the drug with enzyme. The binding affinity is inversely proportional to the *K*i value. The lesser the *K*i the more effective are inhibitors.The drug can be deleterious if it exceeds the maximal plasma drug concentration a patient can tolerate; an ideal effective *K*i should always be in the concentration lesser or within the typical prescribed dose limits [148].

2.4.7 Cytotoxicity studies

The cytotoxic effects of AP4, EAC, PSU and DRV on Vero cell were determined by MTT assay. Cell viability remained above 80% at 1 mM concentrations which was maximum non-toxic dose (MNTD) value for AP4A, EAC and PSU compounds while for DRV compound, MNTD was 0.1 mM (Fig.2.4.8.1 a). Vero cells treated with 0.1% DMSO did not show any cell cytotoxicity which was the final concentration of DMSO in 0.1 mM DRV (data not shown). To eliminate the possibility of compound mediated cytotoxicity, antiviral activity of AP4A, EAC, PSU and DRV was assessed at concentrations lower than their MNTD values.

2.4.8 Assessment of antiviral activity by plaque assay

The antiviral activity of AP4, EAC, PSU and DRV compounds was determined by the plaque reduction assay. To identify the stage of CHIKV infection affected by each compound, Vero cells were treated with compounds at different time points: pre-treatment is 4 h prior to viral infection, simultaneous-treatment is treatment with viral adsorption and post treatment is addition of compounds after the viral infection. The supernatant was harvested at 24 h post-infection. The percentage of plaque inhibition relative to the control (without addition of compound) was evaluated for each compound concentration at different time of infection. All the compounds showed potent antiviral activity based on the percentage of plaque inhibition results except DRV. DRV compound exhibited ambiguous results in plaque assay (data not shown). Interestingly, reduction in the mean virus titer with various concentrations of AP4, EAC and PSU compounds was observed more in post treatment as compared to pre and simultaneous-treatment (Fig.2.4.8.1 b, c and d). These results indicated that AP4, EAC and PSU compounds significantly suppressed plaque formation during post translation at post-entry stages events which demonstrates the anti-CHIKV activity. The antiviral assay for DRV was not conducted due to the ambiguous results of plaque assay.

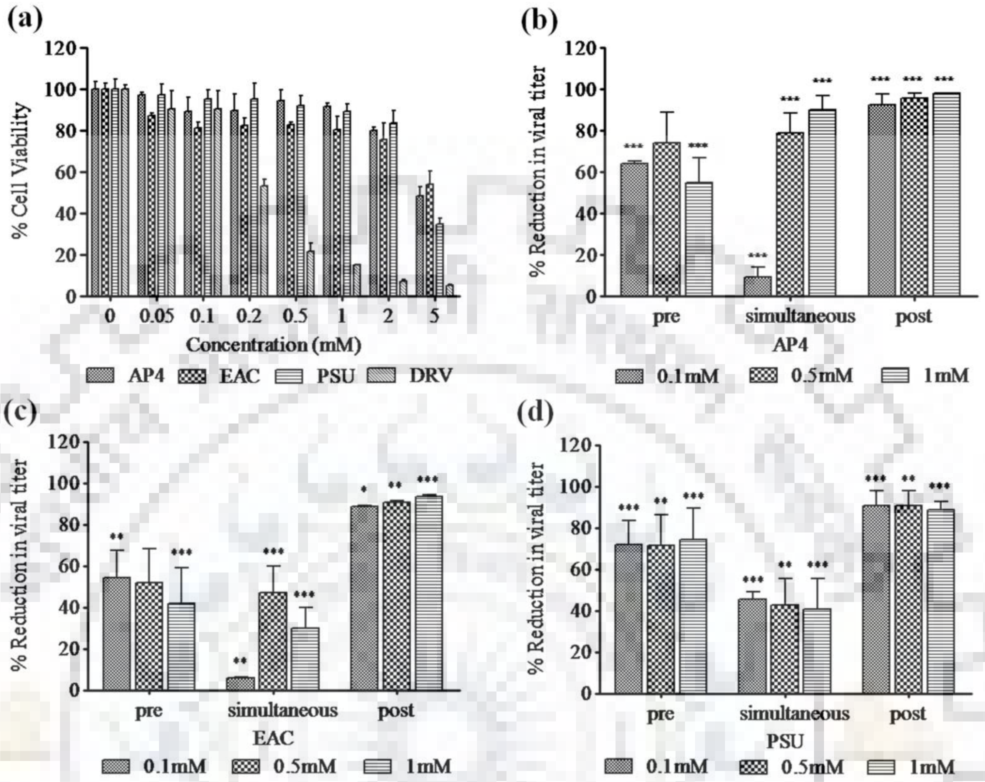


Figure 2.4.8.1Assessment of the compound cytotoxicity by MTT assay (a) and Inhibition of CHIKV with (b) AP4, (c) EAC and (d) PSU compounds in Vero cells by plaque assay. (a) The percentage cell viability was calculated by MTT Assay in Vero cells at varying concentrations (0.05 mM to 5 mM) of each compound (AP4, EAC, PSU and DRV) against the untreated control. 0.1% DMSO was taken as untreated control in DRV compound; Inhibition assay was performed at pre, simultaneous and post treatment with 0.1 mM, 0.5 mM and 1 mM concentration of each compounds. The plaque assay was done and titers of CHIKV were determined in the form of percentage reduction in viral titer with respect to untreated control and graph was plotted using Graph Pad Prism Version 5. Data points represent the mean ± standard deviation of triplicate experiments. The statistical analysis was done by one way – ANOVA and Dunnett’s method. All the data were statistically significant as the p value was less than 0.05 (***;$p \leq 0.001$, **;$p \leq 0.01$, *; $p \leq 0.05$).

2.4.9 Inhibitory potency

The inhibitory potency of AP4, EAC and PSU compounds was evaluated by dose-dependent inhibition study by quantifying infectious viral titer in plaque assay. Dose-response curves were generated by measuring pfu/ml at different range of compound concentrations. Six concentrations (0 µM-100 µM) for each compound were used to generate inhibition curves.The EC$_{50}$of AP4, EAC, and PSU were10.66µM ± 2.25, 4.01µM ± 1.96 and 22.91µM± 3.83, respectively (Fig. 2.4.9.1 a, b and c).

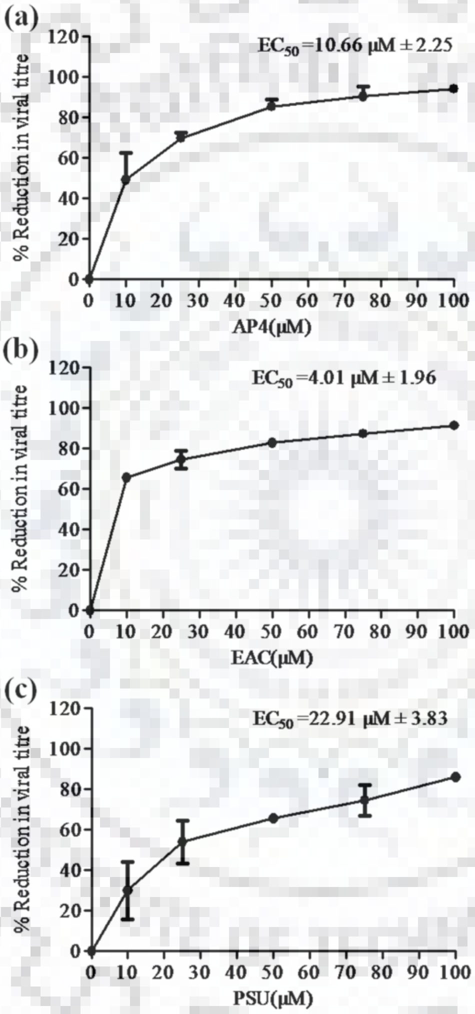


Figure 2.4.9.1 Evaluation of anti-CHIKV activity of (a) AP4, (b) EAC and (c) PSU compound in Dose-dependent manner. CHIKV infected Vero cells were treated with different concentration (0-100 µM) of each compounds in post treatment. Supernatant was harvested after 24 h of infection and plaque assay was done. Values are mean and error bars represent standard deviation from duplicate experiments.

2.4.10 Effects of inhibitors on viral RNA synthesis

In order to verify the intracellular mode of action of compounds that led to reduction in the virus titer in infected cells treated with compounds, the levels of viral RNA in the infected cells treated and untreated with compounds at 12 hpi was analyzed. To detect the levels of viral RNA, a minus-strand specific RT-PCR assay was used. Vero cells were infected with CHIKV at an MOI of 10 at 37 °C with 5% CO2. After incubation of cells with virus for 1.15 h, 100 µM of each compound was in DMEM with 2% FBS was added to respective well (post treatment). At 12 hpi, the infected cells treated with compounds were harvested. RNA was extracted and used in the two-step RT-PCR assay [147]. The levels of viral RNA for infected cells treated with the inhibitors and untreated cells infected with virus as control was examined based on the intensity of the amplified DNA band. The intensity of viral RNA-derived PCR product up to 12 hpi for all the inhibitors was found to be similar to the virus control (without inhibitor) (Fig-2.4.10). These results clearly demonstrate that the compounds have no inhibitory effect on viral RNA synthesis up to 12 hpi in the presence of the inhibitors indicating that inhibitors do not inhibit the enzyme activity of the replication complex. It reveals that these inhibitors have no effect on the early stage of CHIKV life cycle in infected cells.

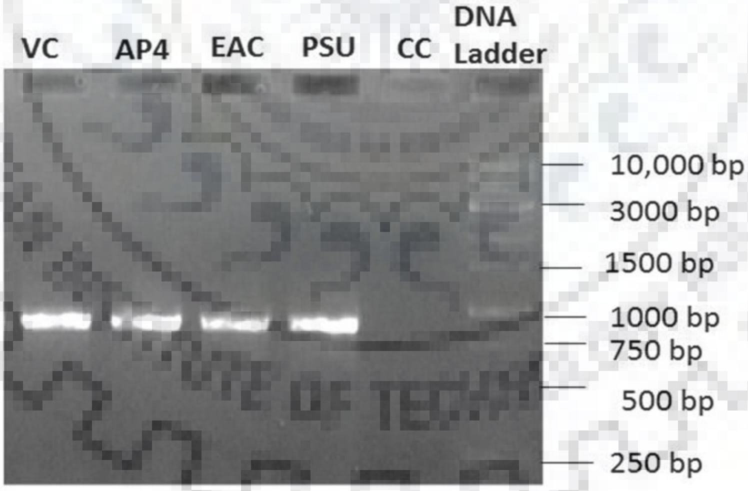
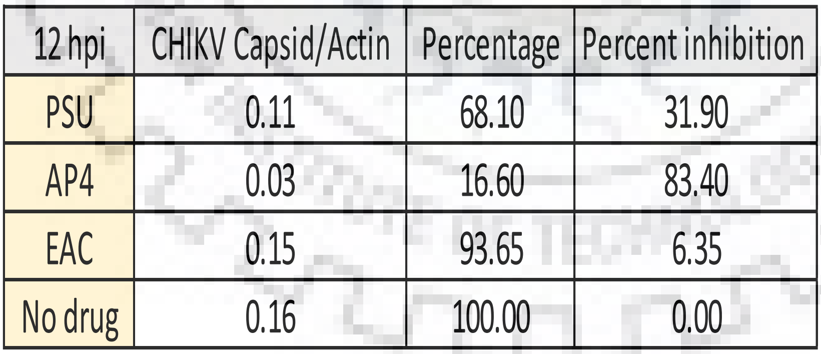


Figure 2.4.10Agarose gel electrophoresis profile of RT-PCRproducts of 12 hpi. Lane 1: VC (Virus Control; infected cells without inhibitor), Lane 2: AP4, Lane3: EAC, Lane 4: PSU, Lane 5: CC (Cell control; uninfected cells) and Lane 7: 1kb DNA ladder. Uniform band of 914bp were observed for all the RT-PCR products including the control which suggests no effect of CHIKV inhibitors on the minus strand RNA synthesis. This reveals that the CHIKV inhibitors have no effect on the CHIKV replication i.e in early stage of infection.

2.4.11 Western immunoblot analysis:

A western blot analysis was performed on CHIKV infected treated and untreated Vero cells to validate that these inhibitors target the serine protease activity of CHIKV capsid. For validating the inhibitory potential of PSU, AP4 and EAC on CHIKV, the levels of CP released as a result of the autoproteolytic serine protease activity of capsid in infected cells was analysed by doing a western blot analysis on lysates of infected Vero cells which were treated with 100 µM of compounds. After 12 hpi, cells infected with CHIKV in presence of the antiviral compounds were lysed using cell lysis buffer (RIPA buffer), samples were separated on a 10% SDS-PAGE and proteins were transferred onto a nitrocellulose membrane. Then these were probed with SINV-specific rabbit polyclonal antibody anti-CP and mouse anti-actin monoclonal antibody (used as an internal control) followed by Infrared-labelled goat anti-rabbit and goat anti-mouse secondary antibodies for the detection of proteins. Finally, the blots were scanned with the Odyssey infrared imager and protein bands were quantified using Odyssey software version 3. The expression of CP after 12 hpi is significantly reduced to approx. 32%, 83% and 6% in case of PSU, AP4 and EAC when compared to no compound treated CHIKV infected cells (Fig. 2.4.11). The level of CP is determined by obtaining the ratio between β-actin and CP level. Here, β-actin acts as internal control for the experiment.



12 hpi	CHIKV Capsid/Actin	Percentage	Percent inhibition
PSU	0.11	68.10	31.90
AP4	0.03	16.60	83.40
EAC	0.15	93.65	6.35
No drug	0.16	100.00	0.00

Figure 2.4.11.1 Western blot band quantification raw data

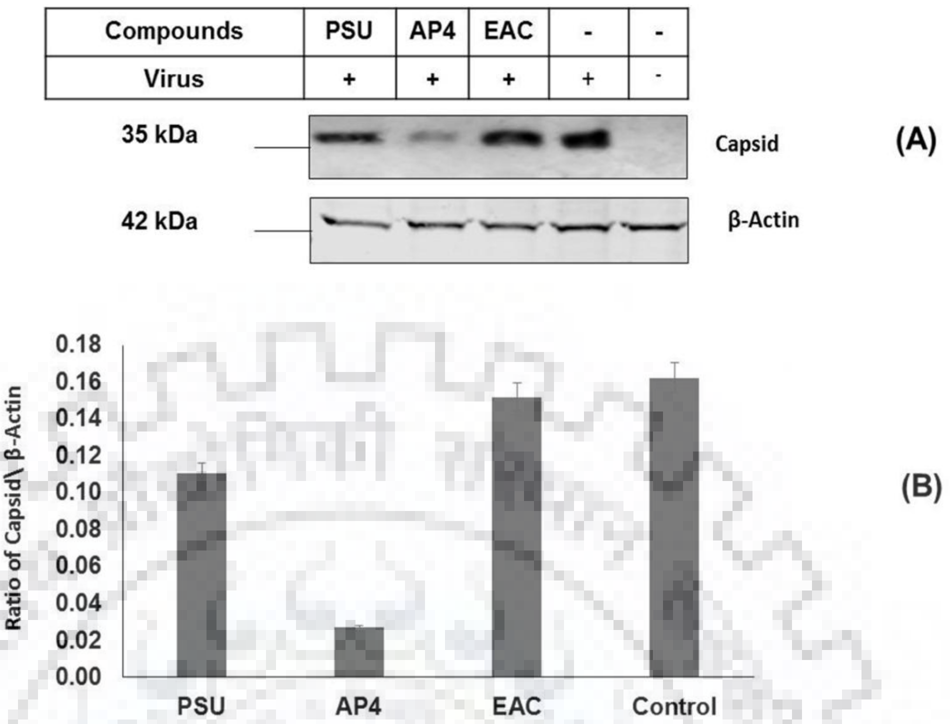


Figure 2.4.11.2. Analysis of CHIKV CP expression in Vero cells. (A) Western Blot analysis. CHIKV CP was probed by SINV-specific rabbit polyclonal antibody anti-CP, detected by infrared-labelled (IRDye 680 and IRDye 800) goat anti-rabbit antibody. Mouse anti-actin monoclonal antibody (1:3000) (Chemicon) was used as an internal control and was detected by goat anti-mouse secondary antibodies (Li-Cor). The expression of CP was reduced in the presence of compounds (AP4, EAC and PSU). (B) Bar graph analysis. The ratio between CP and actin level was used to observe the level of reduction in the expression of CP and was found to be the highest in PSU treated CHIKV infected vero cells followed by AP4 and EAC. The depletion of expression level of CP observed in the presence of compounds confirms that the inhibitors specifically target CP.

2.4.12 Immuno-fluorescence assay

The immuno-fluorescence assay revealed significant reduction in CHIKV load when treated with AP4, EAC and PSU in post-treatment. In this experiment, the alphavirus envelope protein (E2) was detected by immuno-fluorescence assay as an indication of successful CHIKV infection and replication. The cell control (without infection) was included as negative control which ensured the specificity of the primary antibody. The reduction in fluorescence was observed with all the three compoundsat 100µM concentration (Fig. 2.4.12.1 a to e) which confirmed the inhibition of CHIKV replication and translation.

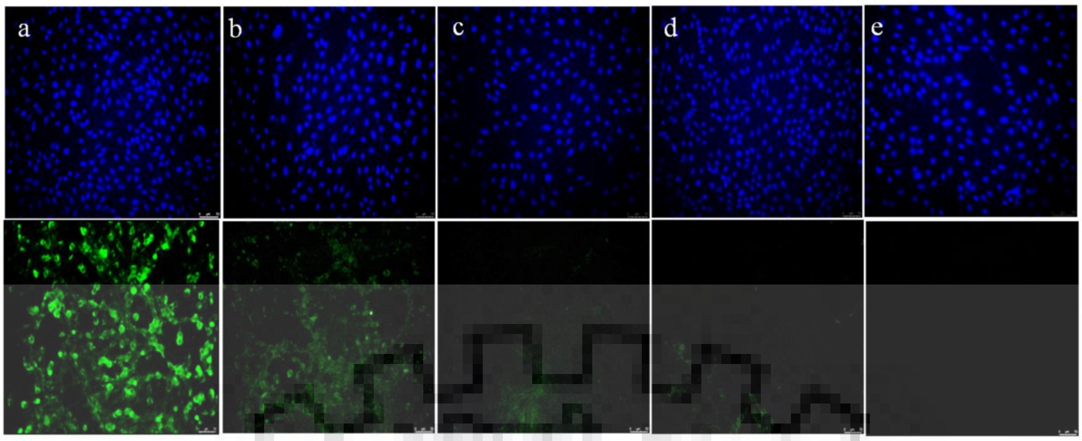


Fig. 2.4.12.1 Effect of AP4, EAC and PSU treatment on CHIKV infection was assessed by Immuno-fluorescence assay. Vero cells were treated with 100µM concentration of each compound in the post treatment assay. At 36 h post-infection, the fixed and stained cells were observed to show the nucleus (blue) and CHIKV E2 protein (green) under the Inverted fluorescence microscope (Lieca) with 20X magnifications. (a)Virus Control (without inhibitor), (b) CHIKV infected Vero cells treated with PSU, (c) CHIKV infected Vero cells treated with AP4, (d) CHIKV infected Vero cells treated with EAC, and (e) Cell control without CHIKV infection.

