Understanding the viral protein structures of clinically relevant arboviruses and fragment-based screening for antivirals

Farhana Abu Bakar

2019

Farhana Abu Bakar. (2020). Understanding the viral protein structures of clinically relevant arboviruses and fragment-based screening for antivirals. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/142588>

<https://doi.org/10.32657/10356/142588>

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0).

Downloaded on 26 Apr 2023 23:11:10 SGT

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

SINGAPORE

**UNDERSTANDING THE VIRAL PROTEIN
STRUCTURES OF CLINICALLY RELEVANT
ARBOVIRUSES AND FRAGMENT-BASED
SCREENING FOR ANTIVIRALS**

FARHANA BINTE ABU BAKAR

SCHOOL OF BIOLOGICAL SCIENCES

2019

**Understanding the viral protein structures of
clinically relevant arboviruses and fragment-
based screening for antivirals**

Farhana Binte Abu Bakar

SCHOOL OF BIOLOGICAL SCIENCES

A thesis submitted to the Nanyang
Technological University in partial fulfilment
of the requirement for the degree of Doctor
of Philosophy

2019

Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research done by me except where otherwise stated in this thesis. The thesis work has not been submitted for a degree or professional qualification to any other university or institution. I declare that this thesis is written by myself and is free of plagiarism and of sufficient grammatical clarity to be examined. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

24 September 2019

.....
Date

.....
Farhana Binte Abu Bakar

Supervisor Declaration Statement

I have reviewed the content and presentation style of this thesis and declare it of sufficient grammatical clarity to be examined. To the best of my knowledge, the thesis is free of plagiarism and the research and writing are those of the candidate's except as acknowledged in the Author Attribution Statement. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

24 September 2019

.....
Date

.....
A/Prof. Julien Lescar

Authorship Attribution Statement

This thesis contains material from 1 number paper published in the following peer-reviewed journal(s) / from papers accepted at conferences in which I am listed as an author.

Part of the introduction in Chapter 3 is published as Bakar, F. A. and Ng, L. F. P. (2018) 'Nonstructural proteins of alphavirus —potential targets for drug development', *Viruses*, 10(2), pp. 1–15.

The contributions of the co-authors are as follows:

- I wrote the manuscript drafts with Prof Lisa F.P. Ng. The manuscript was proofread by Dr Eng Kai Er.

24 September 2019

.....
Date

.....
Farhana Binte Abu Bakar

ACKNOWLEDGEMENTS

I would like to take this opportunity to firstly express my deepest gratitude to my two supervisors, A/Prof Julien Lescar in NTU and Prof Lisa Ng in SIgN for their continued support and encouragement throughout the four years. It has been a tough experience but their guidance and patience have greatly helped me through the challenges and made this journey a fulfilling one. My heartfelt gratitude also extends to Prof Laurent Renia, who gave me the encouragement and recommendation to embark on this journey. And to my Thesis Advisory Committee members, A/Prof Konstantin Pervushin and Dr Katja Fink, who have provided useful ideas and important suggestions that have greatly improved the quality of my thesis work.

I would like to thank the current members of the JL team in NTU - Dr El Sahili Abbas, Dr Qian Xinlei, Dr Goh Boon Chong, Ms Wong Yee Hwa, Mr Lin Jianqing, Ms Jia Huan and not forgetting the past members – Ms Swathi Pasunooti, Ms Grace Chan, Mr Liu Chong, and Mr Yeo Tiong Han for their support as we worked together in a team. I am also grateful to the current members of the LN team in SIgN - Dr Carissimo Guillaume, Dr Fong Siew Wai, Dr Poh Chek Meng, Mr Chan Yi Hao, Ms Cheryl Lee, Mr Anthony Torres, Ms Siti Naqiah Amrun, Ms Rhonda Chee and Mr Nicholas Yeo. It has been a privilege to cross paths with everyone and I am grateful for having had the chance to work with such collaborative and creative individuals. It was a privilege to work alongside many of these enthusiastic scientists and I truly enjoyed my stay in NTU and SIgN.

Finally, I would like to thank my husband, Nor Azman, and my family, whose love and devotion are without limit. I am eternally grateful for their encouragement and support throughout every step I have taken in life. It is purely having them in my life that makes anything else matter. Thank you for all you have given me.

TABLE OF CONTENTS

Statement of Originality	i
Supervisor Declaration Statement	ii
Authorship Attribution Statement	iii
Acknowledgements	iv
Summary	xii
List of Publications	xiv
List of Abbreviations	xvi
List of Tables	xix
List of Figures	xx
Chapter 1 Introduction	
1.1 Introduction to Chikungunya virus (CHIKV)	1
1.1.1 General overview of Alphaviruses	1
1.1.2 Epidemiology and evolution of CHIKV	2
1.1.3 Characteristics of Chikungunya fever	3
1.1.4 CHIKV pathogenesis	3
1.1.5 CHIKV genome and replication cycle	6
1.1.6 Interplay of viral and host factors during CHIKV infection	7
1.1.7 O'nyong'nyong virus, a close relative of CHIKV	13
1.1.8 Functions of nonstructural proteins (nsPs)	17
1.2 Available structures of nonstructural proteins (nsPs)	21
1.3 Current strategies for inhibition of CHIKV infection	26
1.3.1 Drugs inhibiting viral entry	27
1.3.2 Inhibitors targeting viral genome replication	28
1.3.3 Non-specific inhibitors	29

1.3.4	Antivirals against nsP2 protease	30
1.3.5	Treatment strategies targeting CHIKV-induced inflammation	32
1.3.6	Current limitations	35
1.4	Aims and approaches	36
Chapter 2 Materials and Methods		
2.1	Materials	39
2.1.1	Primers	39
2.1.2	Fragment library and compounds	39
2.1.3	Reagents for bacteria culture	39
2.1.3.1	Antibiotics	39
2.1.3.2	Media	40
2.1.4	Protein extraction materials	40
2.1.4.1	Buffers	40
2.1.4.2	Chromatography resin / columns	41
2.1.5	Crystallization kits	41
2.1.6	Peptide substrates	41
2.1.7	Cell lines	42
2.1.8	Virus stocks	42
2.2	Methods	44
2.2.1	Cloning of alphavirus nsP2 for <i>Escherichia coli</i> expression	44
2.2.2	Expression and purification of nsP2 proteins	45
2.2.3	SDS-PAGE	46
2.2.4	NTPase activity assay	46

2.2.5	Cell-free FRET-based protease assay	47
2.2.6	Fluorescence Extinction Coefficient determination	47
2.2.7	Measurement of steady-state kinetic parameters	48
2.2.8	Compound screening on CHIKV nsP2 protease	49
2.2.9	Thermal shift assay	49
2.2.10	IC ₅₀ determination	50
2.2.11	<i>In vitro</i> antiviral assay	50
2.2.12	Cell viability assay	51
2.2.13	Murine model of CHIKV-induced joint footpad inflammation	51
2.2.14	Viral RNA extraction	52
2.2.15	CHIKV viral RNA quantification	52
2.2.16	Protein crystallization screen	53
2.2.17	Crystal soaking and freezing	54
2.2.18	X-ray data collection and structure refinement	54
2.2.19	Small angled X-ray scattering and data analysis	54
2.2.20	Negative staining and electron microscopy image analysis	55
2.2.21	Statistical analyses	56

Chapter 3 Functional and Structural Characterization of Alphavirus nsP2

3.1	Introduction	57
3.1.1	Importance of N-terminal nsP2 in 2/3 site processing	57
3.1.2	Importance of revealing nsP2 structure of closely-related alphaviruses	58

3.1.3	Structural biology techniques to elucidate viral protein structures	61
3.1.4.	Study approach	61
3.2	Results	63
3.2.1	Generation of nsP2 protein constructs	63
3.2.2	Validation of nsP2 protease and ATPase activities of purified nsP2 proteins	66
3.2.3	Low-resolution structure of CHIKV-vc038	72
3.2.4	Purification, crystallization and data collection of ONNV-vc002	79
3.2.5	Crystal structure of ONNV-vc002	86
3.3	Discussion	91
Chapter 4 Identification of Small Molecules as Alphavirus Protease Inhibitors		
4.1	Introduction	94
4.1.1	Fragment to lead optimization	94
4.1.2	Study Approach	100
4.2	Results	101
4.2.1	Expression and purification of CHIKV, ONNV and VEEV nsP2FL proteins	101
4.2.2	Determination of Fluorescence Extinction Coefficient of CHIKV and VEEV nsP1/2 peptide substrates	106
4.2.3	Fragment-based screening against CHIKV nsP2FL	108
4.2.4	NTPase activity not affected by compound treatment	115