2.4.13 Quantitative real time (qRT) PCR

To determine the antiviral effect of compoundsin terms of the reduction in CHIKV RNA copy number in the infected cells, qRT-PCR was performed. Cells treated with the compounds showed reduction in viral RNA for all the three compounds. The prominent reduction in viral RNA copy number was observed only in post treatment. The viral RNA was significantly reduced 90.87%, 93.93% and 71.9% in post-treatment of AP4, EAC and PSU, respectively (Fig. 2.4.13.1). The fold reduction in viral RNA amount was observed indicating that AP4, EAC and PSU inhibited the CHIKV replication at post stage.

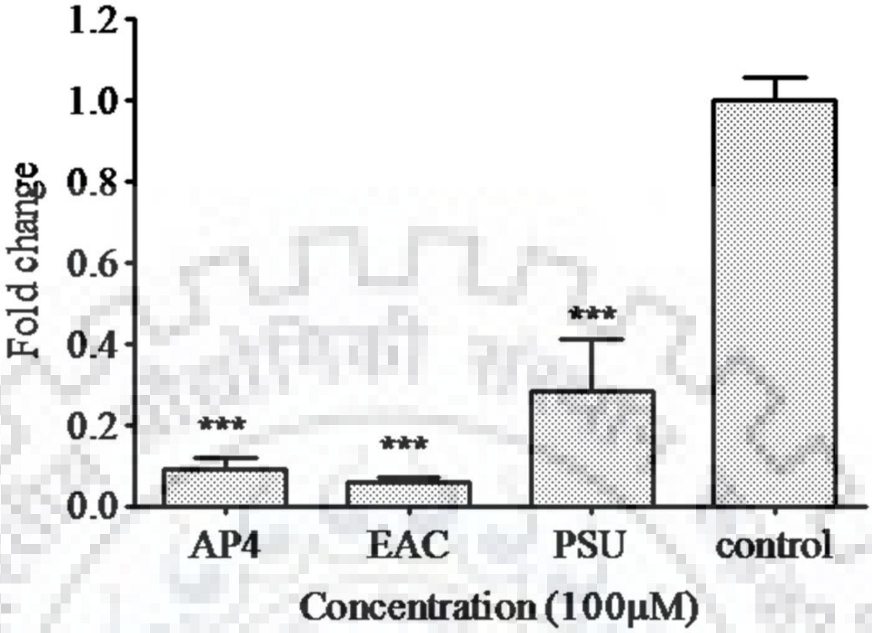


Fig.2.4.13.1Reduction of CHIKV RNA in Vero cells by AP4, PSU and EAC in post treatment assay. The post treatment with 100μM concentration of each compound was done in CHIKV infected cells and viral RNA was quantified 24h post-infection using beta action as endogenous control in triplicate. Quantitative real-time RT- PCR data depicting fold change with respect to control and presented logarithmically. The statistical analysis was done by one way – ANOVA and Dunnett's method. All the data were statistically significant as the p value is less than 0.05 (***; $p \leq 0.001$).

2.5 Conclusion

In the current study, the crystallisation trial of CVCP active was done but diffractable cystals were not obtained. Due to unavailability of crystal structure of CVCP active, the C-terminal W261 residue was removed from the coordinate file of the crystal structure of the CVCP inactive protein (PDB 4AGK) to make the active site accessible for the docking of small molecules against it. Structure-based *in silico* screening of FDA approved drugs and molecular docking of compounds into the active site of CVCP and AVCP led to the identification of four potential inhibitors of AVCP and CVCP which already serve as aminogycoside (PSU), HIV protease inhibitor (DRV), eosinophil derived neurotoxin inhibitor (AP4) and antiplatelet drug (EAC). These four molecules were further *in vitro* tested separately for their antiviral potential against the purified active CP enzyme using peptide (substrate) specific FRET based assay wherein the fluorogenic peptide has scissile Trp- Ser bond which is specifically cleaved by the protease. In the peptide specific FRET based assay, all the four compounds showed their inhibitory effects on the CP proteolytic activity of CVCP active by showing reduced RFU (relative fluorescence unit) per minute when compared with the control (without inhibitor). The double reciprocal plot for all the inhibitors showed different mode of inhibition. For the inhibitor AP4 and DRV, mode of inhibition was competitive which showed unaltered *V*max and reduced *K*m. EAC showed mixed inhibition with the reduced *K*m and increased *V*max , also the α value for this inhibition was 5.34 which proclaims the inhibition to be more of a competitive and less of a noncompetitive. The PSU showed non competitive mode of inhibition as *V*max reduced on increasing the PSU concentration but the *K*m remain unaltered suggesting the possibility of the existence of the allosteric site in the CVCP protease. The *K*i (inhibitory constant) of all the four compounds were in micromolar concentrations.The K_i values calculated for compounds AP4, EAC, DRVand PSU were 133.6±27.45 µM, 135±76.25 µM, 147.4 ±10.93 µM and 353.0±17.14 µM respectively. The lesser the *K*i of a compound, the more potential it attains to claim as a drug. These inhibitors has the potential to classify as antiviral drug if their *K*i does not exceed the maximal plasma drug concentration a patient can withstand in the typical dosing. However, if this fails still improvements can be done to label them as anti CHIKV drugs by designing their structural analogues and testing their *K*i. Although their *K*i is within the permissible range as they are lower than the MNTD of the drugs determined by MTT based assay except for DRV. The *K*i of DRV (147.4±10.93 µM) >MNTD of DRV (100 µM). The antiviral activity was further validated by cell-based antiviral assays in Vero cells which demonstrated the reduction in CHIKV infection in the form of viral load and viral RNA titer.Prior to the cell based antiviral assay, MTT based assay was done for each of the compounds to determine their non cytotoxic dose which was upto 1 mM for AP4, EAC and PSU and 100 µM for DRV. The plaque reduction assay showed reduction in the viral titer (Post treatment with compound for 24h after 1.5 h of CHIKV infection) for all the compounds except for DRV.AP4, EAC and PSU showed upto 90 % inhibition at 100 µM concentrations.The EC_{50} for each compound were calculated based on plotting graph between percentage inhibitions vs. compound concentration. The maximum EC_{50} value among the approved antiviral drugs (from the past 50 years) is 9µM [146]. Thus EAC can be an effective potential anti CHIKV drug with the EC_{50} value 4.0 µM. AP4 can also be a good anti-CHIKV drug with EC_{50} value 10 µM. Designing structural analogues of AP4 and testing their efficacy against CHIKV can be promising. To detect the effect of these inhibitors on viral RNA synthesis, a minus-strand specific RT-PCR assay was conducted for compound (AP4/ EAC/ PSU) treated infected cells and it was found the level of RT-PCR products including the control were same by observing uniform band of 914 bp on EtBr stained agarose gel. This suggests, CHIKV inhibitor has no effect on the initial viral RNA synthesis i.e in the early stage of infection. To establish the fact these compounds specifically target CP, western-immunoblot analysis was done where the reduction in the release of CP was observed for AP4, EAC and PSU when compared with control (CHIK infected Vero cells without any treatment with the compound) by using polyclonal anti-CP antibody.The inhibitors tend to target CHIKV at the post entry stage. This was further validated by conducting Immunofluoresnce assay (IFA) where significant reduction in CHIKV load was observed by AP4, EAC and PSU in post treatment. E2 glycoprotein was targeted by using anti-alphavirus (E2) mouse monoclonal antibody. The reduction in CHIKV E2 was seen in the form of depleted FITC intensity when probed by anti-mouse IgG fluorescein isothiocyanate (FITC) antibody. The control had Vero cells with CHIKV infection. This suggests the compounds targetted CP which hindered polyprotein processing. The q-RT PCR showed significant reduction in viral RNA copy number in post treatment only for the compounds AP4, EAC and PSU. The constant reduction in CHIKV infection and viral RNA load in the post treatment only, suggests the compounds actively target the stage of virus release not at the stage of replication. All these experiments sum up to suggest that the compounds AP4, EAC and PSU are potential CVCP protease inhibitors. These reported compounds against CHIKV can pave their way in clinical trials by conducting *in vivo* experimental studies in the future aspect and can be futher refined to classify as effective anti-CHIKV drugs.

भारतीय प्रौद्योगिकी संस्थान रुड़की
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

CHAPTER 3

Structural and functional perspective of the N-terminal CP domain of AVCP and CVCP

3.1. Chapter summary

The chapter 3 strives to study the structural and functional significance of N-terminal domain of alphavirus CP. The N-terminal domain of alphavirus CP is disordered and till now no crystal structures of it has been attained. The N-terminal domain has positively charged residues in abundance. The absence of structure of N-terminal domain is suggestive of the fact, the domain intrudes inside the NC and electrostatically interacts with the viral genomic RNA [51], [150], [151]. The CP of SINV is categorized into three major regions (Region I, Region II and Region III). Region I is comprised of positively charged amino acid residues which binds with the genomic RNA non-specifically through charge neutralization [56]. Region II is comprised of amino acid residues from ~ 81 to 113 which bind to the genomic RNA where encapsidation signal is present [69]. Region III constitutes the C-terminal domain of CP which is a protease domain. Region III is involved in the generation of capsomeres. The N-terminal domain of CP of VEEV is shown to play an essential requirement in the formation of virus- like particles (VLPs). The structure of a VLP resembles with the structure of a wild-type virion [6]. For the generation of an NC core structure with a defined icosahedral symmetry, interaction of 240 copies of CP with the genomic RNA is required. The crystal structure of SINV CP shows the N-terminal domain highly disordered which makes it difficult to unravel CP- RNA interaction at the atomic level and understand the process of NC assembly [107], [152].

In this chapter, attempts have been made to unravel the three dimensional structure of N-terminal domain of alphavirus CP for which cloningand expression of AVCP19 (19-265 residues) and CVCP20 (20-259 residues) were done. The purification of CVCP20 was not successful therefore n crystallization trials were done for it. Co- crystallisation trial of AVCP 19 was done with the 12 mer DNA. Another construct of AURAV namely AVCP 80 (80-267 residues) was used for the co-crystallisation trial with the 12-mer DNA. In the literature reviewedof SINV CP, region comprising of 81-113 residues is shown to specifically bind the encapsidation signal present on the viral genomic RNA [69]. To study the CP–DNA (48- mers) interactions in the AVCP19, size exclusion chromatography and DSC techniques were used. This study is the first report of the alphavirus CP- DNA interaction through Size exclusion chromatography and DSC techniques. The terminal residue W261 was deliberately removed while making AVCP19 and CVCP20 constructs to study the *trans*proteolytic activity of DNA bound CP and determine the effect of nucleic acid binding on the protease activity of CP.

3.2. Introduction

The worldwide distribution of alphavirus is a grim matter of concern especially in poverty dwelling areas where the measures to deal with it become limited. The alphavirus infection is majorly transmitted by mosquitoes while their perpetuation in poverty stricken areas remains unchecked because of limitation in basic sanitation amenities caused by economy crunch.The alphavirus infection accompanies infectious arthritis, encephalitis, rashes and fever. Alphavirus is categorized in the group IV belonging to Togaviridae family. Alphaviruses exhibit icosahedral symmetry and are enveloped membrane bound (host cell derived). The viral envelope proteins E1 and E2 interact to form heterodimer; these heterodimers assemble to for, trimmers which give rise to spike protrusions. The cytoplasmic domain of heterodimer E1-E2 (transmembrane proteins) interacts with the nucleocapsid (NC) which makes the virus bud out of the cell. The genome of alphavirus constitutes 49S genomic RNA which is a 12 kb approx. long RNA bearing positive-sense. Apart from viral genomic RNA, there is also 26S subgenomic RNA which encodes the structural proteins. The genomic RNA constitutes two open reading frames (ORFs); the N-terminal ORF encodes the nonstructural polyprotein whereas the C-terminal ORF encodes the structural polyprotein. CP is a structural protein which has serine proteinase present at the C-terminus. The protease cleaves the CP from the precursor structural polyprotein precursor via *cis*-autoproteolysis. The C-terminal Trp residue of CP turns off the proteolytic activity of CP protease by binding to the active site of the protease. Post the cis-cleavage of CP from the structural polyprotein, the N-terminal of CP interacts with the viral genomic RNA which is followed by the RNA encapsidation and generation of the NC assembly. A typical alphavirus particle contains NC core encircled by the host cell derived phospholipid bilayer insidewhich trimeric spikes are embedded.The CP present inside the NC core binds with the cytoplasmic tail of E2-E1 which triggers the budding of viruses.The NC assembly of alphavirus is accompanied by many steps. The primary step in the assembly is the binding of nucleic acid to the dimer of CP. The region spanning 1 to 80 amino acid residues in the N-terminal domain of CP non-specifically binds to the viral genomic RNA [53]. The N-terminal domain of CP which constitutes 81 to 113 amino acid residues specifically binds to the encapsidation signal present on the viral genomic RNA [53], [57], [153]. Binding of CP to RNA is lost on mutating the 76-107 amino acid residues of N-terminal of CP [154], [155]. In SINV, the N-terminal of CP exhibits α-helix (Helix I) (38 to 55amino acid residues) [151]. The Helix I constitute the leucine zipper comprising of three conserved Leu residues. Partially truncating the Helix I or by substituting the conserved Leu residues results in reduced viral replication. The Helix I is speculated to participate in inter-capsomeric interaction and stabilize the assembly intermediates [107]. *In vitro* studies has predicted the involvement of amino acid residue 99, 103 and 105 in the capsid dimer formation via coiled-coil interlink of Helix I from each CP monomer [53], [69]. Additionally, in the premature NC assembly, the coiled-coil interaction is predicted to regulate the assembly [107].

3.3 Materials and methods

3.3.1 Cloning of CVCP20 (20-259 residues)

The gene encoding CVCP20 (amino acid residues 20-259) was amplified using the plasmid containing the structural polyprotein of chikungunya (procured from Dr. Shailly Tomar Lab) as template by polymerase chain reaction using Taq polymerase. The restriction endonuclease site for the gene was *Nde*I and *Xho*I.The primers for PCR amplification of CVCP20 was designed manually and analyzed by Oligoanalyzer tool of IDT technologies. The forward and reverse primers designed for CVCP 20 are:

Forward: 5'–CAAGCATATGACTCCGCGCTCTACTATCCAAATCATCAGG–3'

Reverse : **5' –CTAGAATCTCGAGCTATTCGGCCCCCTCGGG–3'**respectively.

The expression vector used in the molecular cloning was plasmid pET28c having cleavage site for the TEV- protease comprising of a 6XHis tag at the N-terminus. The plasmid was isolated from the 10 ml culture of DH5α cells using the plasmid miniprep kit of Qiagen. The PCR amplification of the target gene was done in a thermal cycler (Biorad) at the optimum conditions. The reaction was set in duplicate 50 µl each. The 100 µl PCR product was purified using PCR purification kit of Qiagen. The PCR purified product was eluted in 40 µl nuclease free water. The PCR purified product and the expression vector was digested by *Nde*I and *Xho*I to generate sticky ends. The digested products were run on 0.8% low melting agarose gel and bands comprising the digested gene and vector were excised. The excised products comprising DNA were extracted from the gel using the gel extraction kit of Qiagen. The extracted DNA fragments were then ligated at 16 ºC for 16 h using T4 DNA ligase. The ligated product was introduced into the chemically competent DH5α cells on a Luria Bertani (LB) agar plate supplemented by 50 µg ml^{-1} kanamycin. The plate was incubated at 37 ºC for 14 h. After 14h, the plate was observed with the lawn of uniform colonies. A distinct colony was picked from the grown lawn and was inoculated in 10 ml LB supplemented with 50 µg/ml kanamycin overnight at 37 ºC for plasmid isolation. The plasmid isolated was screened for the clones by PCR amplification and restriction digestion. The confirmation of gene insert in the pET-28c was also done by sequencing the plasmid.

The isolated plasmid having correct insert was introduced in the chemically competent *Rosetta* cells for the expression of the gene and the culture was spread on LB agar plate supplement with 50 µg ml^{-1} kanamycin and 35 µg ml^{-1} chloramphenicol. The plate was incubated at 37 ºC for 14h. After 14 h, plate had lawn of uniform colonies. From the lawn, one colony was picked and inoculated in 10 ml LB supplemented with 50 µg ml^{-1} kanamycin and 35 µg ml^{-1} chloramphenicol. The primary culture was grown at 37ºC for 14h at constant agitation of 200 rpm. After 14 h, secondary culture of 500ml each was grown by taking inoculum of 200µl from the primary culture supplemented with 50 µg ml^{-1} kanamycin and 35 µg ml^{-1} chloramphenicol. Secondary culture was grown at 37 ºC at constant agitation of 200 rpm until its optical density reached 0.7 at 600 nm. At this point the culture was induced by 0.4 mM isopropyl ß-D-1 thiogalactopyranoside (IPTG) and one culture of 500 ml was grown at 37 ºC for 5 h at 200 rpm while the other secondary culture of 500 ml was grown for 14 h at 18 ºC at 200 rpm. The grown culture was subjected to centrifugation at 7000 rpm for 10 min at 4 ºC. The cells were pelleted down and the expression of CVCP 20 was analyzed on 12 % SDS PAGE.

3.3.2 Expression of CVCP 20

The pelleted cells were resuspended in chilled ice cold 20mM potassium phosphate buffer having pH 7.6. The resuspended cells were then subjected to the French press (Constant Systems Ltd, Daventry UK) in which cells were broken down at high pressure 2700 bar. The broken cells were then centrifuged at 12,000 rpm for 90 min at 4 ºC to segregate the soluble protein from the cell debris. The supernatant, pellet and protein ladder were run on 15 % SDS PAGE to check the solubility and expression of CVCP20. The gels were stained with Coomassie Brilliant Blue R-250 dye followed by destaining.

3.3.3 Purification of CVCP 20

The CVCP20 was purified from the cell pellet of 1L batch of secondary culture .The pellets were resuspended in chilled (4 ºC) 20 mM phosphate buffer having pH 7.6. Post resuspension of pellets, high pressure of 40 KPSI (2700 bar) was given to the cells by the French press. After the disruption of cell walls due to high pressure, the soluble part was segregated from the cell debris by centrifuging the lysate at 12000 rpm at 4ºC for 90 min. The supernatant having soluble CVCP 20 was mixed with Ni-NTA (Nickel- Nitrilotriacetic acid) agarose beads which were pre- equilibrated with 20mM phosphaphate buffer (pH 7.6).

The incubation of supernatant with the Ni-NTA beads was done for 30mins. The soluble CVCP20 was purified using IMAC. Initial two consecutive wash were given with KCl (Wash1 had 20 mM phosphate buffer; pH 7.6 and 500 mM KCl; Wash 2 had 20 mM phosphate buffer; pH 7.6 and 1M KCl). Wash 3 to 5 comprised of 20 mM phosphate buffer, pH 7.6 with the ascending concentration of imidazole (20 mM to 100 mM). The final elution buffer for the His-tagged CVCP 20 had 250 mM imidazole in 20 mM phosphate buffer (pH 7.6). Same procedure was used for the purification of CVCP20 using another buffer i.e.50 mM Tris and 100 Mm NaCl; pH 7.6.

3.3.4 Cloning of AVCP19 (19-265)

The gene encoding AVCP 19 (amino acid residues 19-265) was amplified using the plasmid containing the full length CP protein previously cloned by a graduate student in the lab as template in the polymerase chain reaction. The restriction endonuclease sitesused for cloning the gene were*Nde*I and *Xho*I.

The primers for PCR amplification of AVCP 19 were manually designed and analyzed by Oligoanalyzer tool of IDT technologies. The forward and reverse primers designed were

5'- CTG GAA TTC ATA TGA TAG CAT GGA GGC CAA GAC G -3' and**5'- CTA GAA TCT CGA GCT ATA CAG TAT CTT CGT GGG TGG -3'** respectively.

The expression vector used in the molecular cloning was plasmid pET-28c having cleavage site of TEV- protease for removing the 6x-His tag residues from the N-terminus of recombinant protein. The template plasmid was isolated from the 10 ml culture of DH5α cells using miniprep kit of Qiagen. The PCR amplification of the target gene was done in a thermal cycler (Biorad) at the optimum conditions. The reaction was set in duplicate 50 µl each. The 100 µl PCR product was purified from the PCR purification kit from Qiagen. The PCR purified product was eluted in 40 µl nuclease free water. The PCR purified product and the expression vector was digested by *Nde*I and *Xho*I to generate sticky ends. The digested products were run on 0.8% low melting agarose gel and bands comprising the digested gene and vector were excised. The excised products comprising DNA were extracted from the gel using the extraction kit of Qiagen. The extracted DNA fragments were then ligated at 16 ºC for 16 h using T4 DNA ligase. The ligated product was introduced into the chemically competent DH5α cells on a Luria Bertani (LB) agar plate supplemented with 50 µg ml^{-1} kanamycin (for selecting transformed colonies). The plate was incubated at 37 ºC for 14 h. After 14 h, the plate was observed with the lawn of uniform colonies. A distinct colony was picked from the grown lawn and was inoculated in 10 ml LB supplemented with 50 µg ml^{-1} kanamycin overnight at 37 ºC for plasmid isolation. The plasmid isolated was screened for the clones by PCR amplification and restriction digestion. The insert gene fragment in the pET-28c was further confirmed by DNA sequencing.

The isolated plasmid having positive clones was introduced in the chemically competent *E. coli* strain*Rosetta* cells for the expression of the gene and the culture was spread on LB agar plate supplement with 50 µg ml^{-1} kanamycin and 35µgml^{-1} chloramphenicol. The plate was incubated at 37 ºC for 14 h. After 14 h, plate had lawn of uniform colonies. From the lawn, one colony was picked and inoculated in 10 ml LB supplemented with kanamycin 50 µg ml^{-1} and 35 µg ml^{-1} chloramphenicol. The primary culture was grown at 37 ºC for 14 h at constant agitation of 200 rpm. After 14 h, secondary culture of 500 ml each was grown by taking inoculum of 200 µl from the primary culture supplemented with 50 µg/ml kanamycin and 35 µg/ml chloramphenicol. Secondary culture was grown at 37ºC at constant agitation of 200 rpm until its optical density reached 0.7 at 600nm. At this point the culture was induced by 0.4 mM isopropyl ß-D-1 thiogalactopyranoside (IPTG) and one culture of 500ml was grown at 37 ºC for 5 h at 200 rpm while the other secondary culture of 500 ml was grown for 14h at 18 ºC at 200rpm. The grown culture was subjected to centrifugation at 7000 rpm for 10 min at 4 ºC. The cells were pelleted down and the expression of AVCP19 was analyzed on 12 % SDS PAGE.

3.3.5 Expression and solubility of AVCP19

The pelleted cells were resuspended in ice cold 20 mM potassium phosphate buffer having pH 7.6. The resuspended cells were then subjected to the French press (Constant Systems Ltd, Daventry UK) in which cells were broken down at high pressure 2700 bar. The broken cells were then centrifuged at 12,000rpm for 90 min at 4ºC to segregate the soluble protein from the cell debris. The supernatant, pellet and protein ladder were run on 15 % SDS PAGE to check the solubility and expression of AVCP19.The gels were stained with Coomassie Brilliant Blue R-250 dye followed by destaining.

3.3.6 Purification of AVCP19

The AVCP19 was purified from the cell pellet of 1L batch of secondary culture .The pellets were resuspended in chilled (4 ºC) 20 mM phosphate buffer having pH 7.6. Post resuspension of pellets, high pressure of 40 KPSI (2700 bar) was given to the cells by the French press. After the disruption of cell walls due to high pressure, the soluble part was segregated from the cell debris by centrifuging the lysate at 12000 rpm at 4 ºC for 90 min. The supernatant comprising the predicted soluble AVCP19 was mixed with Ni-NTA (Nickel- Nitrilotriacetic acid) agarose beads which were pre- equilibrated with 20 mM phosphaphate buffer (pH 7.6).

The incubation of supernatant with the Ni-NTA beads was for 30mins. The soluble AVCP 19 was purified using IMAC (Immobilized metal ion affinity chromatography). Initial two consecutive wash were given with KCl (Wash 1 had 20 mM phosphate buffer; pH 7.6 and 500 mM KCl; Wash 2 had 20 mM phosphate buffer; pH 7.6 and 1M KCl). Wash 3 to 5 comprised of 20 mM phosphate buffer, pH 7.6 with the ascending concentration of imidazole (20 mM to 100 mM). The final elution buffer for the His-tagged AVCP19 had 250 mM imidazole in 20 mM phosphate buffer (pH 7.6). The eluted fractions were pooled with TEV protease in 1:25 ratio for the cleavage of His-tag in the dialysis buffer at 4 ºC for 12 h. Dialysis membrane had 10 kDa molecular weight cutoffs (Thermo Scientific). Reverse Ni-NTA was done to remove the cleaved His-tag (affinity –tag) and His –tagged TEV protease. All the fractions of AVCP19 were run on 15% SDS PAGE. The fraction containing AVCP19 was concentrated in the Amicon® Ultra-15 Centrifugal Filter Unit having 3 kDa cutoff. The concentration of purified protein was measured via UV absorbance spectroscopy. The molar extinction of AVCP19 was calculated online from Protparam tool (https://web.expasy.org/protparam/). The molar extinction coefficient of AVCP19 is 22460

3.3.7 Co - crystallisatisation trial of AVCP 19 with the 12- mer DNA

The purified AVCP 19 was concentrated up to 7 mg/ml at 4°C usingAmicon® Ultra-15 Centrifugal Filter Unit having 3 kDa cut off at 4000 rpm. Equimolar concentration (0.4 mM) of AVCP 80 and 12- mer DNA (5'-CCGTTAATGCAT-3') was incubated overnight at room temperature. Co crystallization trial of AVCP 19 and DNA was done using varied concentration of PEG 3350 and Magnesium Chloride. The drop size was 3ul (2µl sample: 1µl reservoir buffer) against 100 µl reservoir buffer. Sitting drop method was used for the crystallisation by exploiting vapour diffusion phenomenon. The trays were kept unperturbed for a month at 20 ºC in the vibration free chamber.

3.3.8 Purification of AVCP80 (capsid protein; residues 80-267)

The AVCP80 cloned in pET-28c vector preceded by 6 His-tag in the N-terminal, expressed in the *Rosetta* cells, purified from the cell pellet of 1L batch of secondary culture. The pellets were resuspended in chilled (4ºC) 50 ml purification buffer which comprised of 50mM Tris buffer and 100 mM NaCl having pH 7.6. Post re-suspension of pellets, high pressure of 40 KPSI (2700 bar) was given to the cells by the French press. After the disruption of cell walls due to high pressure, the soluble part was segregated from the cell debris by centrifuging the lysate at 12000 rpm at 4ºC for 90 min. The supernatant comprising the predicted soluble AVCP80 was mixed with Ni-NTA (Nickel- Nitrilotriacetic acid) agarose beads which were pre- equilibrated with 50ml purification buffer.

The incubation of supernatant with the Ni-NTA beads was for 30 min. The soluble AVCP 80 was purified using IMAC (Immobilized metal ion affinity chromatography). Initial two consecutive wash were given with NaCl (Wash 1 had 500 mM KCl and 20 mM imidazole in 50 ml purification buffer ; Wash 2 had 500 mM NaCl and 50 mM imidazole in 25 ml purification buffer). The final elution buffer for the His-tagged CVCP 20 had 250 mM imidazole in 30ml purification buffer which were collected in 6 equal fractions (pH 7.6). The eluted fractions were pooled with TEV protease in 1:25 ratio for the cleavage of His-tag in the dialysis buffer at 4 ºC for 12h. Dialysis membrane had molecular weight cut off of 10 kDa (Thermo Scientific). Reverse Ni-NTA was done to remove the cleaved His-tag (affinity –tag) and His –tagged TEV protease. All the fractions of AVCP80 were run on 15% SDS PAGE. The fraction containing AVCP 80 was concentrated in the Amicon® Ultra-15 Centrifugal Filter Unit having 3kDa cutoff. The concentration of purified protein was measured via UV absorbance spectroscopy. The molar extinction of AVCP80 was calculated online from Protparam tool (https://web.expasy.org/protparam/). The molar extinction coefficient of AVCP 80 is 22460.

3.3.9 Co crystallization trial of AVCP80

The purified AVCP 80 was concentrated up to 7mg/ml at 4°C usingAmicon® Ultra-15 Centrifugal Filter Unit having 3kDa cut off at 4000rpm. Equimolar concentration (0.4 mM) of AVCP 80 and 12 mer DNA was incubated overnight at room temperature. Co crystallization trial of AVCP80 and DNA was done using varied concentration of PEG 3350 and Magnesium Chloride. The drop size was 3ul (2µl sample: 1µl reservoir buffer) against 100 µl reservoir buffer. Sitting drop method was used for the crystallisation by exploiting vapour diffusion phenomenon. The trays were kept unperturbed for a month at 20 ºC in the vibration free chamber.

3.3.10 Size Exclusion Chromatography to study AVCP19-DNA interaction

AVCP 19 of concentration 3 mg/ml was incubated with 3 mg/ml single stranded DNA of 48 bases (purchased from IDT technologies, India) overnight at room temperature in 20 mM phosphate buffer; pH 7.6, 50 mM KCl and 5% glycerol. The protein-DNA mix was subjected to size exclusion chromatography using HiLoad 16/ 60 prep grade Superdex 75 column (GE Healthcare) to detect the AVCP19-DNA interactions. The control in this experiment was 3mg/ml AVCP 19 without DNA. The sequence of 48 mer single stranded DNA, purchased from IDT technologies (T_m of DNA is 68 ºC) used in the experiment was:

5' –CCGTTAATGCATGTCGAGATATATAAAGCATAAGGGACATGCATTAACGG-3'

The sequence of the 48 mer DNA was in reference to the study of SINV CP that showed *in vitro* assembly of core like particles by cross linking the inter capsomeres dimer[161].

3.3.11Differential Scanning Calorimetry to study AVCP19 and DNA interaction

To study the effect of DNA binding to the N-terminal of CP of alphavirus, differential scanning calorimetry experiments were done and the change in melting temperature Tm was calculated. In the DSC experiment, the buffer used was composed of 20 mM phosphate buffer (pH 7.6), 50 mM KCl and 5% glycerol. The buffer and the samples were degassed prior to the DSC experiment. For conducting the DSC experiment, prior to the addition of sample, 3 buffer scans were done at the scan rate 90º C/h. The DNA and AVCP 19 were incubated overnight at room temperature.Table 3.1 details about the DSC parameters taken for conducting the experiment.

Table 3.1DSC parameters used in conducting the experiment to detect the stability of AVCP 19

No. of scans	10
Post cycle thermostat	1°C
Starting temperature	10°C
Final temperature	90°C
Scan rate	90 °C/h
Prescan thermostat	10 min
Post scan thermostat	2.0 min

Table 3.2Number of samples and their concentrations used in the DSC experiment

Serial number	Samples	Concentration
Sample 1	AVCP19 (Control)	14 µM
Sample 2	Equimolar concentration of AVCP19 and 48 mer DNA	14 µM

3.3.12 AVCP19 *trans*-protease activity

The *trans*- activity of AVCP19 was detected by employing substrate specific FRET based assayas described by Aggarwal et al., 2015 [5]. Brieflly, the sequence of the substrate-peptide used wasDABCYL-Gly-Ala-Glu-Glu-Trp↓Ser-Leu-Ala-Ile-Glu-EDANS (Biolinkk New Delhi, India) having the scissile bond between Trp and Ser which is highly conserved among alphaviruses. The FRET pair used is EDANS (fluorophore) and DABCYL (Quencher) located at C and N terminal respectively. EDANS, has excitation maxima at 340 nm and emission maxima at 490 nm, while DABCYL absorbs in the range of 470 -520 nm. The proteolytic assay was conducted in a 96- well black round bottom non polystyrene plates at 25 ºC in 100µL reaction mixture containing 20mM HEPES pH7.0 and 2µM purified protein with a fixed concentration of substrate (0.6 µM) and fluorescence was analysed at excitation and emission wavelength of 340 nm and 490 nm respectively using a multimode plate reader Cytation 3 (BioTek Instruments, Inc.). The negative control was reaction without the enzyme and all the reactions were conducted in triplicate. The enzyme kinetics for AVCP 19 was determined using a varied range of substrate concentrations (0.6 µM, 1 µM, 4 µM, 8 µM and 16 µM) and a fixed protein concentration 2 µM. The kinetic parameters were obtained by Michaelis-Menten and Lineweaver burk plot. Graph-pad prism software was used to calculate the Vmax, Km and catalytic efficiency.

3.3.13 AVCP 19- 48 mer DNA complex *trans*-protease activity

The *trans*- protease activity of the complex AVCP 19-48 mer DNA (5'–CCGTTAATGCATGTCGAGATATATAAAGCATAAGGGACATGCATTAACGG-3') was determined in the similar fashion as described in 3.3.12. The complex had equimolar concentration of AVCP 19 with 48 mer DNA (2 µM).

3.4 Result

3.4.1 Cloning of CVCP 20 (residues 20 to 259)

CVCP20 (residues 20-259) of 738 base pairs was cloned in the pET 28c vector. The PCR amplification of the positive clone was seen around 738 bp when run on 0.8 % agarose gel. This was further confirmed on digesting the clone by *Nde*I and *Xho*I which showed the insert of 738bp when run on 0.8% agarose gel. The figure 3.4.1.1 (a –b) shows the amplified gene and insert (*Nde*I and *Xho*I digested) of 738 bp on 0.8 % agarose gel. The cloning of CVCP 20 was further confirmed by the Sanger's sequencing.

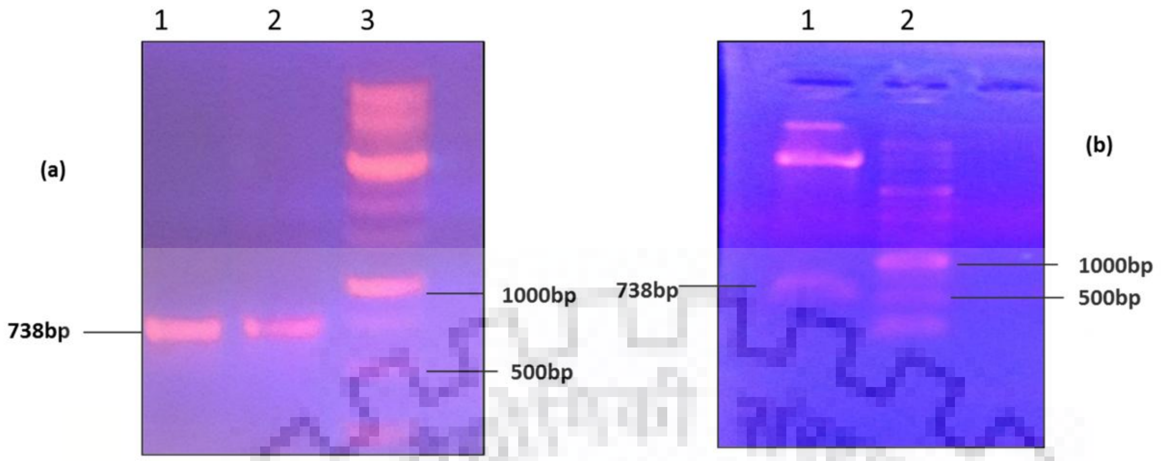


Figure 3.4.1.1: Cloning confirmation of CVCP20 by PCR amplification and RE digestion by *Nde*I & XhoI PCR. **(a)** amplification of the plasmid from the screened clone shows amplified band of 738 bp which corresponds to the size of CVCP 20. **(b)** RE digestion gave insert of 738 bp which confirmed the cloning of CVCP20

3.4.2 Expression and purification of CVCP20

Rosetta cells were transformed by the cloned plasmid for the expression of CVCP20. The transformed *cells* were checked for expression by growing the cells at two different temperatures (37 ºC/ 4h and 30 ºC/ 5 h) after their log phase. The expression profile of both the culture grown at 2 different temperatures was checked on 15% SDS PAGE. CVCP20 was expressed and was soluble at 30 ºC. CVCP 20 was purified from 1L culture (from the pellet of expressed CVCP20 grown in *Rosetta* cells) by IMAC technique. Figure 3.4.2.1 (A, B) shows the SDS PAGE (15 %) profile of the expression and purification of CVCP 20.