4.2.5	Anti-protease activity of CHIKV primary hits are dose-dependent	117
4.2.6	Similar inhibitory effect on nsP2 protease of ONNV and VEEV	117
4.2.7	Effect of combinatorial treatment on protease inhibition	122
4.2.8	Amine at meta position of Compound 2 is crucial to its bioactivity	124
4.2.9	Kinetic analysis of inhibition by Compound 2	126
4.2.10	Compound 2 destabilizes CHIKV nsP2 protease	132
4.2.11	<i>In vitro</i> antiviral efficacy against CHIKV isolates	136
4.2.12	Antiviral efficacy on other pathogenic alphaviruses	137
4.2.13	Efficacy of compounds <i>in vivo</i>	143
4.2.14	Crystallization attempts of CHIKV protease-Compound 2 complex	146
4.3	Discussion	149
Chapter 5 Conclusions and future perspectives		155
References		159
Appendix		195
1.	Constructs list	195
2.	Primers list	196
3.	Peptide substrate list	197
4.	Sequence alignment of CHIKV, SINV and VEEV nsP2 proteases	198
5.	Sequence alignment of ONNV-vc002 with CHIKV nsP2	

protease	199
6. Sequence alignment of CHIKV, ONNV and VEEV full length nsP2	200
7. Sequence alignment of SGP11, CNR20235, LR2006-OPY1 and IND91 nsP2 proteases	201
8. Publication	202

SUMMARY (410 words)

Arboviruses are viruses that are transmitted to the human hosts by virus-infected arthropods. In the past decade, the unprecedented emergence of arboviral diseases transmitted by *Aedes* mosquitoes have caused debilitating illnesses that has affected millions of people worldwide. Chikungunya virus (CHIKV) is an arthritogenic alphavirus that causes a debilitating musculoskeletal inflammatory disease that could persist for months to years after the virus is cleared. To date, there are no specific therapeutics available to prevent or mitigate the disease.

In recent years, antiviral drug discovery has increasingly relied on the use of high-resolution structural information of important viral proteins and have led to the development of potent inhibitors. CHIKV nonstructural protein 2 (nsP2) protease structure is routinely targeted due to its essential polyprotein processing function and although the crystal structure of CHIKV nsP2 helicase has recently been determined, the dynamics between the two domains and overall structure of nsP2 remains unclear. Thus, the first objective of this thesis was to gain structural and functional insights into the global architecture and dynamics of the full length nsP2 by integrating several techniques including Small-angle X-Ray Scattering (SAXS), negative stain electron microscopy (EM) and X-ray crystallography as well as biochemical activity assays. This could give a more holistic view of the protein's function and mechanism, and open new avenues for antiviral development in the future. Another arthritogenic alphavirus, O'nyong'nyong virus (ONNV), most closely related to CHIKV, has been

proposed to share similar nsP2 protease structural folds due to its high sequence homology, but this has never been shown. Therefore, another aim of this study was to reveal its structure through X-ray crystallography.

The second objective of this thesis was to discover new chemical motifs that could effectively target alphavirus nsP2 protease. Inhibition of this protein was frequently found to be effective in suppressing viral replication *in vitro*, while the efficacy is limited by drug dissemination to viral replication sites *in vivo*. Herein, fragment-based screening was applied to CHIKV nsP2 protease to identify novel start points for drug design. Biochemical and biophysical assays were employed to determine the inhibition modes of hits and *in vitro* antiviral assays were performed to investigate their efficacies. Another aspect of this study was to evaluate the fragment-protein complex structure and the information could be useful for the optimization of these hits into high affinity leads in the future.

Overall, results obtained from this thesis could be used as a platform for the design of novel therapeutics specifically targeting alphavirus nsP2.

LIST OF PUBLICATIONS

Below is a list of publications currently in preparation:

1. **Abu Bakar F**, Chua TK, Amrun SN, Carissimo G, Lescar J, Ng LF. “Small fragments potentially target Chikungunya virus nsP2 protease”. (2019) (*to be submitted*)
2. **Abu Bakar F**, El Sahili A, Lin JQ, Amrun SN, Carissimo G, Lescar J, Ng LF. “Structure of O’nyong’nyong nonstructural protein 2 protease”. (2019) (*in preparation*)

Publications not included in this thesis:

1. Carissimo G, Chan YH, Utt A, Chua TK, **Abu Bakar F**, Merits A, Ng LF. “VCP/p97 is a proviral host factor for replication of chikungunya virus and other alphaviruses”. *Front. Microbiol.* 10:2236. (2019)
2. Chan YH, Teo TH, Utt A, Tan JJ, Amrun SN, **Abu Bakar F**, Yee WX, Becht E, Lee CY, Lee B, Rajarethinam R, Newell E, Merits A, Carissimo G, Lum FM, Ng LF. “Mutating chikungunya virus non-structural protein produces potent live-attenuated vaccine candidate”. *EMBO Mol Med.* Jun;11(6). (2019)
3. Amrun SN, Yee WX, **Abu Bakar F**, Lee B, Kam YW, Lum FM, Tan JJ, Lim VW, Watthanaworawit W, Ling C, Nosten F, Renia L, Leo YS, Ng LF. “Novel differential linear B-cell epitopes to identify Zika and dengue virus infections in patients”. *Clin Transl Immunology.* Jul 26;8(7):e1066. (2019)

4. Kam YW, Leite JA, Amrun SN, Lum FM, Yee WX, **Bakar FA**, Eng KE, Lye DC, Leo YS, Chong CY, Freitas ARR, Milanez GP, Proença-Modena JL, Rénia L, Costa FTM, Ng LFP; Zika-Unicamp Network. “ZIKV-Specific NS1 Epitopes as Serological Markers of Acute Zika Virus Infection”. *J Infect Dis*. Jun 19;220(2):203-212. (2019)

LIST OF ABBREVIATIONS

ATP	Adenosine triphosphate
ATPase	Adenosine triphosphatase
BSA	Bovine serum albumin
CHIKV	Chikungunya virus
DABCYL	4-[[4-(dimethylamino) phenyl]-azo} benzoic acid
DENV	Dengue virus
DMSO	Dimethyl sulfoxide
DTT	Dithiothreitol
dRFU/dT	Change in fluorescence level with respect to per unit change in temperature
<i>E. coli</i>	<i>Escherichia coli</i>
EC ₅₀	Half-maximal effective concentration
ECSA	Eastern, central, southern African
EDANS	5-[(2'-aminoethyl)-amino] naphthalenesulfonic acid
FRET	Fluorescence Resonance Energy Transfer
Gluc	<i>Gaussia princeps</i> luciferase
GTP	Guanosine triphosphate
IC ₅₀	Half-maximal inhibitory concentration
IPTG	Isopropyl-β-D-thiogalactopyranoside
Kb	Kilobase
kDa	Kilodalton
MOI	Multiplicity of infection
MWCO	Molecular weight cutoff
nsP	Nonstructural protein

NTP	Nucleotide triphosphate
NTPase	Nucleoside triphosphatase
OD	Optical density
ONNV	O'nyong'nyong virus
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate-buffered saline
PBST	PBS with Tween 20
PDB	Protein data bank
PEG	Polyethylene glycol
PFU	Plaque forming unit
Pi	Phosphate ion
PPP	Protein Production Platform
PCR	Polymerase chain reaction
qRT-PCR	Quantitative real time PCR
RFU	Relative fluorescence unit
RLU	Relative light unit
RMSD	Root-mean-square deviation
Ro3	Rule of Three
RRV	Ross River virus
SAR	Structure-activity relationship
SBDD	Structure-based drug design
s.c	Subcutaneous
SD	Standard deviation
SDS	Sodium dodecyl sulfate

SEC	Size exclusion chromatography
SEM	Standard error of mean
SFV	Semliki Forest virus
TCEP	Tris-(2-carboxyethyl)phosphine
T _m	Melting temperature
TSA	Thermal shift assay
VEEV	Venezuelan Equine Encephalitis Virus
WT	Wildtype

LIST OF TABLES

Table	Title	Page
1.1	Functions of domains within the structural and nonstructural polyproteins	10
3.1	Crystallization conditions of ONNV-vc002	83
3.2	Data collection and refinement statistics of crystal form 3 of ONNV-vc002	85
4.1	Chemical structures of Compounds 1-9	113
4.2	Effect of hit compounds on nsP2 protease of CHIKV, ONNV and VEEV, where their IC ₅₀ values were determined	121
4.3	Kinetic parameters of nsP2FL and nsP2pro in the presence of Compound 2	129
4.4	Inhibitory profile of compounds against multiple luciferase-tagged CHIKV isolates, ONNV and SFV in HEK293T cells	140
4.5	Inhibition of CHIKV-SGP011-Gluc infection by different compounds in HEK293T and human skin fibroblast BJ cells	141

LIST OF FIGURES

Figure	Title	Page
1.1	Phylogeny of CHIKV clades	5
1.2	CHIKV genome organization	9
1.3	Life cycle of CHIKV in infected cells	11
1.4	A model for the whole glycoprotein layer of the alphavirus particle	12
1.5	Comparison of ONNV and CHIKV diseases in mouse model	16
1.6	Structure of papain-like nsP2 protease and conserved motifs within the nonstructural polyprotein of multiple alphaviruses	20
1.7	Available crystal structures of alphavirus nonstructural proteins	25
1.8	Druggable targets and inhibitors	33
1.9	Berberine treatment reduced CHIKV-induced footpad inflammation	34
3.1	Structures of nsP2 protease from the alphaviruses CHIKV, SINV and VEEV	60
3.2	nsP2 protein constructs	64
3.3	Size exclusion profiles of CHIKV-vc038, CHIKV-vc039 and CHIKV-vc045	65
3.4	Schematic representation of mechanism of EDANS-DABCYL FRET pair, using CHIKV peptide substrate as a representative	69
3.5	ATPase activity of CHIKV nsP2 helicase is dependent on the presence of MgCl ₂	71
3.6	Small Angle X-ray Scattering of CHIKV-vc038	76
3.7	Low resolution model of CHIKV-vc038 obtained from SAXS data using a combination of <i>ab initio</i> calculations and rigid body modeling	77

3.8	Single particle analysis of CHIKV-vc038	78
3.9	SDS PAGE and gel filtration profiles of ONNV-vc002	81
3.10	Comparison of protease activities between purified ONNV-vc002 and CHIKV-vc045	82
3.11	Crystal forms obtained from screening of ONNV-vc002	84
3.12	Superposition of ONNV-vc002 with CHIKV, SINV and VEEV nsP2 proteases	88
3.13	Crystal structure of ONNV-vc002	89
3.14	Structural alignment of ONNV-vc002 with CHIKV nsP2 protease (PDB: 4ZTB)	90
4.1	Fragment to lead discovery process	98
4.2	Fragment optimization strategies	99
4.3	SDS PAGE, cation exchange and gel filtration profiles of CHIKV nsP2FL	103
4.4	SDS PAGE and gel filtration profiles of ONNV nsP2FL	104
4.5	SDS PAGE and gel filtration profiles of VEEV nsP2FL	105
4.6	Dose-dependent curves and linear regression lines of CHIKV and VEEV peptide substrates	107
4.7	Reproducibility assessment of cell-free FRET based protease assay	111
4.8	Primary screening of fragment library	112
4.9	Dose-dependent analysis of Compounds 1-9	114
4.10	nsP2 ATPase activity in the presence of 100 μ M compounds	116
4.11	Dose-dependent nsP2 protease inhibition of CHIKV, ONNV and VEEV by hit compounds	119

4.12	Effect of combinatorial treatment on protease activity	123
4.13	Compounds 2-1 and 2-2 do not inhibit CHIKV nsP2 protease activity	125
4.14	MST capillary scan during binding affinity test	128
4.15	Enzyme inhibition kinetics of Compound 2 against alphavirus nsP2 protease	130
4.16	Thermal shift assay schematic	134
4.17	The effect of compounds on melting temperature (T_m) of CHIKV nsP2pro measured by thermal shift assay	135
4.18	Antiviral activities of compounds against several alphaviruses in HEK293T cells	139
4.19	Inhibition of CHIKV infection by Compounds 1 , 2 and Ribavirin in human skin fibroblast BJ cells	142
4.20	Screen of candidate compounds for treatment of acute CHIKV arthritis in mice	145
4.21	Crystals obtained from screening of CHIKVpro-Compound 2 complex	148

CHAPTER 1. Introduction

1.1. Introduction to Chikungunya virus (CHIKV)

1.1.1. General overview of Alphaviruses

Alphaviruses are members of the *Togaviridae* family and include a number of important human and animal pathogens (Strauss & Strauss, 1994). Alphaviruses have been categorized into 'Old World' or 'New World' viruses, based upon the type of disease they cause. Old World alphaviruses are generally associated with rheumatic disease in humans and New World viruses are primarily associated with potentially fatal encephalitic disease (Ryman & Klimstra, 2008). One alphavirus, chikungunya virus (CHIKV), is currently distributed in all continents, capable of causing widespread epidemics associated with severe morbidity (Gérardin et al., 2011; Schilte et al., 2013). CHIKV is an arthritogenic alphavirus and is a member of the Semliki Forest antigenic complex along with the closely-related O'nyong'nyong virus (ONNV), Ross River virus (RRV) and Semliki Forest virus (SFV) (Levinson et al., 1990). Epidemic transmission of CHIKV occurs through urban cycles between humans and primarily by the mosquito species *Aedes aegypti* and *Aedes albopictus* (Peyrefitte et al., 2007; Vazeille et al., 2008). CHIKV was first isolated in Tanzania during the earliest recorded epidemic in 1952 (Lumsden, 1955). Prior to its identification, the similarity in CHIKV-induced disease to that caused by dengue virus (DENV) led to the frequent misdiagnoses of CHIKV cases (Carey, 1971; Velasco et al., 2015).

1.1.2. Epidemiology and evolution of CHIKV

CHIKV outbreaks have been documented throughout Africa and Asia and include four distinct lineages; East/Central/South African (ECSA), Western Africa (WA), Indian Ocean (IOL) and Asian (Figure 1.1) (Wahid et al., 2017). The re-emergence of CHIKV in 2005 in several Indian ocean islands and across many parts of the world caused severe outbreaks and this drove the disease to prominence. In La Réunion alone, over one-third of the island's 785,000 inhabitants were crippled by the disease (Schuffenecker et al., 2006). Interestingly, sequence analysis of CHIKV strains from the ECSA lineage associated with epidemics in La Réunion revealed an alanine to valine mutation at position 226 in the E1 glycoprotein in greater than 90% of isolates (Schuffenecker et al., 2006). This polymorphism contributed to more efficient replication of CHIKV in *Ae. albopictus* and likely enhanced its transmission efficiency during its reemergence. CHIKV continued to spread to India and Southeast Asia to cause overwhelming epidemics (Arankalle et al., 2007). In 2007, autochthonous transmission of CHIKV initiated by imported cases was first reported in Italy and has since spread across the entire globe causing large numbers of epidemics that have infected millions of people in Australia, Europe and the Americas (Amraoui & Failloux, 2016; Liunbruno et al., 2008; Yactayo et al., 2016). The IOL lineage caused the majority of CHIKV outbreaks in 2004–2012, whereas the ECSA and Asian lineages are mainly responsible for the current outbreaks in the Americas (Weaver, 2014; Weaver & Forrester, 2015). As the distribution of the efficient vector *Ae. albopictus* extends into more temperate regions (Her

et al., 2009), it is likely the outbreaks will continue and spread into new geographic areas.

1.1.3. Characteristics of Chikungunya fever

CHIKV causes a debilitating musculoskeletal inflammatory disease commonly referred to as Chikungunya fever (CHIKF). People infected with CHIKV generally experience febrile illness similar to that seen in DENV infections. Polyarthralgia, the most characteristic symptom of the acute phase, is reported in > 85% of cases and it normally resolves within a week (Thiberville et al., 2013). However, when the disease continues into the chronic phase, patients experience prolonged arthritis that primarily affects the peripheral small joints for months to several years after the initial diagnosis, which is the hallmark of CHIKV disease (Ali Ou Alla & Combe, 2011; Borgherini et al., 2007; Borgherini et al., 2008; Manimunda et al., 2010; Pialoux et al., 2007; Rezza et al., 2007; Simon et al., 2007; Suhrbier et al., 2012). This causes serious economic and social impact on both the individual and the affected communities.

1.1.4. CHIKV pathogenesis

CHIKV infects susceptible hosts through inoculation at the skin via bites from infected mosquitoes. Skin resident cells such as skin fibroblasts and dermal macrophages are targets for the virus (Sourisseau et al., 2007; Young et al., 2019). The first round of virus replication occurs and this stimulates the innate immune response to contain the virus at the site of inoculation. Despite this response, CHIKV disseminates into the lymph

node before being released into the circulatory system, spreading to peripheral tissues, such as the muscles, where the virus further replicates (Assunção-Miranda et al., 2013). This triggers a potent inflammatory response driven by type I Interferon (IFN), resulting in a large influx of infiltrating cells and production of antiviral and proinflammatory cytokines and chemokines. Elevated levels of IFN- γ , IL-1 β , IL-6, CCL2, CXCL9 and CXCL10 have been detected during CHIKV infection (Ninla-aesong et al., 2019). Production of these cytokines recruits monocytes, macrophages, NK cells, and T cells to site of infection, which contribute to tissue injury and disease (Gardner et al., 2010).