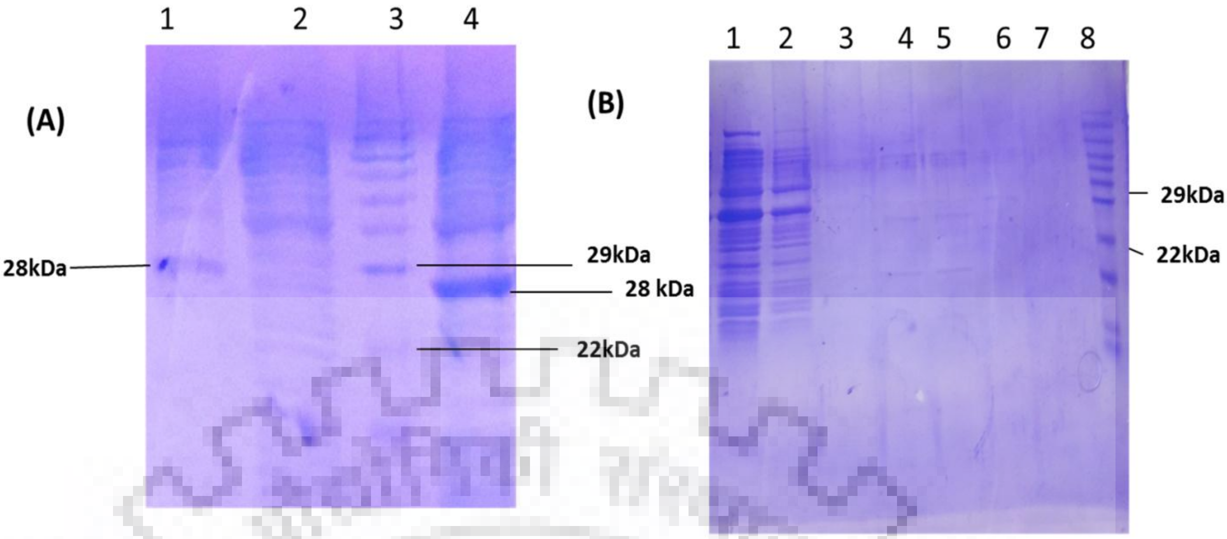


Figure 3.4.2.1 SDS PAGE (15%) profile of expression and purification of CVCP20. (A) lane 1- induced pellet (30 ºC) shows the band of CVCP20 at 28 kDa ; lane 2- uninduced supernatant (30 ºC); lane 4 – induced supernatant (30 ºC) shows the optimum expression of soluble CVCP20 ; Lane 3 – prestained protein ladder (Thermoscientific).**(B)**Purification attempt for CVCP20: Lane1- supernatant of CVCP20 ; lane2 – flowthrough; lane3-5 (wash fractions); lane 6-7 elution fractions ; lane 8- prestained protein ladder (Thermoscientific)

CVCP 20 could not be purified by IMAC as the solubility of was not very good and it seems that the affinity-tag failed to bind the Ni-NTA agarose beads as a result of which the CVCP20 came in flowthrough along with the impurities. This might be due to the instability of the N-terminal domain which is highly disordered and is comprising of positive charged amino acid residues

3.4.3 Cloning of AVCP19

AVCP19 (residues 19-265) of 738 base pairs was cloned in the Pet 28c vector. The PCR amplification of the positive clone was seen around 738bp when run on 0.8% agarose gel. This was further confirmed on digesting the clone by *Nde*I and *Xho*I which showed the insert of 738bp when run on 0.8% agarose gel. The figure 3.4.3.1 (A –B) shows the amplified gene and insert (*Nde*I and *Xho*I digested) of 738bp on 0.8% agarose gel

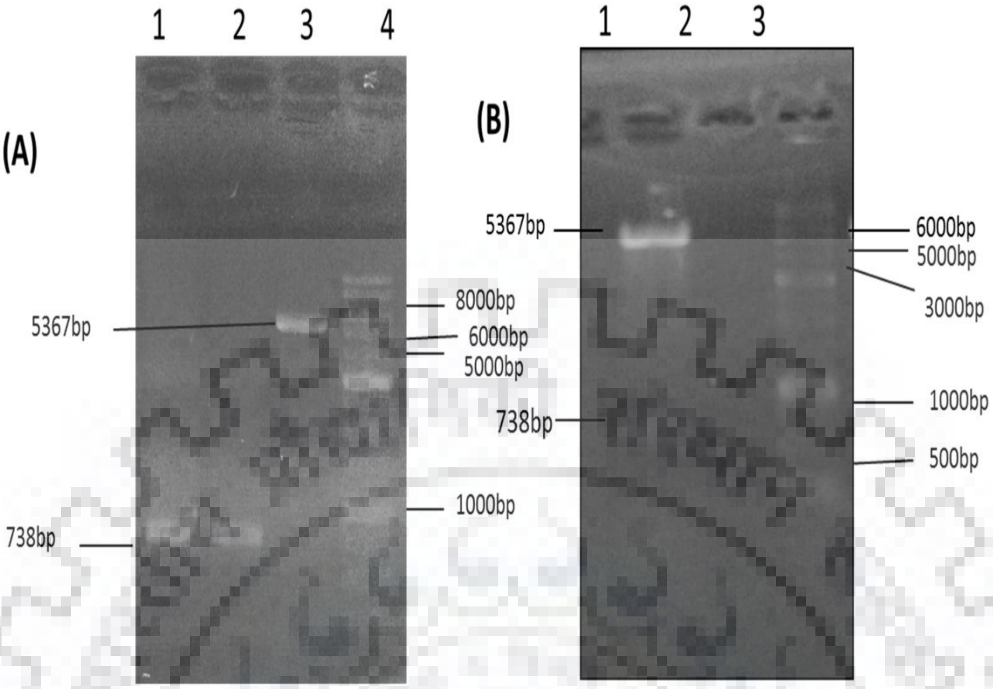


Figure 3.4.3.1: Cloning of AVCP19 (residues 19-265). (A) Polymerase chain reaction (amplification of AVCP 19 (738bp) in lane 1& 2; lane3 shows the pET28c vector, lane4 contains 1Kb DNA ladder (Gene direx). (B)Restriction digestion of the plasmid (clone) by *Nde*I and *Xho*I shows the insert of 739 which corresponds to the size of AVCP19; lane3 contains 1Kb DNA ladder (Gene direx).

3.4.4 Expression and purification of AVCP19

Rosetta cells were transformed by the cloned plasmid for the expression of AVCP19. The transformed *cells* were checked for expression by growing the cells at two different temperatures (37 ºC/4h and 18 ºC/14h) after their log phase. The expression profile of both the culture grown at 2 different temperatures was checked on 15% SDS PAGE. AVCP19 was expressed and was soluble at 37 ºC. AVCP 19 was purified from 1L culture (from the pellet of expressed AVCP19 grown in *Rosetta* cells) by IMAC technique and was eluted at 250mM imidazole. The affinity tag was removed by TEV protease. The cleaved AVCP19 was in the flowthrough and wash 1of reverse Ni-NTA (Fig.3.4.4.1 a, b).

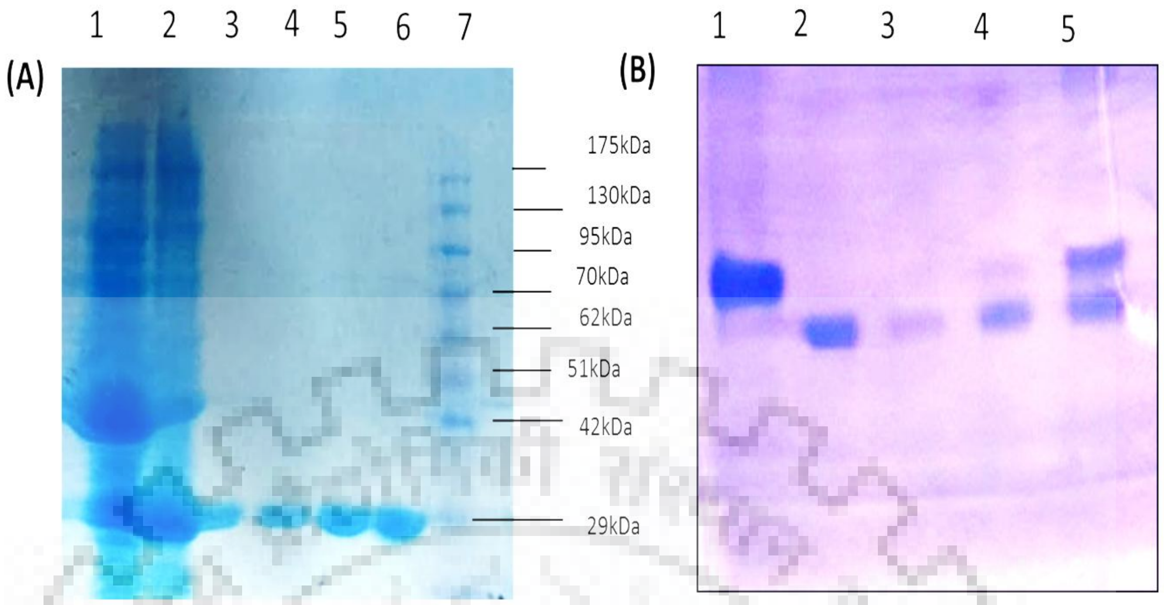


Figure 3.4.4.1: 15 % SDS PAGE of the purification profile of AVCP19. (A) Purification profile of AVCP19**;** Lane1- Supernatant; lane 2- Flowthrough; lane3-6 – elution fractions of AVCP19 ~ 28 kDa; lane7 – pre stained protein ladder (Thermo scientific). (B) Reverse Ni-NTA profile of cleaved AVCP19 ; Lane1- His tagged AVCP19; lane2 – flowthrough (AVCP19 without his tag); lane 3- wash 1; lane4 – wash2; lane 5- wash 3.

3.4.5 Co- crystallisation of AVCP 19

To avail the 3D structure insights of the N-terminal domain of AVCP, co- crystallisation trials of AVCP 19 with 12 mer DNA were done.The reservoir buffer had 100 µl volumes against the 2 µl drop (1µl reservoir buffer: 1µl protein). Tentativediffractable crystals of 0.5 mM AVCP 19-DNA were obtained by vapour diffusion method at 20 ºC in 32 % PEG 3350, 0.1 M $MgCl_2$ which were later discovered to be salt crystals. Co- crystallisation attempts of AVCP 19 with 12 -mer DNA was unsuccessful.

3.4.6 Purification of AVCP 80 (residues 80-267)

AVCP 80 was purified using IMAC technique. AVCP 80 without any impurity was eluted in the final elution of 250 mM imidazole which was confirmed by running the fractions on 15 % SDS PAGE along with the protein marker (prestained protein ladder from Thermoscientific). The affinity tag was removed by adding TEV protease in 1:10 ratio and the protein was dialysed in the dialysis buffer (50 mM Tris, 20 mM NaCl ; pH 7.6) overnight at 4 ºC to remove the imidazole from the protein solution. The cleaved affinity tag was removed along with the

TEV protease via reverse Ni-NTA chromatography. The flowthrough and wash I of reverse Ni-NTA had purified AVCP80 without the affinity tag which was confirmed by running the fractions on 15% SDS PAGE. Figure 3.4.6.1 shows the SDS PAGE profile of purification and reverses Ni-NTA of AVCP80.

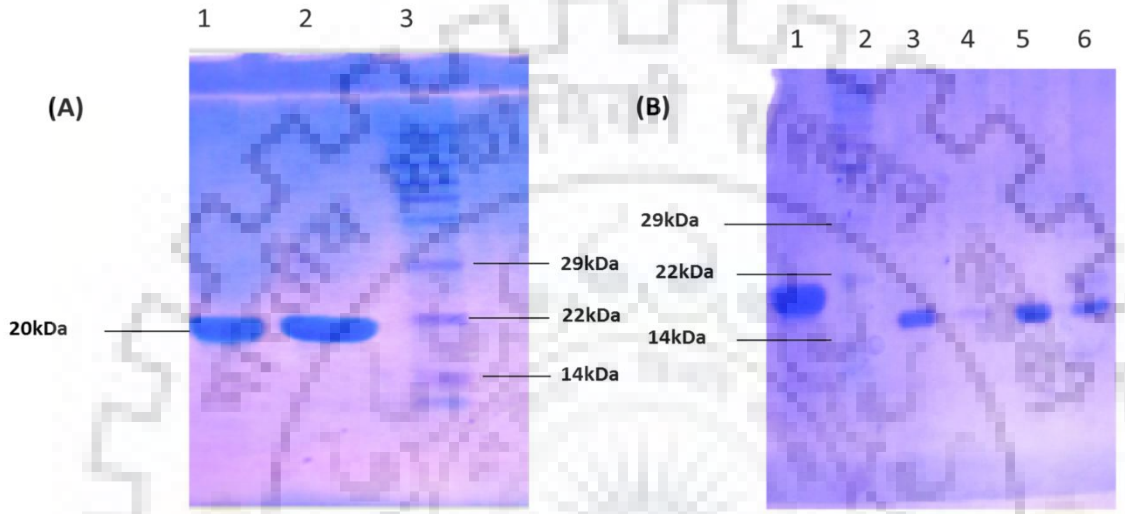


Figure 3.4.6.1: SDS PAGE profile of purification of AVCP 80 and reverse Ni-NTA of AVCP80. (A) Purification of AVCP 80; Lane 1& 2 – elution fractions (250 mM imidazole) of AVCP 80 without impurity, lane 3- protein ladder of Thermoscientific. (B) Reverse Ni-NTA; Lane1- Uncleaved AVCP80 ; lane 2- protein ladder ; lane 3-4 flowthrough fractions ; lane 5 – wash 1(reverse Ni buffer; 50 mM Tris pH 7.6, 20 mM NaCl); lane 6- wash 2 (50 mM imidazole)

3.4.7 Co crystallization of AVCP80 with 12 -mer DNA

An attempt to co-crystallise AVCP80 with 12 mer DNA (5'-CCGTTAATGCAT-3') was made. Crystals were observed after one month in the tray which had 20 % PEG 3350 and 0.15 M $MgCl_2$. The crystal was diffracted at 1.5 Å by beam line ID30A-3 from Synchrotron Grenoble. Prior to the crystal diffraction, the crystal was subjected to the cryoprotectant for enhancing the resolution. The cryoprotectant comprised of 20% ethylene glycol and the mother liquor. The crystal was soaked in the cryoprotectant solution for 2mins. Figure 3.4.7.1 (a, b) shows the crystal image of complex of AVCP80-12mer DNA in the loop and the diffraction pattern of the complex.

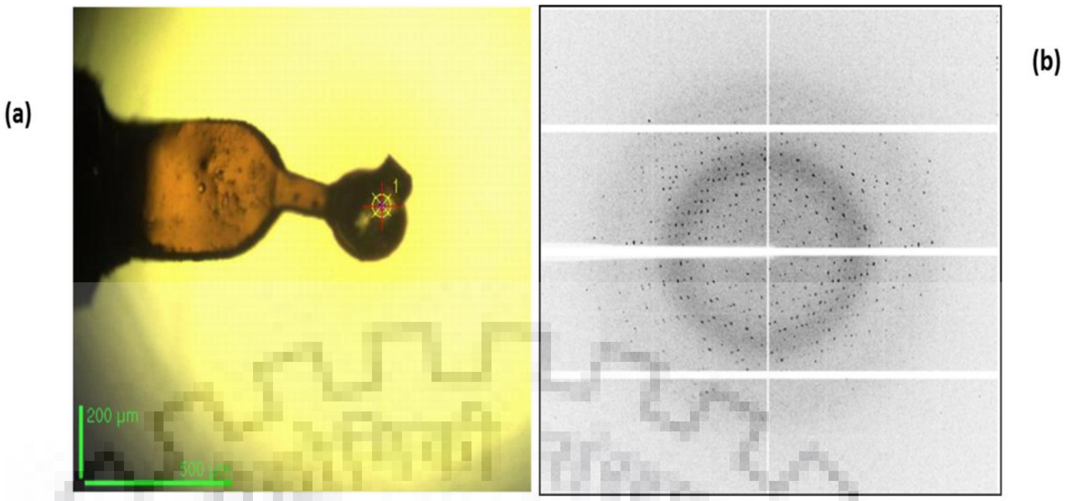


Figure 3.4.7.1: Crystal image of AVCP 80-12mer DNA and Diffraction pattern of crystal at 1.5 Å. (a) Image of crystal in loop prior to mounting (b) Diffraction pattern of crystal when bombarded with X-ray beams.

The crystal structure of AVCP80-12mer DNA was resolved via molecular replacement method using the crystal structure of AVCP (110-267 residues) (PDB ID 4AGK) as template by employing the MOLREP program [157]. The refinement was attained by the Phenix program [158]. COOT program was used for the analysing the electron density map and for manually buiding the model [144]. For the pictorial representation of the analysed structure PYMOL was used (Version 1.7.4 Schrodinger). Figure 3.4.7.2 shows the solved structure of AVCP80-DNA complex.

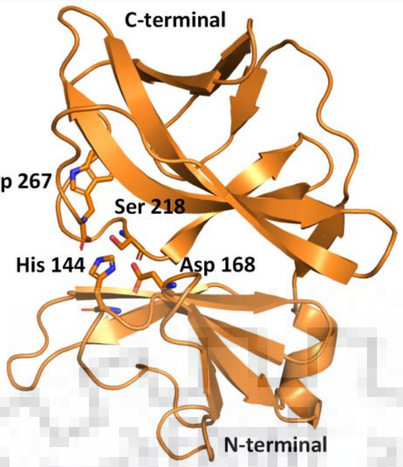

C-terminal
Trp 267
Ser 218
His 144
Asp 168
N-terminal

Figure 3.4.7.2: Structure of AVCP80 -12 mer (DNA). Cartoon view representation of AVCP80- DNA complex comprising of a single chain depicting one molecule in one asymmetric unit.Space group is C121. Electron density of DNA was not found.

The obtained crystal structure of AVCP 80 which had no density of 12-mer DNA was superimposed with the previously reported crystal structure of AVCP 110 (110-267) without DNA having PDB ID 4GEK. The figure 3.4.7.3 shows no major difference. The RMSD of Cα was 0.217.

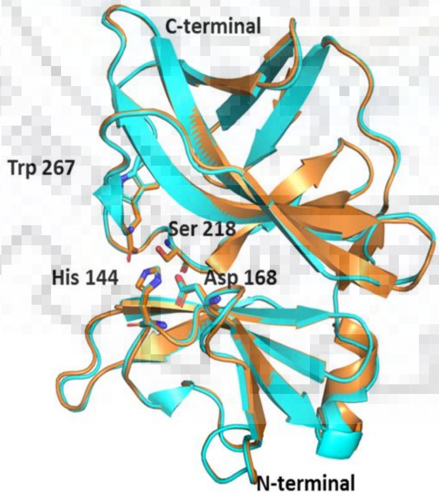

C-terminal
Trp 267
Ser 218
His 144
Asp 168
N-terminal

Figure 3.4.7.3. Cartoon view representation of AVCP 80 superimposed on AVCP 110.AVCP80 is in orange color while AVCP 110 is in cyan color.

Table 3.3 Refinement statistics of resolved crystal structure of AVCP80

Wavelength	32.81 - 1.471 Å
Resolution range	(1.523 - 1.471)
Space group	C 1 2 1
Unit cell	79.03 36.23 49.31 90 101.763 90
Total reflections	162200 (16682)
Unique reflections	23289 (2356)
Multiplicity	7.0 (7.1)
Completeness (%)	99.00 (99.49)
Mean I/sigma(I)	20.57 (2.28)
Wilson B-factor	19.80
R-merge	0.04205 (0.7307)
R-meas	0.04547 (0.789)
R-pim	0.01711 (0.2946)
CC1/2	1 (0.887)
CC*	1 (0.97)
Reflections used in refinement	23255 (2347)
Reflections used for R-free	1161 (95)
R-work	0.1715 (0.2863)
R-free	0.2239 (0.3057)
CC (work)	0.965 (0.866)
CC (free)	0.956 (0.821)
Number of non-hydrogen atoms	1243
Macromolecules	1145
Ligands	11
Solvent	87
Protein residues	152
RMS (bonds)	0.032
RMS(angles)	2.77
Ramachandran favored (%)	99.33
Ramachandran allowed (%)	0.67
Ramachandran outliers (%)	0.00
Rotamer outliers (%)	0.85
Clashscore	2.18
Average B-factor	23.95
	23.07

Macromolecules	
Ligands	43.21
Solvent	33.05

3.4.8 Size Exclusion Chromatography to study AVCP19-DNA interaction

To detect the DNA binding interaction with the N-terminal domain, size exclusion chromatography technique was employed. The protein (AVCP19) was incubated with the 48mer DNA in Equimolar concentration overnight which when run through the 75 Superdex column showed 3 peaks (Fig 3.4.8.2). One among the three peaks (3^{rd}) showed elution at 72.5ml which corresponded to the elution volume of AVCP 19 when passed alone without any DNA (control)(Fig 3.4.8.2). The other two peaks (1^{st} and 2^{nd}) eluted at 54 ml and 63.39 ml respectively. The fractions comprising to their respective peaks were run on 15% SDS PAGE and 0.8% agarose gel.The SDS PAGE profile showed bands for 3^{rd} and 2^{nd} peak which corresponded to the band of AVCP19 (control; purified AVCP19 before the gel filtration). The 2^{nd} peak showing the protein band speculates the AVCP19-DNA interaction (Fig 3.4.8.3). The agarose gel profile showed DNA band for the 2^{nd} peak which confirmed the AVCP19-DNA interaction (Fig 3.4.8.4)

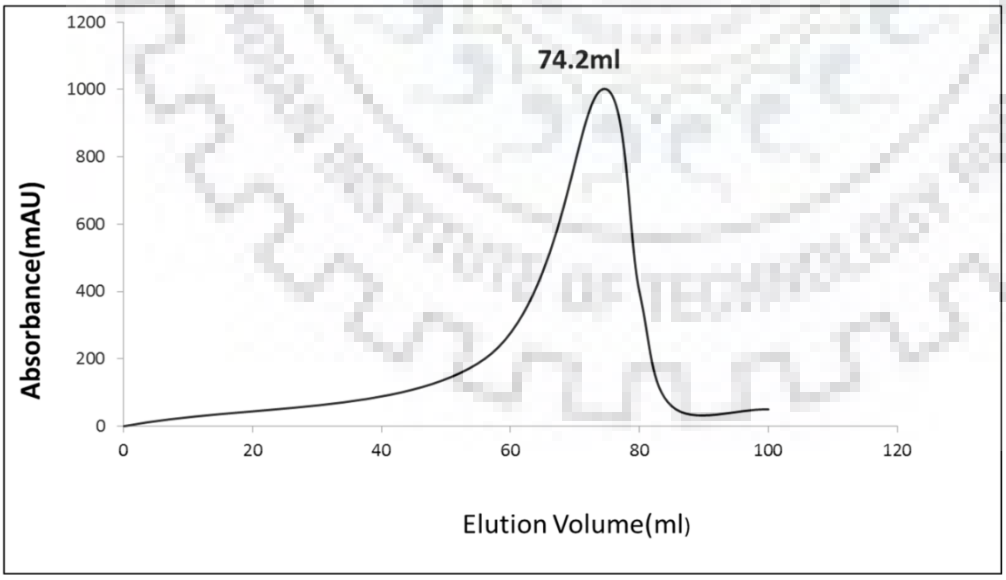
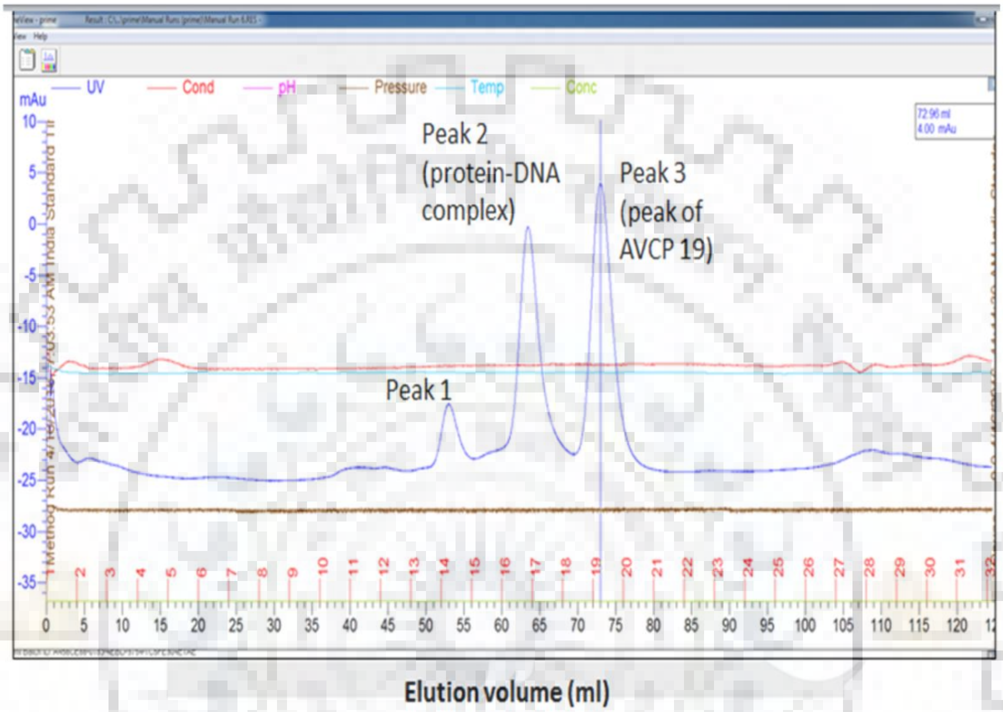
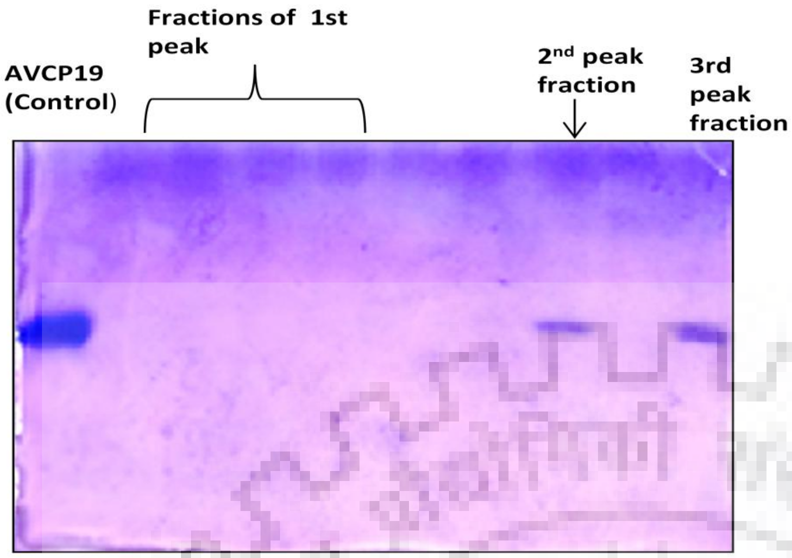


Figure 3.4.8.1: Gel filtration chromatography of AVCP19: The peak of AVCP19 shows elution at 74.2 ml

Figure 3.4.8.2 Gel filtration chromatography of AVCP19 - 48 mer DNA showing the elution volumes peaks. The 2nd peak showing elution at 65 ml suggests the peak of AVCP19 and DNA while the 3rd peak corresponds to the elution volume of AVCP19 at 74.2 ml.



The samples corresponded to their respective peaks were run on 12 % SDS PAGE to confirm the predicted peaks. Lane 8 and 10 showed protein bands of AVCP19 (figure 3.4.8.3).


Fractions of 1st peak
AVCP19 (Control)
2nd peak fraction
3rd peak fraction

3.4.8.3 SDS PAGE profile of eluted samples of gel filtration of AVCP19: Lane1- AVCP19 (before the gel filtration; marker); lane 2-5: fractions of 1st peak; lane 8 – fraction of 2nd peak (DNA bound AVCP 19); lane10 – fraction of 3rd peak (AVCP 19).The predicted peak (2nd) of DNA bound AVCP19 was confirmed by running the sumple on 0.8 % agarose gel which is shown in figure 3.4.8.4.

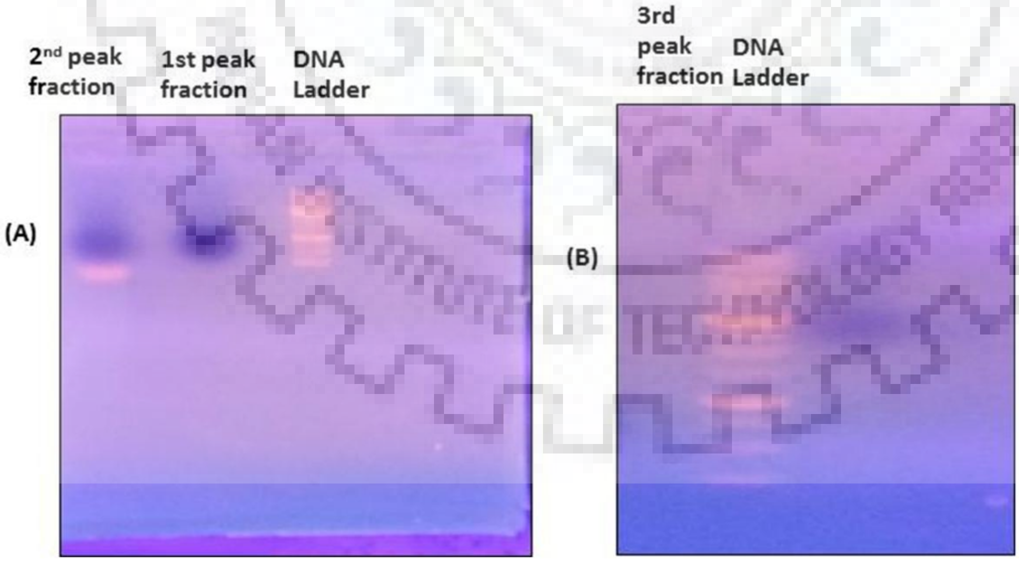

2nd peak fraction
1st peak fraction
DNA Ladder
(A)
3rd peak fraction
DNA Ladder
(B)

Figure 3.4.8.4: Agarose gel profile of eluted samples ofgel filtration of AVCP19: (A) Lane 1 – eluted fraction of 2nd peak; lane 2- eluted fraction of 1stpeak; lane 3- DNA ladder. (B)Lane

1- fraction of 3rd peak,lane 2- DNA ladder.The agarose gel profile shows DNA band in lane 1 which is the eluted fraction of 2nd peak (predicted peak of AVCP 19 – 48- mer DNA complex).

3.4.9 Differential Scanning Calorimetry to study the change in stability of N-terminal domain of AVCP19 when bound to DNA

After the detection and confirmation of the binding of 48 mer DNA to the N-terminal domain of AVCP19, its effect on the stability of the CP was observed by employing differential scanning calorimetry technique. Equimolar concentration of AVCP19 and 48 mer DNA was incubated overnight and its Tm was checked along with the control (AVCP19 only of same concentration) by DSC. The Tm of AVCP19-48 mer DNA was more than the AVCP19 alone. This indicates the interaction of 48-mer DNA with the purified AVCP19 comprising of the basic charged residues in the N-terminal domain. Fig. 3.4.9.1 shows the thermogram of (a) AVCP19 and (b) AVCP19-48mer DNA

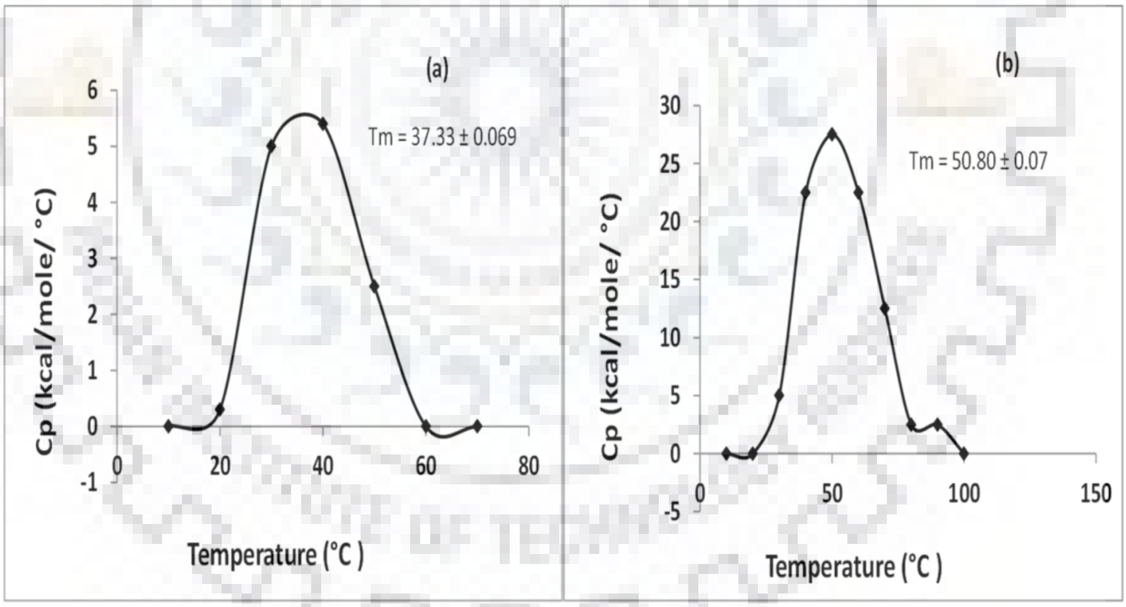


Figure 3.4.9.1: Thermogram of AVCP19 & AVCP19-48mer DNA using differential Scanning calorimetry.(a) Tm of AVCP19 is 37.33ºC ± 0.069. (b)Tm of AVCP19 bound 48-mer DNA was 50.80 ºC± 0.07.The increase in Tm of AVCP19 bound 48 mer DNA confirms the binding of DNA to the N-terminal domain of CP and suggests stability to the N-terminal domain by charge-neutralization

3.4.10 *Trans* –proteaseactivity of AVCP19

The C-teminal trp residue is absent in AVCP 19 which makes the protease domain vulnerable to exhibit *trans*- protease activity. This was confirmed by employing the substrate specific FRET based assay. Upon the addition of 0.6 µM fluorogenic peptide (DABCYL-Gly-Ala-Glu-Glu-Trp↓Ser-Leu-Ala-Ile-Glu-EDANS) in the 20 mM HEPES pH 7.0 containing 2 µM AVCP 19, proteolytic cleavage occurred between the scissile bond trpand ser followed by the increase in fluorescence with respect to the control (no AVCP 19). The final reaction volume was 100 µl for each sample. All reactions were carried out in triplicate. The initial velocity (V_i) of the AVCP 19 (2 µM) was calculated at a varied range of substrate concentration (0.6 µM to 8 µM).By plotting V_i of AVCP 19 at different substrate concentrations in the Lineweaver – burk plot, linear progression was observed with respect to the ascending substrate concentrations (Figure 3.4.10.1). The kinetic parameters of AVCP 19 were calculated using graph pad prism software by fitting the V_i values with respect to their substrate concentrations. The Vmax and *K*m calculated for AVCP 19 were 1.257±0.06 µM min^{-1} and 0.8321±0.158µM. The catalytic efficiency (*k*cat/ *K*m) calculated for AVCP 19 was 0.8µM^{-1} min^{-1}.

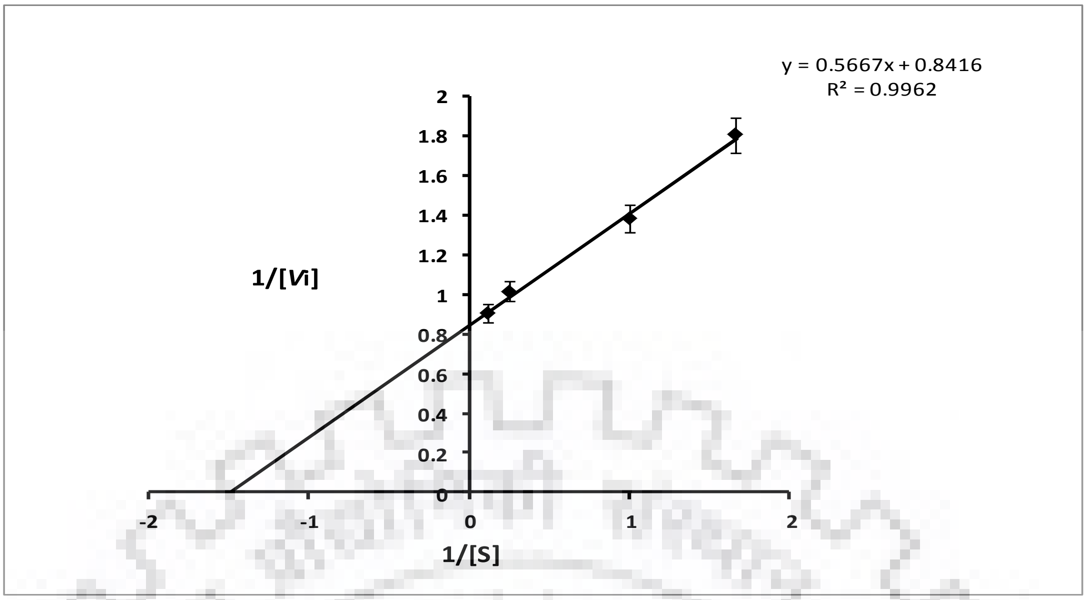


Figure 3.4.10.1 Kinetic analysis of AVCP 19. Kinetic parametes for AVCP 19 were determined using fluorogenic peptide/substrate. The range of substrate concentrations were 0.6µM, 1µM, 4 µM and 8 µM. The inititial velocity (reaction rates; V_i) were plotted at their respective substrate concentrations.

The Line-weaver burk plot was obtained by plotting 1/ [V_i] vs. 1/[S]. The Vmax and Km were determined by using the equation y = mx +c by calculation slope and the intercept. *k*cat was calculated by dividing Vmax by enzyme total (2 µM).

3.4.11 *Trans*- protease activity of AVCP 19 bound 48- mer DNA (2 µM: 2 µM)

The *trans*- protease activity of AVCP19 bound DNA was determined in the same fashion as described in 3.4.10. The calculated Vmax, *K*m and catalytic efficiency (*k*cat/Km) were 1.105.0±0.007 µM min^{-1}, 0.53± 0.01 µM and 1.04μM^{-1} min^{-1}.