Figure 1.1. Phylogeny of CHIKV clades. Phylogenetic organization of CHIKV clades based on the complete coding region of 74 chikungunya virus sequences. Scale bar represents the number of nucleotide substitutions per site. Figure adapted from (Wahid et al., 2017).

1.1.5. CHIKV genome and replication cycle

CHIKV genome consists of a linear, a positive-sense single-stranded RNA (ssRNA) of approximately 11.8 kb. The genome contains two open reading frames (ORFs), encoding for the nonstructural (ns) proteins, required for viral replication, and the structural proteins, which are translated from the internal sub-genomic promoter 26S immediately downstream of the nonstructural ORF. The 5' end of the genome has a 7-methylguanosine cap, while the 3' end is polyadenylated (Figure 1.2). The functions of domains found in the structural and ns polyproteins are summarized in Table 1.1.

The CHIKV replication cycle is essentially similar to the replication cycle of other alphaviruses (Figure 1.3) (Solignat et al., 2009). The virus forms icosahedral virions containing a nucleocapsid surrounded by a host-derived lipid bilayer envelope (Lescar et al., 2001). The envelope is studded with 80 trimeric spikes arranged in a $T = 4$ lattice, where each spike is composed of three heterodimers of the viral E1 fusion and E2 attachment glycoproteins (Figure 1.4) (Cheng et al., 1995; Mancini et al., 2000; Paredes et al., 1993; Voss et al., 2010). E2 mediates the attachment to the host cell via cell-surface receptors, which remain unidentified for CHIKV and many other alphaviruses. Upon receptor binding, virus particle enters via clathrin-dependent endocytosis (Bernard et al., 2010; Sourisseau et al., 2007). The acidic environment of the endosome triggers the E1-mediated fusion of the viral envelope and endosomal membrane, followed by the release of nucleocapsid and

uncoating of the viral genome in the cytosol (Kielian & Rey, 2006; Marsh & Helenius, 2006).

Viral genome replication is carried out by the nsPs that are initially produced as part of ns polyproteins P123 and P1234, which are later sequentially cleaved by the viral nsP2 protease to ultimately generate mature nsP1, nsP2, nsP3 and nsP4 (Lemm et al., 1994; Vasiljeva et al., 2003). The structural proteins are generated by translation of the subgenomic mRNA to produce a polyprotein that are processed to produce a capsid protein and glycoproteins (E1, E2 and E3). Following translation of the subgenomic RNA, the capsid protein is proteolytically cleaved from the polyprotein (Khan et al., 2002; Strauss & Strauss, 1994). The remainder is transported through the endoplasmic reticulum (ER) and Golgi where the 6K peptide is liberated by host cellular proteases, and the E1 and pE2 proteins are glycosylated (Simizu et al., 1984). During egress, the cellular protease, furin, cleaves pE2 to release the E3 peptide and generate the mature heterodimer of E1 fusion and E2 attachment proteins (De Curtis & Simons, 1988; Voss et al., 2010). The structural proteins encapsidate the viral genome, forming progeny virions near the plasma membrane, which bud from the host cell to infect neighboring cells and disseminate throughout the host (Jose et al., 2009).

1.1.6. Interplay of viral and host factors during CHIKV infection

Some host factors are known to be hijacked by CHIKV during infection, either to facilitate viral replication processes or to evade host immune mechanisms. The nuclear import and nuclear export signals found on

CHIKV capsid protein was found to interact with karyopherin α (Kara α) and chromosomal maintenance 1 (CRM1) respectively, to mediate CHIKV's translocation to the nucleus (Thomas et al., 2013). The E1 protein was postulated to interact with ATPase Na1/K1 transporting beta 3 polypeptide (ATP1B3), facilitating the fusion of viral envelope to host membrane during viral entry (Dudha et al., 2015). Other host factors such as protein tyrosine phosphatase non-receptor type 2 (PTPN2), and actin gamma 1 (ACTG1) were shown to interact with the E2 protein, but their mechanisms remain unknown (Dudha et al., 2015). Recently, the cell adhesion molecule Mxra8 was speculated to bind to E2 protein and facilitate CHIKV's entry into the cell (Zhang et al., 2018). NDP52 human autophagy receptor, was shown to interact with CHIKV nsP2 in the cytoplasm and postulated to prevent the nsP2 from localizing into the nucleus (Judith et al., 2013). This could delay cell death and allowed CHIKV more time to replicate. Furthermore, CHIKV relied on nsP2 for inducing host transcriptional and translational shutoff, whereby it suppressed the host UPR that was triggered by the production of CHIKV envelope proteins in the endoplasmic reticulum (Fros et al., 2015). The host UPR was also found to be antagonized by nsP4 through the suppression of eukaryotic translation initiation factor, alpha subunit (eIF2 α), phosphorylation (Rathore et al., 2013).

Figure 1.2. CHIKV genome organization. Diagram of the genome, showing the 5' cap, 5' untranslated region, nonstructural polyprotein open reading frame, and the major functions of the individual proteins, subgenomic promoter (26S), structural polyprotein open reading frame, 3' untranslated region and poly(A) tail.

Table 1.1. Functions of domains within the structural and nonstructural polyproteins

Proteins	Structural and/or Enzymatic Functions	Role in virus lifecycle
nsP1	Membrane association, guanosine-7-methyltransferase (MTase), guanylyltransferase (GTase)	mRNA capping
nsP2	N-terminal domain functions as nucleotide triphosphatase, helicase C-terminal domain functions as a cysteine protease	P1234 polyprotein processing
nsP3	Macrodomain	Unknown
nsP4	RNA-dependent RNA Polymerase	Production of viral RNAs
Capsid	Forms nucleocapsid core with the genomic RNA, trypsin-like protease	Formation of nucleocapsid
E1	Glycoprotein which forms part of the isocahedral protein shell on the virion	Mediation of membrane fusion
E2	Glycoprotein which forms part of the isocahedral protein shell on the virion	Interaction with host cell surface receptors
E3	Peripheral glycoprotein	Regulates spike assembly
6K	Residual polypeptide chain	Assists in E1 folding

Drug Discovery Today

Figure 1.3. Life cycle of CHIKV in infected cells. The virus particle attaches to cellular receptor through E2 and enters the cell via endocytosis. A decrease in pH in the endosome induce fusion of the virus to the host cell membrane, and the nucleocapsid is released into the cytosol. Following virus entry, two rounds of translation occur. Positive-sense genomic RNA acts directly as mRNA and is partially translated (5' end) to produce non-structural proteins (nsPs). These proteins are responsible for replication and formation of a complementary negative strand, the template for further positive-strand synthesis. Subgenomic mRNA (26S) replication occurs through the synthesis of full-length negative intermediate RNA, which is regulated by nsP4 and p123 precursor in early infection, and later by mature nsPs. Translation of the 26S sub-genomic RNA results in production of 5' structural proteins (C, E3, E2, 6k, E1), which are required for viral encapsidation and budding. Assembly occurs at the cell surface, and the envelope is acquired as the virus buds from the cell. Figure is adapted from (Kaur & Chu, 2013).

Figure 1.4. A model for the whole glycoprotein layer of the alphavirus particle. The structure of the CHIKV E2–E1 heterodimer fitted into the 9 Å resolution cryo-EM map of the Sindbis virus (SINV) virion (Mukhopadhyay et al., 2006), which is represented as a grey transparent surface. **(A)** The alphavirus icosahedral surface glycoprotein shell. Atomic model of the 240 chikungunya E2–E1 heterodimers arranged as 80 spikes. E2 domain A is colored cyan, B green and ,and E1 domains are sandy brown. **(B)** Top view of the spikes. The $T = 4$ icosahedral symmetry and quasi-symmetry axes are indicated. **(C)** Side view of a spike. Red spheres mark the transitional epitopes on E2 domain B and on E1 domain III, in a region buried at interspike contacts. The stem region, transmembrane (TM) domains, capsid and viral membrane are indicated. Figures are adapted from (Voss et al., 2010).