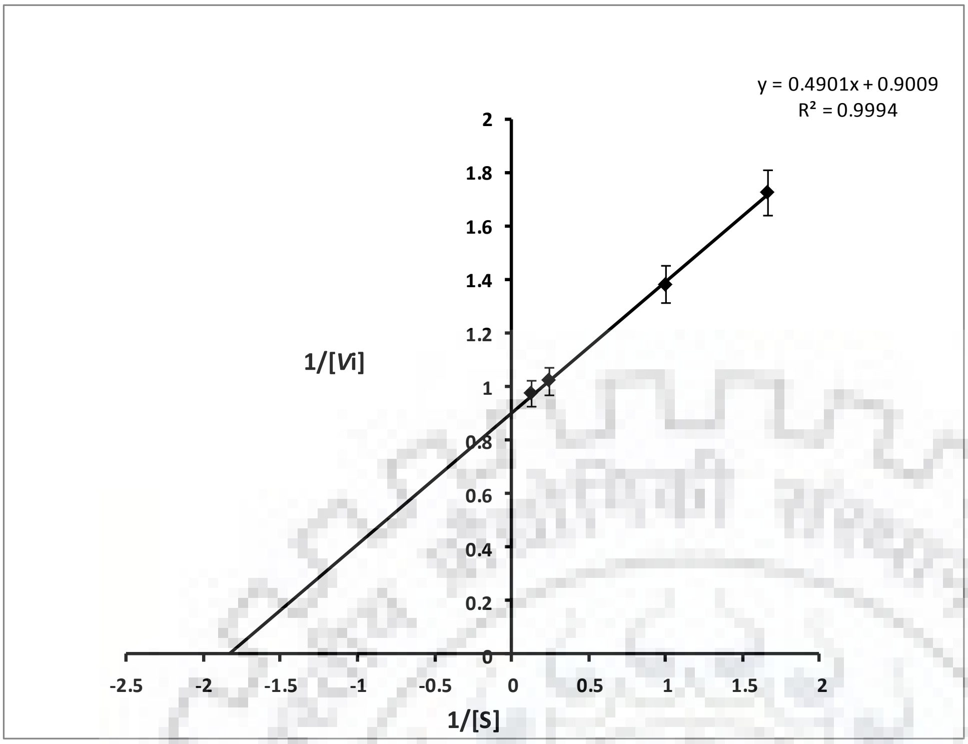


Figure 3.4.11.1 Kinetic analysis of AVCP 19 bound 48- mer DNA. Kinetic parametes for AVCP 19-DNA complex were determined using fluorogenic peptide/substrate. The range of substrate concentrations were 0.6 µM, 1µM, 4 µM and 8 µM. The inititial velocity (reaction rates; V_i) were plotted at their respective substrate concentrations. The Line-weaver burk plot was obtained by plotting 1/ [V_i] vs. 1/[S]. The Vmax and *K*m were determined by using the equation y = mx +c by calculation slope and the intercept. *k*cat was calculated by dividing Vmax by enzyme total (2 µM).

The catalytic efficiency of AVCP 19-DNA (1.04 $\mu M^{-1} min^{-1}$) complex was more than AVCP 19 (0.81 $\mu M^{-1} min^{-1}$).But the difference is very negligible. This suggests, the *trans*- protease activity of AVCP 19 is independent of the DNA/ nucleic acid interactions.

3.5 Conclusion

This study attempts to decipher the 3-dimensional structure of N-terminal domain of alphavirus CP and to study the effect on the protease activity of CP on inding to the nucleic acid. AVCP-19 crystals could not be obtained by co- crystallising it with 12 mer-DNA. Salt crystals were frequently obtained under all tried conditions with varied pH. This could be due to the fact the protein was in phosphate buffer. AVCP 19 yield and stability is very poor in Tris buffer.

However it is know that small constructs are easy to crystallize so an attempt was made to crystallize AVCP80 with DNA. Crystals of AVCP 80 were obtained by co- crystallising it with DNA but no electron density of N-terminal domain was found, also density of DNA was absent. This suggests the binding condition of AVCP 19 and AVCP 80 with DNA need to be improvised during co- crystallisation. This study reports for the first time the CP-DNA interaction in alphavirus by size exclusion chromatography and DSC. The DSC-based result for AVCP 19- DNA (48-mer) complex showed shift in the Tm for AVCP19-DNA complex. This suggests the interactions provide stability to the N-terminal domain. These *in vitro* studies of CP-DNA interaction can help in developing a HTP assay which can be exploited to screen inhibitors that can block the molecular interactions between N-terminal domain of CP and Nucleic Acid. In this chapter, *trans*- protease activity of AVCP19 was detected along with the DNA bound AVCP 19. The catalytic efficiency of DNA bound AVCP 19 did not differ much from the unbound AVCP 19 which suggests the interactions between nucleic acid and N-terminal domain does not have any effect on the *trans*– proteolytic activity of AVCP 19.

CVCP20 could not be purified by IMAC which might be because of positive charge residues in the N-terminal domain creating hindrance in the outward exposure of His Tag to the Ni –NTA column.

भारतीय प्रौद्योगिकी संस्थान रुड़की
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

CHAPTER 4

Discussion (Alphavirus capsid protease – a potential antiviral target)

4.1 Structural and functional significance of N-terminal domain of alphavirus CP

The genus alphavirus constitutes at least 8 pathogenic viruses which hold medical importance. All these viruses are mosquito borne.These pathogens pose severe health threats to humans and livestock and cause range of diseases which include rash, infectious arthritis, febrile illness and encephalitis. These notoriety acclaimed viruses are CHIKV, SINV, SFV, WEEV, EEEV, VEEV, and RRV. They are geographically distributed worldwide which is quite alarming as they are spread mainly by mosquito bite [18]. The developing nations mostly bear the brunt of the epidemic caused by these viruses. VEEV has been categorized among the bioweapons which is attributed by the easy transmission of virus through aerosols of infected mosquitoes [159]. The resurgence of CHIKV has posed a grim worldwide health threat and till date no vaccines have been developed, nor are any drug is available to combat the disease [25]. CHIKV categorized among the old alphavirus requires nsP2 protein for inhibiting the host transcription and is associated with the symptoms mainly skin rash, polyarthalgia, headache, nausea and fever. The new-world alphavirus, EEEV and VEEV, require capsid protein for inhibiting the host transcription and are mostly associated to cause encephalitis [14]. But as per recent reports, CHIKV infection has also been claimed to cause neurological complications including encephalitis [30], [160]. The CP of alphavirus possesses multiple functions all of which are essential for viral life cycle. There are two major functional domains of CP, the N-terminal domain and the C-terminal domain [57], [41]. The C-terminal of CP constitutes serine protease which exhibits *cis*-proteolytic activity. Once the structural polyprotein is translated from the 26S viral subgenomic RNA, the *cis*- proteolytic activity of CP leads to the self-cleavage of CP from the structural polyprotein. This step is an essential part of structural polyprotein processing [43]. After the cleavage of CP from the polyprotein, the polyprotein acquires signal sequence in its N-terminus which guides it to translocate towards ER and Golgi apparatus for its post translation modification and further processing. The cleaved polyprotein is further processed by host cell furin and signal peptidases to generate functional structural proteins E3, E2, E1 and 6K all of which are essential for the budding of virus and in the viral infection [161]. The N-terminal domain of CP is laden with positive charged amino acid residues which help in the encapsidation of genomic and subgenomic RNA to the CP [107], [51]. This CP-

RNA binding promotes NC assembly and NC cores [51]. The genetic content of the virus remains intact by encapsidation and is released during the viral entry in the host cell [56]. The cytoplasmic domain of E2 interacts with the hydrophobic pocket present opposite to the protease domain of CP which triggers the budding of virus [41], [67]. The structure of N-terminal domain is poorly understood due to the unavailability of crystal structures. This is attributed due to the fact that the N-terminal of CP is highly disordered which makes the detection of electron density of the domain cumbersome [5]. The RNA binding to the N-terminal domain of alphavirus CP promotes charge neutralization. The region underlying the N-terminal domain of CP where the viral RNA nonspecifically interacts could be a potential antiviral drug target because this recognition supposedly triggers NC cores assembly by protein-protein interactions via helix I. In the current study the N-terminal domain of AVCP19 (residues 19-265) and its interaction with the DNA has been studied by the size exclusion chromatography technique. The DNA used in the study is a 48-mer single stranded DNA. The sequence of the 48 mer used is in reference to the study of SINV CP that showed *in vitro* assembly of core like particles by cross linking the inter capsomeres dimer [162]. AVCP19 bound 48-mer DNA was eluted earlier than the unbound AVCP19. This finding was in corollary to the studies carried out by Kuhn et al., 2000 and Perera et al., 2001 [54], [107], [162]. This nonspecific DNA CP interaction was further correlated with the stability of CP by analyzing thermogram of unbound AVCP19 and AVCP19 bound 48-mer DNA. It was found; the Tm of DNA bound AVCP 19 was more than the unbound AVCP19. This speculates the non- specific interaction of the nucleic acid to the N-terminal domain of CP provides stability to the CP structure by charge neutralization. Further *trans*-protease activity was checked for AVCP19 bound 48-mer DNA using substrate specific FRET based assay employing similar technique as conducted by Aggarwal et al., 2015 on CVCP active (residues 106-259) [5]. It was found, the *trans*- proteolytic activity of AVCP 19 is independent of the DNA (nucleic acid) - N-terminal domain by exhibiting no remarkable difference in the catalytic efficiency between unbound AVCP 19 and DNA bound AVCP 19. An attempt was made to conduct these experiments on CVCP20 (residues 20-259) owing to the medical importance of CHIKV. Due to the difficulty in purification of his tagged CVCP20 by IMAC (heterologous expressed in *Rosetta* cells, cloned in pET28c vector) the experiments could not be executed. This might be because of the positively charged amino acid residues constituting N-terminal domain of CVCP creating hindrance in the outward exposure of his tag (affinity tag). In SINV CP residues 1-114th is involved in RNA binding while the protease domain starts from residue 115-267 [107], [151]. The crystal structure of AVCP (residues 110-259) having PDB ID 4AGK shows electron density for amino acid residues from 116th; the density for 6 residues 110th- 115th is not reported [5]. An attempt was made to co-crystallize AVCP 19 and AVCP 80 with the 12- mer DNA separately to find out the disordered structure of N-terminal domain. In accordance to the hypothesis that the charge neutralization will provide stability to the disordered region and probability of finding the electron density for the N-terminus could be achieved by co-crystallising it with DNA/ RNA. The crystal of AVCP 19 with 12-mer DNA could not be obtained. Crystal of AVCP 80 with 12-mer DNA was obtained after 1 month of incubationin room temperature in an optimum condition by exploiting vapor diffusion phenomenon. Crystal was diffracted at 1.5Å. The crystal data did not have electron density for N-terminus of AVCP80 and 12-mer DNA but it showed electron density for extra 4 residues for (111- 114) which are not reported in the crystal structure of AVCP having PDB ID 4AGK. This is still suggestive of the fact that the DNA in the crystallization complex might have minimal interactions with the N-terminal domain of AVCP which provided stability to disordered region and due to which electron density of extra 4 residues of the N-terminus was found. The optimum crystal condition of AVCP 80 has been found which can also help in improvising co-crytallisation conditions with DNA/ RNA.

All the above studies provide insights into the structural and functional importance of N-terminal domain of alphavirus CP.

Future prospective of these studies in gist:

(a) Conducting similar experiments on CP of pathogenic alphaviruses. The CVCP purification can be optimized by other chromatographic techniques like ion exchange apart from IMAC.
(b) Co crystallization of AVCP 80 and AVCP 19 with the DNA/RNA having sequence similar to the viral RNA which specifically binds to the region 81-114 as reported by Weiss et al., 1994, BR et al.,2004 and Owen et al.,1996 [50,165,166].
(c) The DNA binding region can be explored as antiviral drug target if three dimensional structure of N-terminal domain is revealed.

4.2 Structure based screening and evaluation of C-terminal domain of CVCP protease inhibitors for drug repurposing

The C- terminal region has protease domain, having a hydrophobic pocket wherein the cd E2 glycoprotein binds to it, enabling virus budding [57], [108]. Dioxane, piperazine and picolinate are hydrophobic molecules which apparently bind to the conserved hydrophobic pocket in CP, and have been proposed to inhibit CP–cdE2 interaction which nips the budding process [41], [45]. The C-terminal domain of CP has cis-autoproteolytic activity which allows the detachment of CP from the structural polyprotein, an essential requirement for the processing of the structural polyprotein and life cycle of the virus and the *in vitro* trans protease activity of CVCP has been determined by peptide/substrate specific FRET based assay [4], [5], [140].

The objective of this study was to find anti-CHIKV drugs (from the library of FDA approved drugs) which can bind the active site of capsid protease and inhibit the proteolytic activity of the CP thus, eventually putting halt to the viral infection. The catalytic triad has conserved His, Ser and Asp confined in a cleft interspersed among ß barrel chains [164]. Various potential inhibitors for AVCP (110-265) were *in silico* screened using structure based approach and four compounds which showed the best binding affinity with the active site were chosen. Due to the failed crystallization attempts of CVCP active and unavailability of crystal structure of CVCP active, the conserved terminal Trp261 of CVCP present in the native structure of CVCP (PDB ID: 5H23) was computationally removed from the structure to make the active site available for the potential inhibitors to bind to the active site in the molecular docking. By molecular docking, four compounds were identified which showed best hits. Each compounds showed binding and molecular interactions with the active site or substrate binding residues of CVCP. Anti-protease activity of these was further validated by an *in vitro* approach implying the substrate specific FRET based enzyme inhibition assay. The proteolytic activity of CVCP was inhibited by increasing concentration of compounds at different substrate concentrations. Remarkable decrease in RFU/min was observed when compared with the RFU/min of the control (proteolytic assay without inhibitor), conferring the binding of these compounds to the active site of the CVCP protease. The inhibition kinetics for compounds AP4, EAC, PSU and DRV was set by analyzing the rate of proteolytic cleavage of fluorogenic peptide by CVCP in the substrate specific FRET based assay using 4 different concentrations of substrate ranging from 2μM to 12μM along with the increasing concentration of compounds. The double-reciprocal plots for the AP4 and DRV showed reduced K_m and unaltered V_{max} which signified competitive mode of inhibition. The double reciprocal plot of compound PSU showed variation in the V_{max} for the fluorogenic peptide upon increasing the concentration of PSU while the K_m remained unaltered, suggesting the mode of inhibition to be non-competitive which is contrary of the conformation of the CVCP active site impeding the peptide catalysis. This attributes to a possible allosteric binding site in the CVCP for the PSU binding. The FRET based inhibition assay for compound EAC showed mixed mode of inhibition where K_m was reduced with increased V_{max}. Theα value of this inhibition was 5.34 (calculated from graph pad prism) whichwas greater than 1. This suggested the fact the inhibition is more of a competitive and less of non-competitive. EAC binding to the allosteric site is less than the EAC binding to the active site of CVCP. The inhibitor constant K_i of all the compounds were determined and found to be in micromolar concentration. The inhibitory potency of these compounds was further verified by cell culture-based antiviral assays –plaque based reduction assay, qRT-PCR, immuno-fluorescence assay and western immuno blot assay.

The cytotoxic effects of AP4, EAC, PSU and DRV on Vero cells were conducted by MTT assay. The effective concentration of AP4, EAC, PSU and DRV for antiviral activity was much lower than MNTD (maximum non-toxic dose) to eliminate the possibility of compound mediated cytotoxicity. A time-of-addition assay was performed using different concentrations of AP4, EAC, PSU and DRV compounds to determine the stage of infection where the maximum antiviral effect was observed. AP4, EAC and PSU showed remarkable reduction in viral load based on the plaque assay while DRV showed ambiguous reduction. For these compounds, reduction in CHIKV infection was more observed in the post treatment as compared to pre and simultaneous treatment. These results suggested that AP4, EAC and PSU significantly suppressed plaque formation at post-entry events. Further, the inhibitory potency of AP4, EAC and PSU was evaluated to quantify EC_{50} value in low micromolar concentration which concretes the antiviral potency of these compounds. To check the effect of these inhibitors on viral RNA synthesis, minus strand specific reverse transcription was done and it wound these inhibitors don't affect the level of negative strand RNA. This suggests the inhibitors don't show any effect in initial stage of viral RNA synthesis. Results of Western-immunoblot assay by using anti CHIKV CP antibody showed the inhibiotrs specifically target CP. To quantify the reduction of viral RNA copy number in the cell lysate, the quantitative qRT-PCR was performed using ΔΔCt method. The prominent reduction in viral RNA copy number was observed in post treatment only while there was no significant reduction at pre and simultaneous-treatment. Further, anti-CHIKV activity of these compounds at post-entry stage was measured by IFA. The significant reduction in CHIKV E2 glycoprotein in the form of reduced FITC intensity was observed in the CHIKV infected Vero cells when treated with the all three compounds (post treatment). Down regulation of E2 glycoprotein clearly demonstrated the inhibition occurred at post entry stage events durion the virus release. All the disussed antiviral assays suggested the inhibitors targeted the CHIKV at post entry stage events.

Future prospective:

The inhibitory potency of DRV could not be further tested as it did not show inhibition in the plaque reduction assay. However its potency can be further refined by using it in combination with other three compounds or any other potential antiviral drug. Designing its structural analogues and testing them in cell based antiviral assays might lead to the finding of new potential anti CHIKV drug. Co-crystallising these compounds with the CVCP active can help in better understanding of the interactions of these compounds with the residues of the active site. The other three reported inhibitors can be further tested for *in vivo* experiments which can build a tarmac for the introduction of the potential effective anti CHIKV drugs in the market.

भारतीय प्रौद्योगिकी संस्थान रुड़की
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

भारतीय प्रौद्योगिकी संस्थान रुड़की
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

REFERENCES

[1] A. Mendes and R. J. Kuhn, “Alphavirus nucleocapsid packaging and assembly.,” viruses, vol. 10, no. 3, 2018.

[2] L. Dalgarno, C. M. Rice, and J. H. Strauss, “Ross River virus 26 s RNA: complete nucleotide sequence and deduced sequence of the encoded structural proteins.,” Virology, vol. 129, no. 1, pp. 170–87, Aug. 1983.

[3] C. M. Rice and J. H. Strauss, “Nucleotide sequence of the 26S mRNA of Sindbis virus and deduced sequence of the encoded virus structural proteins.,” Proc. Natl. Acad. Sci., vol. 78, no. 4, pp. 2062–2066, Apr. 1981.

[4] H. K. Choi et al., “Structure of Sindbis virus core protein reveals a chymotrypsin-like serine proteinase and the organization of the virion.,” Nature, vol. 354, no. 6348, pp. 37–43, Nov. 1991.

[5] M. Aggarwal, R. Sharma, P. Kumar, M. Parida, and S. Tomar, “Kinetic characterization of trans-proteolytic activity of chikungunya virus capsid protease and development of a FRET-based HTS assay,” Sci. Rep., vol. 5, no. 1, p. 14753, Dec. 2015.

[6] V. Lulla, D. Y. Kim, E. I. Frolova, and I. Frolov, “The amino-terminal domain of alphavirus capsid protein is dispensable for viral particle assembly but regulates rna encapsidation through cooperative functions of its subdomains,” J. Virol., vol. 87, no. 22, pp. 12003–12019, Nov. 2013.

[7] K. E. Owen and R. J. Kuhn, “Identification of a region in the Sindbis virus nucleocapsid protein that is involved in specificity of RNA encapsidation.,” J. Virol., vol. 70, no. 5, pp. 2757–63, May 1996.

[8] D. Baltimore, “Expression of animal virus genomes.,” Bacteriol. Rev., 1971.