1.1.7. O'nyong'nyong virus, a close relative of CHIKV

O'nyong'nyong virus (ONNV) is closely-related to CHIKV, both antigenically and genetically, and it causes similar, but milder clinical symptoms in patients (Kiwanuka et al., 1999; Lanciotti et al., 1998; Powers et al., 2000; Rwaguma et al., 1997). Like CHIKV, there is no specific treatment for ONNV fever, and the illness is usually self-limiting. The similarity in their initial symptoms and signs makes laboratory diagnosis difficult due to serologic cross-reactions with CHIKV and possibly other alphaviruses co-circulating in Africa (Kiwanuka et al., 1999). However, their outbreak patterns and global distribution are quite distinct. While CHIKV has repeatedly been reported in Africa, Asia, and even Europe, ONNV has never been detected outside of sub-Saharan Africa (Powers et al., 2000). In particular, they are not transmitted by the same mosquito vectors (Saxton-Shaw et al., 2013). ONNV is unique among alphaviruses and one of very few arthropod-borne viruses transmitted by anopheline mosquitoes (Brault et al., 2004). These include *Anopheles funestus* and *Anopheles gambiae*, which are also malaria parasite vectors (Kiwanuka et al., 1999; Williams et al., 1965).

Like all other alphaviruses, ONNV is an enveloped virus with a single-stranded positive-sense RNA genome that has both the ns and structural regions. All four of its ns proteins are essential for RNA replication and polyprotein processing, and their functions are generally similar across the different alphaviruses. However, an earlier study demonstrated the presence of an opal termination codon between nsP3 and nsP4 provided fitness advantage to ONNV in anopheline mosquitoes

(Myles et al., 2006). Additionally, recent findings with chimeric viruses suggested that ONNV nsP3 could sufficiently restore CHIKV infection in *An. gambiae* (Saxton-Shaw et al., 2013).

And most recently, comparative studies between ONNV and CHIKV pathogenesis in mice demonstrated that ONNV exhibited less severe joint swelling, lower viremia and faster viral clearance (Figures 1.5A-B), despite showing similar joint peak swelling (at 6 dpi) and viremia peaks (at 2 dpi) (Figures 1.5C-D). Nevertheless, these observations from the various studies could be attributed to numerous aspects of the virus-host interactions and virus-virus regulatory functions, whereby they are still not well defined in ONNV infection. And since nsP3 remains poorly understood, further investigations are needed to reveal the exact mechanisms and functions of the highly-variable protein. Moreover, none of the ONNV proteins are structurally characterized to date; thus, it will be interesting to study and compare them with the solved protein structures of the other alphaviruses as they can further aid in our understanding of their conformations and functionality during RNA replication.

Given that CHIKV and ONNV share 85% homology in their structural proteins (Khan et al., 2002; Levinson et al., 1990) and are antigenically related, they are likely to share antigenic sites that are identical and/or conformationally similar. Antigenic studies have demonstrated that live-attenuated CHIKV vaccine candidates (Chan et al., 2019; Partidos et al., 2012) and neutralizing CHIKV-specific monoclonal antibodies can confer protection against ONNV infection by inducing cross-protective humoral immune responses (Fox et al., 2015). These

findings not only demonstrated but further highlighted the potential of CHIKV vaccines to protect humans against ONNV-induced disease.

Figure 1.5. Comparison of ONNV and CHIKV diseases in mouse model. Three-week-old WT C57BL/6 mice were infected with 1E6 PFU of ONNV-WT, CHIKV-WT or CHIKV-ZsGreen at the metatarsal region of the footpad. **(A)** Joint inflammation and **(B)** viremia progression in ONNV-infected mice were monitored over 2 weeks. Viremia was detected with ONNV nsP1 probe via qRT-PCR. Figure adapted from (Chan et al., 2019). **(C)** Joint inflammation and **(D)** viremia progression in CHIKV-infected mice were monitored over 2 weeks. Viremia was detected with CHIKV nsP1 probe via qRT-PCR. Figure adapted from (Carissimo & Ng, 2019).

1.1.8. Functions of nonstructural proteins (nsPs)

Although all functions of the individual nsPs have yet to be determined, the topic has been well studied (Kääriäinen & Ahola, 2002; Rupp et al., 2015; Strauss & Strauss, 1994).

The nsP1 protein is essential for initiating synthesis of viral RNA of negative polarity (Ahola et al., 1997; Hahn et al., 1989; Wang et al., 1991). The nsP1 protein possesses both guanine-7*N*-methyltransferase and guanylyltransferase activities, where they direct the capping of genomic and subgenomic viral RNAs (Ahola & Kääriäinen, 1995; Cross, 1983; Laakkonen et al., 1994; Li et al., 2015). The cap structure protects viral mRNA from cellular 5'-3' exonucleases and promotes the initiation of RNA translation into viral proteins. In addition, the cap structure, together with conserved hairpin structure present in the 5'UTR, hides viral RNA from detection by sensors of the innate immunity, such as RIG-I and/or MDA-5 (Hyde et al., 2014). The presence of α -helical amphipathic loop and palmitoylation sites allows the protein and nsP1-containing replication complex to anchor onto the host's plasma membrane, possibly through nsP1 interaction with the membrane's anionic phospholipids (Ahola et al., 2000, 1999; Laakkonen et al., 1996; Lampio et al., 2000; Spuul et al., 2007).

The nsP2 protein contains two main domains and they individually possess distinct functional roles. The N-terminal domain is an RNA helicase with both nucleotide triphosphatase (NTPase) and RNA triphosphatase (RTPase) activities (Das et al., 2014; Karpe et al., 2011). The helicase unwinds RNA secondary structures formed during viral RNA

replication and it is dependent on the NTPase activity (Gomez De Cedrón et al., 1999; Rikonen et al., 1994). The RTPase activity removes γ -phosphate from the 5'-end of nascent positive-sense RNAs, which enables the RNA to act as a substrate for the nsP1-mediated capping reaction (Karpe et al., 2011; Vasiljeva et al., 2000). The C-terminal region is a papain-like cysteine protease that is responsible for polyprotein processing and it recognizes conserved motifs within the polyprotein (Figure 1.6) (Gomez De Cedrón et al., 1999; Hardy & Strauss, 1989; Merits et al., 2001; Vasiljeva et al., 2001). The C-terminal subdomain of the protease is an S-adenosyl-L-methionine-dependent (SAM) RNA methyltransferase (MTase) domain with a classical MTase fold, but is enzymatically non-functional (Hu et al., 2016; Narwal et al., 2018; Russo et al., 2006). The alphavirus nsP2 protein has also been described as a virulence factor responsible for the transcriptional and translational shutoff in infected host cells and the inhibition of interferon-mediated antiviral responses contributing to the controlling of translational machinery by viral factors (Bhalla et al., 2016; Breakwell et al., 2007; Frolov et al., 2009; Gorchakov et al., 2005; Lulla et al., 2006; Ramakrishnan et al., 2017).

The precise role(s) of nsP3 is not completely understood, but it is heavily phosphorylated in the hypervariable domain and is essential for virus replication (Hahn et al., 1989; Vihinen et al., 2001). The deletion of this domain resulted in low viral pathogenicity, suggesting its importance in viral RNA transcription regulation (Lastarza et al., 1994; Tuittila et al., 2000). In addition, mutations in nsP3 have been shown to exhibit defects

in the initiation of minus-strand synthesis or subgenomic RNA synthesis (LaStarza et al., 1994; Rupp et al., 2011; Wang et al., 1994).

The presence of a characteristic GDD motif has led to the assignment of nsP4 as the viral RNA polymerase, and genetic studies are consistent with this assignment (Chen et al., 2017). The core RNA-dependent RNA polymerase (RdRp) domain at the C-terminal end was determined to be solely responsible for the RNA synthetic properties of the viral replication complex (Rubach et al., 2009). The RdRp participated in replicating the genomic RNA via a negative strand RNA and transcribing the 26S subgenomic RNA. The N-terminal domain appeared to be important for the interaction with polyprotein P123 to form replication complexes that are capable of synthesizing minus strands from plus-strand templates (Rupp et al., 2011; Tomar et al., 2006). Deletion and mutation of the RdRp domain demonstrated terminal adenylyltransferase (TATase) activity, suggesting it has a potential role in maintenance of the 3' poly-A tail at the 3'-end of positive-sense RNAs (Tomar et al., 2006).

Figure 1.6. Structure of papain-like nsP2 protease and conserved motifs within the nonstructural polyprotein of multiple alphaviruses. (A) Illustration of superposed structures of papain and chikungunya virus (CHIKV) nsP2 proteases (Protein Data Bank (PDB) 9PAD and 3TRK respectively). The structure of papain protease is presented as a solid, blue ribbon. The structure of CHIKV nsP2 protease is presented as solid, green ribbon. The domain common for both proteases is highlighted in dark colors and is enlarged subsequently to show the conserved catalytic dyad. **(B)** Alignment of nsP2 cleavage sites. The nomenclature of Berger and Schechter is used to identify residues on the amino (P1, P2, etc.) or carboxy (P1', P2', etc.) termini of the scissile bond. The arrow indicates the location of the cleavage site. Cleavage sites between the ns proteins contain a common motif, AG(A/C)↓(G/Y/A) (A. Lulla et al., 2006). Abbreviations: Chikungunya virus (CHIKV), o'nyong'nyong virus (ONNV), semliki forest virus (SFV), ross river virus (RRV), sindbis virus (SINV) and Venezuelan equine encephalitis virus (VEEV). Adapted from (Bakar & Ng, 2018).

1.2. Available structures of nonstructural proteins (nsPs)

Out of the four nsPs, crystal structures are available for the nsP2 helicase (PDB ID: 6JIM), nsP2 protease (PDB ID: 3TRK and 4ZTB) and P23 precursor (PDB ID: 4GUA). They not only provided structural insights into the catalytic and substrate binding sites, but also provided essential information about the functioning of the replicase proteins. Furthermore, they allow a deeper understanding of the processes involved in ns polyprotein processing and RNA replication.

The structure of CHIKV nsP2 helicase bound to the conserved 3' end of the genomic RNA and the nucleotide analog ADP-AIF4 revealed the arrangement of multiple domains. They include a small hairpin looped N-terminal domain (NTD), the superfamily 1 (SF1) accessory domains 1B and α -helical Stalk, two parallel Rec-A-like domains designated RecA1 and RecA2, and an extended α -helical connector (Figure. 1.7A) (Law et al., 2019). The NTD-Stalk-1B interacts with both RecA1 and the 3' half of the bound RNA to create a cover for the ssRNA-binding groove. Mutation and deletion studies have indicated that the interaction of NTD with other viral proteins might be important for ns polyprotein processing and RNA infectivity (Atasheva et al., 2007; Das et al., 2014; Lulla et al., 2018). The accessory domains, 1B and stalk, mediate the interaction of ssRNA to the helicase core, which is formed by the RecA1 and RecA2 domains. The helicase core is structurally similar to that of SF1 RNA helicase from Tomato mosaic virus (Nishikiori et al., 2012), where the RecA domains work in parallel as a motor for RNA binding and NTP hydrolysis (Singleton et al., 2007). The RNA-binding motifs and ATPase binding site lay on top

and between the two domains respectively. The connector, which connects the helicase region to the nsP2 protease domain, wraps around the helicase core close to the ATPase active site and might have a role in closing the ATP binding cleft.

The nsP2 protease was the first ns protein to be structurally solved. To date, the crystal structures of CHIKV, SINV and VEEV nsP2 proteases are available (Cheung et al., 2011; Narwal et al., 2018; Russo et al., 2006; Shin et al., 2012). Relative to CHIKV nsP2 protease, SINV and VEEV nsP2 proteases showed only 65% and 40.5% amino acid identity respectively (Appendix 4). Despite the low sequence identity, their overall structure is similar and their catalytic site residues remain conserved. Alphavirus nsP2 protease characteristically consisted of a papain-like cysteine protease subdomain that is connected to the C-terminal SAM-dependent MTase-like subdomain via a flexible linker (Figure 1.7B) (Russo et al., 2006). The N-terminal protease subdomain is dominantly helical and contains seven α -helices and four β -strands, whereas the C-terminal subdomain is comparatively larger in size and contains five α -helices and seven β -strands aligned with each other (Narwal et al., 2018). The catalytic dyad residues Cys478, which is positioned at the N-terminal of a α 1-helix and His548, is present beneath the variable loop between β 1 and β 2 strands are contributed by the protease subdomain. The substrate binding site comprising of the catalytic dyad appears as a shallow depression in a deep long groove at the interface of the two subdomains. And recently, Asn547, a residue located in the variable loop,

was determined to be crucial in regulating the binding of substrate peptide at the active site (Hu et al., 2016; Russo et al., 2010).

To date, the function(s) of nsP3 remains poorly understood. However, based on conserved motif studies, nsP3 is known to organize into three domains: an amino-terminal conserved macrodomain (Gorbalenya et al., 1991), a central alphavirus-unique domain (AUD), and a hypervariable-sequence carboxyl terminus that is devoid of any secondary structure (Tuittila & Hinkkanen, 2003). The resolved structure of a precleavage form of SINV polyprotein P23 revealed the presence of the macrodomain and zinc-coordination site within the AUD of nsP3, in addition to the protease and MTase-like subdomains from nsP2 protease (Figure 1.7C) (Shin et al., 2012). The macrodomain consists of a central, twisted six β -strands surrounded by three α -helices. The ribose-binding site is found close to positively charged residues within the nsP3 macrodomain. Moreover, these basic patches could be involved in the binding of the negatively charged ADP-ribose and longer substrates. The central AUD contains an antiparallel α -helical bundle, two parallel β -strands, and a zinc-coordination site. Four conserved cysteine residues are involved in zinc coordination; C263 and C265 are located in the loop between the two α -helices, whereas C288 and C306 are at the carboxyl termini of the two parallel β -strands. Mutational studies performed indicated that zinc ion plays a structural role as zinc-coordination mutants failed to express the ns polyprotein and prevented efficient SINV infection. Additionally, the structure also revealed that the P2/3 cleavage site is located at the base of a narrow cleft formed by the MTase-like domain of

nsP2 and the macro domain of nsP3 (Figure 1.7C, black arrow), away from the nsP2 protease active site. The inaccessibility of the P23 cleavage site itself indicated that access is tightly regulated.

Figure 1.7. Available crystal structures of alphavirus nonstructural proteins. **(A)** Schematic representation of CHIKV nsP2 helicase (PDB ID: 6JIM). The domains are arranged with the N-terminal domain (NTD) in pink, domain 1B in yellow, the stalk domain in orange, RecA1 domain in green, RecA2 domain in cyan, and the connector in blue. Figure is adapted from (Law et al., 2019). **(B)** Structure of CHIKV nsP2 protease (PDB ID: 3TRK). The N-terminal protease subdomain is colored blue, where it contains the catalytic dyad residues, Cys 478 and His548 (shown in sticks and circled). The C-terminal MTase-like subdomain is colored cyan and the linker connecting the two subdomains is shown in purple. **(C)** Structure of SINV P23 precursor protein (PDB ID: 4GUA). The structure consists of four domains: protease and methyltransferase-like domains of nsP2, which are colored in blue and cyan respectively; the macro and zinc-binding domain of nsP3, which are colored in yellow and red respectively. The P2/3 cleavage site is labelled with black arrow.

1.3. Current strategies for inhibition of CHIKV infection

To date, no effective vaccines or antivirals are commercially available for the treatment of CHIKV infection. As a result, many are being developed and are either in the early discovery or the various stages of the clinical pipeline. Currently, 16 CHIKV vaccine candidates are being developed using diverse vaccine approaches. These include live-attenuated vaccines, whole-virus inactivated vaccines, replication competent and defective recombinant-virus vectored vaccines, virus-like particles, recombinant protein subunits, and DNA vaccines (Goyal et al., 2018). In addition, several specific compounds have been reported to inhibit CHIKV infection *in cellulo*, but they have not shown significant therapeutic benefits in preclinical models or clinical cases (Lamballerie et al., 2008). Instead, virus-infected individuals are given non-steroidal anti-inflammatory drugs and analgesic agents to control fever and joint pains (Parashar & Cherian, 2014). Therefore, there is a need to develop novel antivirals that are potent and could specifically inhibit CHIKV replication. Given that CHIKV pathology is driven by host immune response, drugs targeting specific immune components such as CD4+ T cells and NK cells can be explored as possible CHIKV interventions. Research efforts are currently focused on either repurposing FDA-approved inhibitors known to be effective for viral therapy or designing antivirals that target virus entry, viral genome replication, or functions of specific viral proteins (Figure 1.8).

1.3.1. Drugs inhibiting viral entry

Inhibitors of viral entry presents an attractive therapeutic strategy because they can minimize the damage caused by virulence factors during intracellular viral replication. In addition, these inhibitors might target extracellular components like host-cell receptors, which are more accessible, thereby reducing the required dosages and limiting cytotoxicity (Teissier et al., 2011). Chloroquine, an antimalarial drug, (Delogu & de Lamballerie, 2011; Delogu et al., 2011) was suggested to disrupt endosome-mediated CHIKV internalization, possibly through the prevention of endosomal acidification. While *in vitro* inhibition of CHIKV replication is strongly established, two clinical trials that use different treatment regimens presented conflicting data (Brighton, 1984; de Lamballerie et al., 2009). However, the patient cohort for both studies was too small for the results to be conclusive. Therefore, more studies need to be performed to determine whether changes in treatment regimens would have an impact on viremia and arthritis symptoms. One other example is Arbidol, a broad-spectrum antiviral approved for the treatment of influenza and other respiratory infections, where it was suggested to interfere with the early stages of CHIKV infection by preventing virus adsorption to target cells (Delogu et al., 2011). However, the precise mechanism remains to be fully elucidated and its clinical efficacy as a prophylactic treatment was never examined.