[9] C. H. Hoke, “History of U.S. military contributions to the study of viral encephalitis,” mil. med., 2015.

[10] M. La Linn et al., “Arbovirus of marine mammals: a new alphavirus isolated from the elephant seal louse, *Lepidophthirus macrorhini*,” J. Virol., vol. 75, no. 9, p. 4103, May 2001.

[11] E. Petterson, M. Sandberg, and N. Santi, “Salmonid alphavirus associated with *Lepeophtheirus salmonis* (Copepoda: Caligidae) from Atlantic salmon, Salmo salar L.,” J. Fish Dis., vol. 32, no. 5, pp. 477–479, May 2009.

[12] T. P. Monath, T. W. Scott, and G. S. Bowen, “A field study on the effects of fort morgan virus, an arbovirus transmitted by swallow bugs, on the reproductive success of cliff swallows and symbiotic house sparrows in morgan county, colorado, 1976 *,” Am. J. Trop. Med. Hyg., vol. 33, no. 5, pp. 981–991, Sep. 1984.

[13] C. Wilks, “Alphaviruses: current biology. , Edited by S Mahalingam, LJ Herrero and BL Herring. Caister Academic Press, Norfolk, UK. vol. 94, no. 3, pp. 59–59, Mar. 2016.

[14] N. Garmashova, R. Gorchakov, E. Volkova, S. Paessler, E. Frolova, and I. Frolov, “The old world and new world alphaviruses use different virus-specific proteins for induction of transcriptional shutoff,” J. Virol., vol. 81, no. 5, pp. 2472–2484, Mar. 2007.

[15] L. Wesula Olivia et al., “Global emergence of alphaviruses that cause arthritis in humans,” Infect. Ecol. Epidemiol., vol. 5, no. 1, p. 29853, Jan. 2015.

[16] A. Paredes, S. Weaver, S. Watowich, and W. Chiu, “Structural biology of old world and new world alphaviruses,” in infectious diseases from nature: mechanisms of viral emergence and persistence, Vienna: Springer-Verlag, 2005, pp. 179–185.

[17] C. H. Calisher et al., “Proposed antigenic classification of registered arboviruses i. togaviridae, &lt;i&gt;Alphavirus&lt;/i&gt;,” Intervirology, vol. 14, no. 5–6, pp. 229–232, 1980.
[18] A. L. Schmaljohn and D. McClain, Alphaviruses (Togaviridae) and Flaviviruses (Flaviviridae). University of Texas Medical Branch at Galveston, 1996.
[19] L. V. Simon and M. A. Fischer, Western Equine Encephalitis. StatPearls Publishing, 2019.
[20] N. A. Bergren, A. J. Auguste, N. L. Forrester, S. S. Negi, W. A. Braun, and S. C. Weaver, “Western equine encephalitis virus: evolutionary analysis of a declining alphavirus based on complete genome sequences.,” J. Virol., vol. 88, no. 16, pp. 9260–7, Aug. 2014.
[21] S. Kurkela, T. Manni, J. Myllynen, A. Vaheri, and O. Vapalahti, “Clinical and laboratory manifestations of sindbis virus infection: Prospective Study, Finland, 2002–2003,” J. Infect. Dis., vol. 191, no. 11, pp. 1820–1829, Jun. 2005.
[22] H. Jöst et al., “Medical importance of Sindbis virus in south-west Germany,” J. Clin. Virol., vol. 52, no. 3, pp. 278–279, Nov. 2011.
[23] A. M. Powers et al., “Evolutionary relationships and systematics of the alphaviruses.,” J. Virol., vol. 75, no. 21, pp. 10118–31, Nov. 2001.
[24] C. S. Hahn, S. Lustig, E. G. Strauss, and J. H. Strauss, “Western equine encephalitis virus is a recombinant virus.,” Proc. Natl. Acad. Sci., vol. 85, no. 16, pp. 5997–6001, Aug. 1988.
[25] V. K. Ganesan, B. Duan, and S. P. Reid, “Chikungunya Virus: Pathophysiology, mechanism, and modeling.,” Viruses, vol. 9, no. 12, 2017.
[26] “Chikungunya.” [Online]. Available: https://www.who.int/en/news-room/fact-sheets/detail/chikungunya. [Accessed: 18-Apr-2019].
[27] C. Caglioti, E. Lalle, C. Castilletti, F. Carletti, M. R. Capobianchi, and L. Bordi, “Chikungunya virus infection: an overview.,” New Microbiol., vol. 36, no. 3, pp. 211–27, Jul. 2013.
[28] V. Rougeron, I.-C. Sam, M. Caron, D. Nkoghe, E. Leroy, and P. Roques, “Chikungunya, a paradigm of neglected tropical disease that emerged to be a new health global risk,” J. Clin. Virol., vol. 64, pp. 144–152, Mar. 2015.
[29] P. Gérardin et al., “Multidisciplinary prospective study of mother-to-child chikungunya virus infections on the island of la réunion,” PLoS Med., vol. 5, no. 3, p. e60, Mar. 2008.
[30] P. Gérardin et al., “Chikungunya virus-associated encephalitis: A cohort study on La Réunion Island, 2005-2009.,” Neurology, vol. 86, no. 1, pp. 94–102, Jan. 2016.
[31] T. Couderc and M. Lecuit, “Chikungunya virus pathogenesis: From bedside to bench,” Antiviral Res., vol. 121, pp. 120–131, Sep. 2015.
[32] T. E. Morrison, “Reemergence of chikungunya virus,” J. Virol., vol. 88, no. 20, pp. 11644–11647, Oct. 2014.
[33] A. L. P. Mosimann, M. K. de Siqueira, L. F. Ceole, and C. Nunes Duarte dos Santos, “A new Aura virus isolate in Brazil shows segment duplication in the variable region of the nsP3 gene,” Parasit. Vectors, vol. 11, no. 1, p. 321, Dec. 2018.
[34] C. H. Calisher, N. Karabatsos, J. S. Lazuick, T. P. Monath, and K. L. Wolff, “Reevaluation of the Western equine encephalitis antigenic complex of alphaviruses (family Togaviridae) as determined by neutralization tests.,” Am. J. Trop. Med. Hyg., vol. 38, no. 2, pp. 447–52, Mar. 1988.
[35] S. C. Weaver et al., “Recombinational history and molecular evolution of western equine encephalomyelitis complex alphaviruses,” 1997.
[36] T. Rümenapf, E. G. Strauss, and J. H. Strauss, “Aura Virus Is a New World Representative of Sindbis-like Viruses,” Virology, vol. 208, no. 2, pp. 621–633, Apr. 1995.

[37] J. H. Strauss and E. G. Strauss, "The alphaviruses: gene expression, replication, and evolution.," Microbiol. Rev., vol. 58, no. 3, pp. 491–562, Sep. 1994.

[38] J. Jose, J. E. Snyder, and R. J. Kuhn, "A structural and functional perspective of alphavirus replication and assembly," Future Microbiol., vol. 4, no. 7, pp. 837–856, Sep. 2009.

[39] A. J. Snyder and S. Mukhopadhyay, "The alphavirus E3 glycoprotein functions in a clade-specific manner.," J. Virol., vol. 86, no. 24, pp. 13609–20, Dec. 2012.

[40] P. Melancon and H. Garoff, "Processing of the Semliki Forest virus structural polyprotein: role of the capsid protease.," J. Virol., vol. 61, no. 5, pp. 1301–9, May 1987.

[41] J. H. Strauss, E. G. Strauss, and R. J. Kuhn, "Budding of alphaviruses.," Trends Microbiol., vol. 3, no. 9, pp. 346–50, Sep. 1995.

[42] C. Erwin, D. T. Brown, I. Frolov, and E. Frolova, "Requirement of cell nucleus for Sindbis virus replication in cultured *Aedes albopictus* cells.," J. Virol., vol. 45, no. 2, pp. 792–9, Feb. 1983.

[43] S. Tomar and M. Aggarwal, "Structure and function of alphavirus proteases," in Viral Proteases and Their Inhibitors, Elsevier, 2017, pp. 105–135.

[44] T. Rumenapf, E. G. Strauss, and J. H. Strauss, "Subgenomic mRNA of Aura alphavirus is packaged into virions," Journal of Virol. vol. 68, pp.56-62, Jan 1994.

[45] M. Aggarwal et al., "Crystal structure of aura virus capsid protease and its complex with dioxane: new insights into capsid-glycoprotein molecular contacts," PLoS One, vol. 7, no. 12, p. e51288, Dec. 2012.

[46] R. Sharma et al., "Inhibition of chikungunya virus by picolinate that targets viral capsid protein," Virology, vol. 498, pp. 265–276, Nov. 2016.

[47] M. Aggarwal et al., "Evaluation of antiviral activity of piperazine against chikungunya virus targeting hydrophobic pocket of alphavirus capsid protein," Antiviral Res., vol. 146, pp. 102–111, Oct. 2017.

[48] G. Shin, S. A. Yost, M. T. Miller, E. J. Elrod, A. Grakoui, and J. Marcotrigiano, "Structural and functional insights into alphavirus polyprotein processing and pathogenesis."

[49] R. Sharma, P. Kesari, P. Kumar, and S. Tomar, "Structure-function insights into chikungunya virus capsid protein: Small molecules targeting capsid hydrophobic pocket," Virology, vol. 515, pp. 223–234, Feb. 2018.

[50] S. J. Watowich, Y. Xu, A. P. Lemon, and S. Weaver, "Crystal structure of the conserved core domain of Venezualan equine encephalitis capsid protein,"submitted in RCSB PDB, PDB ID: 1EP6, 2003.

[51] L. Lundberg, B. Carey, and K. Kehn-Hall, "Venezuelan equine encephalitis virus capsid-The Clever Caper.," Viruses, vol. 9, no. 10, 2017.

[52] J. M. Reynaud, V. Lulla, D. Y. Kim, E. I. Frolova, and I. Frolov, "The sd1 subdomain of Venezuelan equine encephalitis virus capsid protein plays a critical role in nucleocapsid and particle assembly.," J. Virol., vol. 90, no. 4, pp. 2008–20, 2016.

[53] R. Perera, C. Navaratnarajah, and R. J. Kuhn, "A heterologous coiled coil can substitute for helix I of the Sindbis virus capsid protein.," J. Virol., vol. 77, no. 15, pp. 8345–53, Aug. 2003.

[54] R. Perera, K. E. Owen, T. L. Tellinghuisen, A. E. Gorbalenya, and R. J. Kuhn, "Alphavirus nucleocapsid protein contains a putative coiled coil α-helix important for core assembly," J. Virol., vol. 75, no. 1, p. 1, 2001.

[55] S. Atasheva, N. Garmashova, I. Frolov, and E. Frolova, "Venezuelan equine encephalitis virus capsid protein inhibits nuclear import in mammalian but not in mosquito cells," J. Virol., vol. 82, no. 8, pp. 4028–4041, Apr. 2008.

[56] K. E. Owen and R. J. Kuhn, “Identification of a region in the Sindbis virus nucleocapsid protein that is involved in specificity of RNA encapsidation.,” J. Virol., vol. 70, no. 5, pp. 2757–63, May 1996.
[57] B. R. Linger, L. Kunovska, R. J. Kuhn, and B. L. Golden, “Sindbis virus nucleocapsid assembly: RNA folding promotes capsid protein dimerization.,” RNA, vol. 10, no. 1, pp. 128–38, Jan. 2004.
[58] N. Garmashova, R. Gorchakov, E. Volkova, S. Paessler, E. Frolova, and I. Frolov, “The Old World and New World alphaviruses use different virus-specific proteins for induction of transcriptional shutoff.,” J. Virol., vol. 81, no. 5, pp. 2472–84, Mar. 2007.
[59] N. Garmashova, S. Atasheva, W. Kang, S. C. Weaver, E. Frolova, and I. Frolov, “Analysis of Venezuelan equine encephalitis virus capsid protein function in the inhibition of cellular transcription,” J. Virol., vol. 81, no. 24, pp. 13552–13565, 2007.
[60] J. E. Staples, R. F. Breiman, and A. M. Powers, “Chikungunya Fever: An epidemiological review of a re-emerging infectious disease,” Clin. Infect. Dis., vol. 49, no. 6, pp. 942–948, Sep. 2009.
[61] M. Vazeille, S. Moutailler, F. Pages, F. Jarjaval, and A.-B. Failloux, “Introduction of *Aedes albopictus* in Gabon: what consequences for dengue and chikungunya transmission?,” Trop. Med. Int. Heal., vol. 13, no. 9, pp. 1176–1179, Sep. 2008.
[62] S.-D. Thiberville et al., “Chikungunya fever: Epidemiology, clinical syndrome, pathogenesis and therapy,” Antiviral Res., vol. 99, no. 3, pp. 345–370, Sep. 2013.
[63] W. B. Klimstra, E. M. Nangle, M. S. Smith, A. D. Yurochko, and K. D. Ryman, “DC-SIGN and L-SIGN can act as attachment receptors for alphaviruses and distinguish between mosquito cell- and mammalian cell-derived viruses.,” J. Virol., vol. 77, no. 22, pp. 12022–32, Nov. 2003.
[64] W. B. Klimstra, K. D. Ryman, and R. E. Johnston, “Adaptation of Sindbis virus to BHK cells selects for use of heparan sulfate as an attachment receptor.,” J. Virol., vol. 72, no. 9, pp. 7357–66, Sep. 1998.
[65] M. Kielian, “Membrane fusion and the alphavirus life cycle,” Adv. Virus Res., vol. 45, pp. 113–151, Jan. 1995.
[66] I. Singh’ And and A. Helenius2, “Role of Ribosomes in Semliki Forest Virus Nucleocapsid Uncoating,” 1992.
[67] H. Garoff, M. Sjöberg, and R. H. Cheng, “Budding of alphaviruses,” Virus Res., vol. 106, no. 2, pp. 103–116, Dec. 2004.
[68] J. Y.-S. Leung, M. M.-L. Ng, and J. J. H. Chu, “Replication of alphaviruses: a review on the entry process of alphaviruses into cells,” Adv. Virol., vol. 2011, pp. 1–9, Jul. 2011.
[69] R. Warrier, B. R. Linger, B. L. Golden, and R. J. Kuhn, “Role of Sindbis Virus capsid protein region II in nucleocapsid core assembly and encapsidation of genomic RNA,” J. Virol., vol. 82, no. 9, pp. 4461–4470, 2008.
[70] H. Garoff, M. Sjöberg, and R. H. Cheng, “Budding of alphaviruses,” Virus Res., vol. 106, no. 2, pp. 103–116, Dec. 2004.
[71] J. C. Rupp, K. J. Sokoloski, N. N. Gebhart, and R. W. Hardy, “Alphavirus RNA synthesis and non-structural protein functions,” J. Gen. Virol., vol. 96, no. Pt 9, p. 2483, Sep. 2015.
[72] E. G. Strauss, C. M. Rice, and J. H. Strauss, “Complete nucleotide sequence of the genomic RNA of Sindbis virus.,” Virology, vol. 133, pp. 92–110, Feb. 1984.
[73] J. A. Lemm, A. Bergqvist, C. M. Read, and C. M. Rice, “Template-dependent initiation of sindbis virus rna replication *in vitro*,” Journal of Virol, vol. 72, pp- 46-53, Aug 1998.
[74] J.-H. Ou, E. G. Strauss, J. H. Strauss, and K. Simons, “The 5′-terminal sequences of the genomic RNAs of several alphaviruses,” J. Mol. Biol., vol. 168, no. 1, pp. 1–15, Jul. 1983.

[75] I. Frolov, R. Hardy, and C. M. Rice, “Cis-acting RNA elements at the 5’ end of Sindbis virus genome RNA regulate minus- and plus-strand RNA synthesis.,” RNA, vol. 7, no. 11, pp. 1638–51, Nov. 2001.

[76] P. Pushko, M. Parker, G. V. Ludwig, N. L. Davis, R. E. Johnston, and J. F. Smith, “Replicon-helper systems from attenuated Venezuelan equine encephalitis virus: expression of heterologous genesin vitroand immunization against heterologous pathogens *in vivo*,” Virology, vol. 239, no. 2, pp. 389–401, Dec. 1997.

[77] R. J. Kuhn, Z. Hong, and J. H. Strauss, “Mutagenesis of the 3' nontranslated region of Sindbis virus RNA.,” J. Virol., vol. 64, no. 4, pp. 1465–76, Apr. 1990.

[78] R. W. Hardy and C. M. Rice, “Requirements at the 3' end of the Sindbis virus genome for efficient synthesis of minus-strand RNA,” J. Virol., vol. 79, no. 8, pp. 4630–4639, 2005.

[79] R. W. Hardy, “The role of the 3′ terminus of the Sindbis virus genome in minus-strand initiation site selection,” Virology, vol. 345, no. 2, pp. 520–531, Feb. 2006.

[80] R. J. de Groot, W. R. Hardy, Y. Shirako, and J. H. Strauss, “Cleavage-site preferences of Sindbis virus polyproteins containing the non-structural proteinase. Evidence for temporal regulation of polyprotein processing in vivo.,” EMBO J., vol. 9, no. 8, pp. 2631–8, Aug. 1990.

[81] Y. Shirako and J. H. Strauss, “Cleavage between nsP1 and nsP2 initiates the processing pathway of Sindbis virus nonstructural polyprotein P123,” Virology, vol. 177, no. 1, pp. 54–64, Jul. 1990.

[82] W. Reef and J. H. Strauss, “Processing the nonstructural polyproteins of sindbis virus: nonstructural proteinase is in the C-terminal half of nsp2 and functions both in cis and in trans.”, J. Virol., vol.63, no.11, pp.4563-64, Nov 1989.

[83] J. C. Rupp, K. J. Sokoloski, N. N. Gebhart, and R. W. Hardy, “Alphavirus RNA synthesis and non-structural protein functions.,” J. Gen. Virol., vol. 96, no. 9, pp. 2483–500, Sep. 2015.

[84] S. Mi and V. Stollar, “Expression of Sindbis virus nsP1 and methyltransferase activity in *Escherichia coli*.,” Virology, vol. 184, no. 1, pp. 423–7, Sep. 1991.

[85] T. Ahola and L. Kääriäinen, “Reaction in alphavirus mRNA capping: formation of a covalent complex of nonstructural protein nsP1 with 7-methyl-GMP.,” Proc. Natl. Acad. Sci. U. S. A., vol. 92, no. 2, pp. 507–11, Jan. 1995.

[86] S. Tomar, M. Narwal, E. Harms, J. L. Smith, and R. J. Kuhn, “Heterologous production, purification and characterization of enzymatically active Sindbis virus nonstructural protein nsP1,” Protein Expr. Purif., vol. 79 (2), pp. 277–284, Oct. 2011.

[87] W. Aouadi et al., “Toward the identification of viral cap-methyltransferase inhibitors by fluorescence screening assay.,” Antiviral Res., vol. 144, pp. 330–339, 2017.

[88] J. Peränen, P. Laakkonen, M. Hyvönen, and L. Kääriäinen, “The Alphavirus replicase protein nsp1 is membrane-associated and has affinity to endocytic organelles,” Virology, vol. 208 (2), pp. 610–620, Apr. 1995.