1.3.2. Inhibitors targeting viral genome replication

nsP4 RdRp is another attractive target for antiviral drug design since human replication or transcription enzymes lack RdRp activity. Nucleoside analogs are often used as antiviral agents by prematurely terminating elongation of the nucleic acid chain during viral RNA synthesis (De Clercq & Neyts, 2009). Several studies have attempted to repurpose approved therapeutic nucleoside analogs that are currently in use for the therapy of viral infections of high medical importance, such as HIV and arthropod-transmitted flaviviruses.

One such example is Ribavirin, which is a guanosine analog used for the treatment of hepatitis C, have been shown to significantly inhibit CHIKV and SFV genome replication by depleting GTP pools (Briolant et al., 2004; Rothan et al., 2016, 2015). Another nucleoside analog, β -D-N⁴-Hydroxycytidine, was reported to interfere with CHIKV/VEEV replication with high potency, primarily inhibiting early negative-strand RNA synthesis, either through chain termination or mutagenesis (Ehteshami et al., 2017; Urakova et al., 2017). Favipiravir, a purine analog that is used for the treatment of influenza infections, was found to possess broad-spectrum anti-alphavirus activity *in vitro* and could provide protection in a mouse model of lethal CHIKV infection (Delang et al., 2014). Its active form, favipiravir-ribofuranosyl-5'-triphosphate, competitively inhibits the incorporation of ATP and GTP by the viral RdRp, leading to chain termination (Furuta et al., 2013). The defluorinated analog of favipiravir, T-1105, was also reported to inhibit CHIKV replication, likely through interfering with the function of viral RdRp (Delang et al., 2014). Although

the above-mentioned analogs demonstrated good CHIKV inhibition and were postulated to target nsP4 RdRp, the precise mechanism of inhibition remained unknown and further investigations are required.

1.3.3. Non-specific inhibitors

Several other inhibitors exhibited reasonable antiviral action against multiple alphaviruses but their exact mode of actions are not well understood. Some were speculated to abrogate the synthesis of genomic and subgenomic viral RNA as well as interfere with certain host factors required for efficient viral replication (Abdelnabi et al., 2017; Karlas et al., 2016; Lam et al., 2015; Paessler et al., 2008). An advantage of targeting host factors is that it would prevent the development of antiviral resistance. However, targeting them could be risky, since they perform vital functions in the host cells. One such example is Berberine, which is a plant-derived isoquinoline alkaloid that was found to effectively reduce the number of infectious particles produced (Varghese et al., 2016). Moreover, Berberine was demonstrated to reduce footpad inflammation and postulated to inhibit MAPK signaling pathways in the alphavirus infectious cycle (Figure 1.8) (Varghese et al., 2016). The calmodulin signalling and fatty acid synthesis pathways have also been targeted with great outcomes. A combination therapy of calmodulin inhibitor pimozone and the fatty acid synthesis inhibitor TOFA caused a significant reduction of CHIKV replication *in vitro* and exhibited prophylactic antiviral effect *in vivo* (Karlas et al., 2016). Other inhibitors include prostratin and 12-O-tetradecanoylphorbol 13-acetate, where they displayed potent inhibition

of CHIKV-induced cytopathic effect. However, further investigations are still required to identify the molecular targets that are impacted by these compounds.

1.3.4. Antivirals against nsP2 protease

The nsP2 protease is an attractive target since it performs the essential function of ns polyprotein processing into the four nsPs to make up the replication complex (Chung et al., 2014; Lulla et al., 2006; Lulla et al., 2018; Singh et al., 2018; Vasiljeva et al., 2001). The nsP2 protease domain is by far the most characterized functionally and structurally and has been routinely used as an antiviral target in the development of nsP2 protease inhibitors (Cheung et al., 2011; Narwal et al., 2018; Russo et al., 2006). Moreover, mutations that affected the cleavage site directly or indirectly have shown to produce attenuated or non-infectious viruses which suggest that an inhibitor of the essential nsP2 protease function would make a potent antiviral (Chan et al., 2019).

Structure-activity relationship analysis and *in silico* drug design approaches are often employed in the development of molecules that target the catalytic groove of nsP2 protease. They have led to the identification of numerous non-peptidic inhibitors and experimental validation of some of these compounds demonstrated good anti-protease activity (Bassetto et al., 2013; Das et al., 2016; El-Labbad et al., 2015; Hu et al., 2016). Compound 25 in particular, was designed based on VEEV's nsP2 protease structure and was demonstrated to display good antiviral activity against CHIKV (Bassetto et al., 2013). Although many more

molecules were illustrated to dock in the protease's active site with good affinity *in silico* (Byler et al., 2016; Nguyen et al., 2015; Singh et al., 2012), their biological relevance was never validated.

Other inhibitors having peptide-like structures were also shown to exhibit good protease inhibition. One of such a compound would be Pep-1, which is a peptidomimetic that is specific for nsP2 protease and could efficiently inhibit CHIKV replication (El-Labbad et al., 2015; Singh et al., 2018). Another would be E64d, which is a peptide-like generic cysteine protease inhibitor that was shown to effectively inhibit VEEV protease activity. However, its poor efficacy in inhibiting VEEV and was never demonstrated to inhibit CHIKV replication. Nonetheless, the E64d-bound enzyme structure demonstrated that the inhibitor binds beneath a β -hairpin at the interface of the protease and SAM-dependent MTase-like subdomains (Hu et al., 2016). This representation is a major information that identified key residues that may be important for substrate binding and recognition.

Although these molecules possess decent anti-protease activity and/or exhibit *in vitro* inhibition of CHIKV replication, they are still not proven to be potent inhibitors *in vivo*. Thus, there is still room to discover more chemical motifs that can inhibit both the nsP2 protease function and viral replication with great efficiency. And with the recent characterization of the helicase structure, it provided another excellent target for antiviral development.

1.3.5. Treatment strategies targeting CHIKV-induced inflammation

Given that CHIKV pathology is driven by host immune response, drugs targeting proinflammatory mediators have shown promise in reducing CHIKV-induced disease. One such therapeutic candidate is Bindarit, an inhibitor of monocyte/macrophage-derived CCL2 which could completely abolish CHIKV-induced pathology (Rulli et al., 2009, 2011). Other specific immune components such as the virus-specific CD4⁺ T cells was also targeted. Fingolimod (FTY720), an agonist of sphingosine 1-phosphate receptor, successfully limited the migration of CD4⁺ T cells into the joints and effectively suppressed CHIKV-induced joint pathology (Teo et al., 2017). Furthermore, a combination therapy involving CTLA4-Ig and a neutralizing anti-CHIKV monoclonal antibody reduced T cell accumulation in joints, ameliorated joint inflammation and decreased proinflammatory mediators (Miner et al., 2017).

Drug Discovery Today

Figure 1.8. Druggable targets and inhibitors. Direct-acting inhibitors that target virus entry, viral genome replication, or functions of specific viral proteins. Inhibitors are labelled in red.

Figure 1.9. Berberine treatment reduced CHIKV-induced footpad inflammation. Representation of *in vivo* study showing use of Berberine (BBC), an antiviral alkaloid agent, that can reduce **(A)** footpad inflammation and **(B)** viremic load (Varghese et al., 2016).

1.3.6. Current limitations

Despite the urgent need for effective antiviral drugs, several pertinent issues pose a hindrance to the advancement in antiviral drug discovery. Most of the antivirals are only in the preliminary stages of research and this is partly due to the poor knowledge of molecular mechanisms involved in CHIKV replication, including host-cell receptors as well as cell signalling pathways prior to viral budding. Additionally, the failure of several drug candidates in translating *in vitro* efficacy to *in vivo* is attributed to the gaps in our current understanding of CHIKV biology and pathogenesis in acute and chronic disease conditions. An understanding in these processes will not only provide information on the mechanisms underlying viral replication but may reveal novel, host-derived targets for therapeutic interventions.

Furthermore, the substantial gap in our understanding of the crystallographic structures of key viral nsPs poses another major challenge in antiviral development. Since the elucidation of the nsP2 protease structure, the protein has been frequently targeted for the development of potent protease inhibitors through a combined approach of structure-based pharmacophore modelling, drug docking and/or high-throughput screening. Such research has substantially increased the chances of developing CHIKV protease inhibitors. Therefore, it will be essential to determine the crystal structures of the other nsPs as the information will be an important component for the development of target-based drug discovery and will likely increase the probability of developing a successful alphavirus antiviral drug candidate. Additionally, alphavirus

RNA synthesis requires precise structural arrangements of replication complexes, which would allow interactions and arrangements of nsPs and ns polyproteins to produce the functional shifts necessary during the different replication stages. However, these interactions remain largely uncharacterized. Understanding the interactions within the complexes would be informative for determining how the complex is assembled and matures during the virus replication cycle. Thus, determining the spatial arrangements within the RNA complex would be a large step towards our knowledge of the roles individual nsP plays in each RNA synthetic activity.

1.4. Aims and approaches

The severity of CHIKV disease underscores the need to find new antivirals for new targets for antiviral development. As nsP2 is an essential component in polyprotein processing and RNA replication, it is an attractive target for the development of anti-alphaviral drugs. Although the structures of the two domains encompassed within nsP2 were individually available, the overall architecture of nsP2 and the intramolecular dynamics remained poorly characterized. Moreover, antiviral compounds targeted towards nsP2 catalytic activity remained moderately potent and they lacked antiviral efficacy as compared to compounds targeting other components that were essential in the viral replication cycle. An attempt to bridge the knowledge gaps defined above was the focus of this thesis.

The first objective of this thesis was to gain structural and functional insights into the global architecture and dynamics of the full length nsP2

by integrating several techniques including Small-angle X-Ray Scattering (SAXS), negative stain electron microscopy (EM) and X-ray crystallography as well as biophysical binding studies and biochemical activity assays. This could provide a more holistic view of the protein's function and mechanism, and open new avenues for antiviral development in the future. In addition, another arthritogenic alphavirus, O'nyong'nyong virus (ONNV), which is most closely related to CHIKV, has been proposed to share similar nsP2 protease structural folds due to its high sequence homology but the three-dimensional (3D) structure has never been shown. Therefore, another aspect of this thesis was to reveal the structure of ONNV nsP2 protease through X-ray crystallography and compare its structural features with the resolved structures of CHIKV, SINV and VEEV nsP2 proteases. These results would be covered in Chapter 3.

The second objective of this thesis was to discover new chemical motifs that could effectively target alphavirus nsP2 protease. The results from chapter 3 provided the scope on how structural and functional information of nsP2 can be relevant to the discovery of novel chemical entities. Inhibition of this protein was frequently found to be effective in suppressing viral replication *in vitro*, while the efficacy is limited by drug dissemination to viral replication sites *in vivo*. A fragment-based screening approach was applied to CHIKV nsP2 protease to identify novel chemical motifs that could serve as a starting point for identifying fragments that were more suitable for optimization. Although hits would likely show relatively lower affinity as compared to other known larger

compounds due to its small molecular size, they could be evolved into high affinity leads in the future by incorporating unique interactions in the binding site, either by expanding with additional binding elements or by joining fragments together. In addition, *in vitro* antiviral experiments were employed to validate the biological relevance of these fragments in mitigating viral replication and to determine their mechanism of action. These results would be covered in Chapter 4.

Overall, structural results obtained from the studies in this thesis could be used as a platform for the design of novel therapeutics. Moreover, the success of identifying hits through fragment-based screening could advocate its use as a go-to screening strategy against the newly resolved nsP2 helicase in the future.

CHAPTER 2. Materials and Methods

2.1. Materials

2.1.1. Primers

Primers used were purchased from IDT (Singapore). Refer to Appendix 2 for the primer sequences.

2.1.2. Fragment library and compounds

The Ro3 fragment library, containing 1000 compounds, was purchased from Maybridge (Thermo Fisher Scientific Inc). The compounds are in the range of 90 to 300 Da in size, with eLogP values between -2.5 and 3.0. The compounds are dissolved in 100% dimethyl sulfoxide (DMSO) to a final concentration of 200 mM and stored at -20 °C.

The compounds Berberine, Obatoclax, Prostratin, Phorbol 12-myristate 13-acetate (PMA), E64d and Ribavirin were purchased from Sigma and Pep-1 from Mimotopes. All compounds were dissolved in 100% DMSO to a final concentration of 200 mM. Compounds 1, 1C, 1D and 9 were provided by Professor Andres Merits from University of Tartu and were dissolved in 100% DMSO to a final concentration of 100 mM. All compounds were stored at -20 °C.

2.1.3. Reagents for bacteria culture

2.1.3.1. Antibiotics

Ampicillin and kanamycin were dissolved in water to final concentrations of 100 mg/mL and 50 mg/mL respectively, and sterile filtered through a 0.22 µm Minisart syringe filter (Sartorius). Aliquots were stored in -20 °C.

Chloramphenicol was dissolved in ethanol at a final concentration of 34 mg/mL.

2.1.3.2. Media

Luria-Bertani (LB) media was used for the expression of recombinant protein in *E. coli*. 10 g bacto tryptone, 5 g yeast extract and 5 g NaCl were dissolved in 1 L of distilled water and autoclaved at 121°C for 15 min. To make LB agar plate, 1.5 % agar was added to the medium before autoclaving. Different antibiotics were added to the medium as required before bacterial culture.

2.1.4. Protein extraction materials

2.1.4.1. Buffers

Lysis buffer:	50 mM HEPES (pH 7.4), 500 mM NaCl, 5% (v/v) glycerol, 1% Triton-X, 2 mM DTT
Elution buffer:	50 mM HEPES (pH 7.4), 500 mM NaCl, 300 mM Imidazole, 5% (v/v) glycerol, 1% Triton-X, 2 mM DTT
SEC buffer:	50 mM HEPES (pH 7.4), 50 mM Arginine, 300 mM NaCl, 5% (v/v) glycerol, 2 mM DTT
NTPase reaction buffer:	50 mM MOPS (pH 7.25), 0.05 mg/ml BSA, 2 mM DTT, 1 mM MgCl ₂
Protease reaction buffer:	20 mM HEPES (pH 7.2), 2 mM DTT
SDS-PAGE running buffer:	25 mM Tris base, 192 mM Glycine, 0.1% SDS

Coomassie Brilliant Blue

staining solution: 0.1% Coomassie® R-250, 40% ethanol,
10% acetic acid

5X Gel loading dye: 5% Sodium dodecyl sulphate, 50% glycerol,
0.1% Bromophenol Blue, 250 mM Tris-HCl
(pH 6.8), and 5% β -mercaptoethanol

2.1.4.2. Chromatography resin / columns

Affinity resin: Ni-NTA Agarose (QIAGEN)

Ion exchange column: HiTrap® SP HP (GE Healthcare)

Size exclusion column: HiLoad® 16/600 Superdex® 75 prep grade (GE
Healthcare), HiLoad® 16/600 Superdex® 200
prep grade (GE Healthcare)

2.1.5. Crystallization kits

Crystallization screening kits were purchased, not exclusively, from
Hampton Research and Molecular Dimensions.

2.1.6. Peptide substrates

Peptide substrates used for the cell-free FRET-based protease assay
were purchased, not exclusively, from BioBasic and Mimotopes. Peptides
were designed based on the nsP1/2 junction sequences unique to the
alphaviruses. Substrates were dissolved in 100% DMSO to a final
concentration of 100 mM. Refer to Appendix 3 for the peptide sequences.

2.1.7. Cell lines

HEK293T cells: Continuous cell line from HEK293 cell line (CRL-1573) transformed with SV40 (Simian Vacuolating Virus 40) large T-antigen (Graham et al., 1977). Purchased from Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ) (ACC635).

BJ cells: Primary cell line from the normal foreskin of a newborn. Purchased from ATCC (CRL-2522).

2.1.8. Virus stocks

SGP11 / SGP11-Gluc: CHIKV isolate from the ECSA lineage obtained from a patient at the National University Hospital, Singapore, during the 2008 CHIKF outbreak in Singapore. This particular CHIKV isolate contains the A226V mutation in the E1 glycoprotein.

LR2006 OPY1-Gluc: CHIKV isolate from ECSA lineage obtained from a French patient returning from a trip to the Reunion Island during the CHIKF outbreak in 2006 (Bessaud, Peyrefitte, Pastorino, Tock, et al., 2006). This particular CHIKV isolate contains the A226V mutation in the E1 glycoprotein.

CNR20235-Gluc: CHIKV isolate from Asia lineage obtained from a patient in a December 2013 Saint Martin Island,

French West Indies, outbreak (Leparc-Goffart et al., 2014).

IND91-Gluc: CHIKV isolate from ECSA lineage obtained from a patient during the CHIKF Indian Ocean outbreak in 2006 (Schuffenecker et al., 2006)

IMTSSA/5163-Gluc: ONNV isolate obtained from a