[89] Y.-F. Wang, S. G. Sawicki, and D. L. Sawicki, “Sindbis virus nspl functions in negative-strand RNA synthesis,” J. Virol, vol. 65(2), pp. 985-8, Feb1991.

[90] A. T. Russo, M. A. White, and S. J. Watowich, “The Crystal Structure of the Venezuelan Equine Encephalitis Alphavirus nsP2 Protease,” Structure, vol. 14, no. 9, pp. 1449–1458, Sep. 2006.

[91] M. Narwal et al., “Crystal structure of chikungunya virus nsP2 cysteine protease reveals a putative flexible loop blocking its active site,” Int. J. Biol. Macromol., vol. 116, pp. 451–462, Sep. 2018.

[92] S. Dhindwal, P. Kesari, H. Singh, P. Kumar, and S. Tomar, “Conformer and pharmacophore based identification of peptidomimetic inhibitors of chikungunya virus nsP2 protease,” J. Biomol. Struct. Dyn., vol. 35, no. 16, pp. 3522–3539, Dec. 2017.

[93] H. Singh et al., “Chikungunya virus inhibition by peptidomimetic inhibitors targeting virus-specific cysteine protease,” Biochimie, vol. 149, pp. 51–61, Jun. 2018.

[94] Y. F. Wang, S. G. Sawicki, and D. L. Sawicki, “Alphavirus nsP3 functions to form replication complexes transcribing negative-strand RNA.,” J. Virol., vol. 68, no. 10, pp. 6466–75, Oct. 1994.

[95] L. Eckei et al., “The conserved macrodomains of the non-structural proteins of chikungunya virus and other pathogenic positive strand RNA viruses function as mono-ADP-ribosylhydrolases,” Sci. Rep., vol. 7, no. 1, p. 41746, Mar. 2017.

[96] H. Malet et al., “The crystal structures of chikungunya and venezuelan equine encephalitis virus nsp3 macro domains define a conserved adenosine binding pocket,” J. Virol., vol. 83, no. 13, pp. 6534–6545, Jul. 2009.

[97] M. D. Panas, T. Ahola, and G. M. McInerney, “The C-terminal repeat domains of nsp3 from the old world alphaviruses bind directly to g3bp,” J. Virol., vol. 88, no. 10, pp. 5888–5893, May 2014.

[98] Y. Gao, N. Goonawardane, J. Ward, A. Tuplin, and M. Harris, “Multiple roles of the non-structural protein 3 (nsP3) alphavirus unique domain (AUD) during chikungunya virus genome replication and transcription,” PLOS Pathog., vol. 15, no. 1, p. e1007239, Jan. 2019.

[99] M. W. LaStarza, J. A. Lemm, and C. M. Rice, “Genetic analysis of the nsP3 region of Sindbis virus: evidence for roles in minus-strand and subgenomic RNA synthesis.,” J. Virol., vol. 68, no. 9, pp. 5781–91, Sep. 1994.

[100] J. K. Rubach, B. R. Wasik, J. C. Rupp, R. J. Kuhn, R. W. Hardy, and J. L. Smith, “Characterization of purified Sindbis virus nsP4 RNA-dependent RNA polymerase activity *in vitro*,” Virology, vol. 384, no. 1, pp. 201–208, Feb. 2009.

[101] E. K. O’reilly and C. Cheng Kao, “Minireview analysis of rna-dependent rna polymerase structure and function as guided by known polymerase structures and computer predictions of secondary structure,” Virol., vol. 252 (2), pp. 287-303, Dec 1998.

[102] S. Tomar, R. W. Hardy, J. L. Smith, and R. J. Kuhn, “Catalytic core of alphavirus nonstructural protein nsp4 possesses terminal adenylyltransferase activity,” J. Virol., vol. 80, no. 20, pp. 9962–9969, Oct. 2006.

[103] L. Carrasco, “Modification of membrane permeability by animal viruses,” Adv. Virus Res., vol. 45, pp. 61–112, Jan. 1995.

[104] P. Melancon, H. Garoff, and R. J. Kuhn, “Processing of the Semliki Forest virus structural polyprotein: role of the capsid protease.,” J. Virol., vol. 61, no. 5, pp. 1301–9, May 1987.

[105] S. K. Jain, S. DeCandido, and M. Kielian, “Processing of the p62 envelope precursor protein of Semliki Forest virus.,” J. Biol. Chem., vol. 266, no. 9, pp. 5756–61, Mar. 1991.

[106] M. Aggarwal et al., “Crystal structure of aura virus capsid protease and its complex with dioxane: new insights into capsid-glycoprotein molecular contacts,” PLoS One, vol. 7, no. 12, p. e51288, Dec. 2012.

[107] E. M. Hong, R. Perera, and R. J. Kuhn, “Alphavirus capsid protein helix i controls a checkpoint in nucleocapsid core assembly,” J. Virol., vol. 80, no. 18, pp. 8848–8855, 2006.

[108] K. E. Owen and R. J. Kuhn, “Alphavirus budding is dependent on the interaction between the nucleocapsid and hydrophobic amino acids on the cytoplasmic domain of the e2 envelope glycoprotein,” Virology, vol. 230, no. 2, pp. 187–196, Apr. 1997.

[109] H. Garoff and K. Simons, “Location of the spike glycoproteins in the Semliki Forest virus membrane.,” Proc. Natl. Acad. Sci. U. S. A., vol. 71, no. 10, pp. 3988–92, Oct. 1974.

[110] M. Lobigs, H. X. Zhao, and H. Garoff, “Function of Semliki Forest virus E3 peptide in virus assembly: replacement of E3 with an artificial signal peptide abolishes spike heterodimerization and surface expression of E1.,” J. Virol., vol. 64, no. 9, pp. 4346–55, Sep. 1990.

[111] K. D. Ryman, W. B. Klimstra, and R. E. Johnston, “Attenuation of Sindbis virus variants incorporating uncleaved PE2 glycoprotein is correlated with attachment to cell-surface heparan sulfate,” Virology, vol. 322, no. 1, pp. 1–12, Apr. 2004.

[112] S. Mukhopadhyay et al., “Mapping the structure and function of the e1 and e2 glycoproteins in alphaviruses,” Structure, vol. 14, no. 1, pp. 63–73, Jan. 2006.

[113] P. Soonsawad et al., “Structural evidence of glycoprotein assembly in cellular membrane compartments prior to alphavirus budding.,” J. Virol., vol. 84, no. 21, pp. 11145–51, Nov. 2010.

[114] C. H. von Bonsdorff and S. C. Harrison, “Hexagonal glycoprotein arrays from Sindbis virus membranes.,” J. Virol., vol. 28, no. 2, p. 578, 1978.

[115] R. L. Knight, K. L. W. Schultz, R. J. Kent, M. Venkatesan, and D. E. Griffin, “Role of N-linked glycosylation for Sindbis virus infection and replication in vertebrate and invertebrate systems.,” J. Virol., vol. 83, no. 11, pp. 5640–7, Jun. 2009.

[116] K. D. Ryman, C. L. Gardner, C. W. Burke, K. C. Meier, J. M. Thompson, and W. B. Klimstra, “Heparan sulfate binding can contribute to the neurovirulence of neuroadapted and nonneuroadapted Sindbis viruses,” J. Virol., vol. 81, no. 7, pp. 3563–3573, Apr. 2007.

[117] W. J. Welch and B. M. Sefton, “Characterization of a small, nonstructural viral polypeptide present late during infection of BHK cells by Semliki Forest virus.,” J. Virol., vol. 33, no. 1, pp. 230–7, Jan. 1980.

[118] P. Liljeström and H. Garoff, “Internally located cleavable signal sequences direct the formation of Semliki Forest virus membrane proteins from a polyprotein precursor.,” J. Virol., vol. 65, no. 1, pp. 147–54, Jan. 1991.

[119] K. Gaedigk-Nitschko, M. Ding, M. A. Levy, and M. J. Schlesinger, “Site-directed mutations in the Sindbis virus 6K protein reveal sites for fatty acylation and the underacylated protein affects virus release and virion structure,” Virology, vol. 175, no. 1, pp. 282–291, Mar. 1990.

[120] G. M. McInerney, J. M. Smit, P. Liljeström, and J. Wilschut, “Semliki Forest virus produced in the absence of the 6K protein has an altered spike structure as revealed by decreased membrane fusion capacity,” Virology, vol. 325, no. 2, pp. 200–206, Aug. 2004.

[121] J. H. Strauss and E. G. Strauss, “Virus Evolution: How does an enveloped virus make a regular structure?,” Cell, vol. 105, no. 1, pp. 5–8, Apr. 2001.

[122] J. E. Snyder et al., “Functional characterization of the alphavirus TF protein.,” J. Virol., vol. 87, no. 15, pp. 8511–23, Aug. 2013.

[123] U. Skoging and P. Liljeström, “Role of the C-terminal tryptophan residue for the structure-function of the alphavirus capsid protein,” J. Mol. Biol., vol. 279, no. 4, pp. 865–872, Jun. 1998.

[124] S. Lee, R. J. Kuhn, and M. G. Rossmann, “Probing the potential glycoprotein binding site of sindbis virus capsid protein with dioxane and model building.,” Proteins, vol. 33, no. 2, pp. 311–7, Nov. 1998.

[125] S. Thomas et al., “Functional dissection of the alphavirus capsid protease: sequence requirements for activity,” Virol. J., vol. 7, no. 1, p. 327, Dec. 2010.

[126] M. Morillas, H. Eberl, F.-T. Allain, R. Glockshuber, and E. Kuennemann, “Novel enzymatic activity derived from the semliki forest virus capsid protein,” J. Mol. Biol., vol. 376, no. 3, pp. 721–735, Feb. 2008.

[127] A. Munjal et al., “Advances in developing therapies to combat zika virus: current knowledge and future perspectives,” Front. Microbiol., vol. 8, p. 1469, Aug. 2017.

[128] P. Agnihotri, A. K. Mishra, S. Mishra, V. K. Sirohi, A. A. Sahasrabuddhe, and J. V. Pratap, “Identification of novel inhibitors of leishmania donovani γ-glutamylcysteine synthetase using structure-based virtual screening, docking, molecular dynamics simulation, and *in vitro* studies,” J. Chem. Inf. Model., vol. 57, no. 4, pp. 815–825, Apr. 2017.

[129] M. I. Hassan, V. Kumar, R. K. Somvanshi, S. Dey, T. P. Singh, and S. Yadav, “Structure-guided design of peptidic ligand for human prostate specific antigen,” J. Pept. Sci., vol. 13, no. 12, pp. 849–855, Dec. 2007.

[130] N. Chhikara, M. Saraswat, A. K. Tomar, S. Dey, S. Singh, and S. Yadav, “Human Epididymis Protein-4 (HE-4): A novel cross-class protease inhibitor,” PLoS One, vol. 7, no. 11, p. e47672, Nov. 2012.

[131] J. Kumari, S. Dhingra, and J. Kumar, “Fluorescence-Based Screening of Membrane Proteins for Structural Studies,” 2016, pp. 141–155.

[132] F. Simon, E. Javelle, M. Oliver, I. Leparc-Goffart, and C. Marimoutou, “Chikungunya Virus Infection,” Curr. Infect. Dis. Rep., vol. 13, no. 3, pp. 218–228, Jun. 2011.

[133] E. S. Gudo, J. F. P. Black, and J. L. Cliff, “Chikungunya in Mozambique: A forgotten history,” PLoS Negl. Trop. Dis., vol. 10, no. 11, p. e0005001, Nov. 2016.

[134] S. C. Weaver and M. Lecuit, “Chikungunya virus and the global spread of a mosquito-borne disease,” N. Engl. J. Med., vol. 372, no. 13, pp. 1231–1239, Mar. 2015.

[135] A. Vega-Rua, K. Zouache, R. Girod, A.-B. Failloux, and R. Lourenco-de-Oliveira, “High level of vector competence of aedes aegypti and aedes albopictus from ten american countries as a crucial factor in the spread of chikungunya virus,” J. Virol., vol. 88, no. 11, pp. 6294–6306, Jun. 2014.

[136] P. J. Hotez et al., “The global burden of disease study 2010: interpretation and implications for the neglected tropical diseases,” PLoS Negl. Trop. Dis., vol. 8, no. 7, p. e2865, Jul. 2014.

[137] P. J. Hotez and A. Kamath, “Neglected tropical diseases in sub-saharan africa: review of their prevalence, distribution, and disease burden,” PLoS Negl. Trop. Dis., vol. 3, no. 8, p. e412, Aug. 2009.

[138] M. Solignat, B. Gay, S. Higgs, L. Briant, and C. Devaux, “Replication cycle of chikungunya: A re-emerging arbovirus,” Virology, vol. 393, no. 2, pp. 183–197, Oct. 2009.

[139] S. Lee et al., “Identification of a protein binding site on the surface of the alphavirus nucleocapsid and its implication in virus assembly.,” Structure, vol. 4, no. 5, pp. 531–41, May 1996.

[140] U. Skoging and P. Liljeström, “Role of the C-terminal tryptophan residue for the structure-function of the alphavirus capsid protein,” J. Mol. Biol., vol. 279, no. 4, pp. 865–872, Jun. 1998.

[141] K. Dhama et al., “Advances in designing and developing vaccines, drugs, and therapies to counter Ebola virus.,” Front. Immunol., vol. 9, p. 1803, 2018.

[142] A. Bitra, A. Biswas, and R. Anand, “Structural basis of the substrate specificity of cytidine deaminase superfamily guanine deaminase,” Biochemistry, vol. 52, no. 45, pp. 8106–8114, Nov. 2013.

[143] Z. Zhao, C. Martin, R. Fan, P. E. Bourne, and L. Xie, “Drug repurposing to target Ebola virus replication and virulence using structural systems pharmacology,” BMC Bioinformatics, vol. 17, no. 1, p. 90, Dec. 2016.

[144] P. Emsley and K. Cowtan, “Coot : model-building tools for molecular graphics,” Acta Crystallogr. Sect. D Biol. Crystallogr., vol. 60, no. 12, pp. 2126–2132, Dec. 2004.

[145] W. L. Delano, “PyMOL: An Open-Source Molecular Graphics Tool.”

[146] A. C. Wallace, R. A. Laskowski, and J. M. Thornton, “LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions.,” Protein Eng., vol. 8, no. 2, pp. 127–34, Feb. 1995.

[147] T. W. Geders, J. L. Smith, and R. J. Kuhn, “Role for conserved residues of sindbis virus nonstructural protein 2 methyltransferase-like domain in regulation of minus-strand synthesis and development of cytopathic infection †,” J. Virol., vol. 82, no. 15, pp. 7284–7297, 2008.

[148] S. M. M. Alsanosi, C. Skiffington, and S. Padmanabhan, “Pharmacokinetic pharmacogenomics,” in handbook of pharmacogenomics and stratified medicine, Elsevier, 2014, pp. 341–364.

[149] E. De Clercq and G. Li, “Approved antiviral drugs over the past 50 years,” Clin. Microbiol. Rev., vol. 29, no. 3, pp. 695–747, Jul. 2016.

[150] E. M. Hong, R. Perera, and R. J. Kuhn, “Alphavirus capsid protein helix i controls a checkpoint in nucleocapsid core assembly,” J. Virol., vol. 80, no. 18, pp. 8848–8855, Sep. 2006.

[151] R. Perera, K. E. Owen, T. L. Tellinghuisen, A. E. Gorbalenya, and R. J. Kuhn, “Alphavirus nucleocapsid protein contains a putative coiled coil -helix important for core assembly,” J. Virol., vol. 75, no. 1, pp. 1–10, Jan. 2001.

[152] H.-K. Choi et al., “Structural analysis of sindbis virus capsid mutants involving assembly and catalysis,” J. Mol. Biol., vol. 262, no. 2, pp. 151–167, Sep. 1996.

[153] G. Wengler, U. Boege, G. Wengler, H. Bischoff, and K. Wahn, “The core protein of the alphavirus Sindbis virus assembles into core-like nucleoproteins with the viral genome RNA and with other single-stranded nucleic acids in vitro.,” Virology, vol. 118, no. 2, pp. 401–10, Apr. 1982.

[154] B. Weiss, U. Geigenmüller-Gnirke, and S. Schlesinger, “Interactions between Sindbis virus RNAs and a 68 amino acid derivative of the viral capsid protein further defines the capsid binding site.,” Nucleic Acids Res., vol. 22, no. 5, pp. 780–6, Mar. 1994.

[155] B. Weiss, H. Nitschko, I. Ghattas, R. Wright, and S. Schlesinger, “Evidence for specificity in the encapsidation of Sindbis virus RNAs.,” J. Virol., vol. 63, no. 12, pp. 5310–8, Dec. 1989.

[156] T. L. Tellinghuisen, R. Perera, and R. J. Kuhn, “In vitro Assembly of Sindbis Virus Core-Like Particles from Cross-Linked Dimers of Truncated and Mutant Capsid Proteins,” J. Virol., vol. 75, no. 6, pp. 2810–2817, 2001.

[157] A. Vagin and A. Teplyakov, “Molecular replacement with MOLREP,” Acta Crystallogr. Sect. D Biol. Crystallogr., vol. 66, no. 1, pp. 22–25, Jan. 2010.

[158] P. D. Adams et al., “PHENIX: a comprehensive Python-based system for macromolecular structure solution.,” Acta Crystallogr. D. Biol. Crystallogr., vol. 66, no. Pt 2, pp. 213–21, Feb. 2010.

[159] M. S. Bronze, M. M. Huycke, L. J. Machado, G. W. Voskuhl, and R. A. Greenfield, “Viral agents as biological weapons and agents of bioterrorism,” Am. J. Med. Sci., vol. 323, no. 6, pp. 316–325, Jun. 2002.

[160] P. Gérardin et al., “Neurocognitive outcome of children exposed to perinatal mother-to-child chikungunya virus infection: the chimere cohort study on reunion island,” PLoS Negl. Trop. Dis., vol. 8, no. 7, p. e2996, Jul. 2014.

[161] M. Kielian, C. Chanel-Vos, and M. Liao, “Alphavirus entry and membrane fusion.,” Viruses, vol. 2, no. 4, pp. 796–825, Mar. 2010.

[162] T. L. Tellinghuisen, A. E. Hamburger, B. R. Fisher, R. Ostendorp, and R. J. Kuhn, “In vitro assembly of alphavirus cores by using nucleocapsid protein expressed in *Escherichia coli*,” 1999.

[163] B. Weiss, U. Geigenmüller-Gnirke, and S. Schlesinger, “Interactions between Sindbis virus RNAs and a 68 amino acid derivative of the viral capsid protein further defines the capsid binding site.,” Nucleic Acids Res., vol. 22, no. 5, pp. 780–6, Mar. 1994.

[164] L. Tong, G. Wengler, and M. G. Rossmann, “Refined structure of Sindbis virus core protein and comparison with other chymotrypsin-like serine proteinase structures,” J. Mol. Biol., vol. 230, no. 1, pp. 228–247, Mar. 1993.

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE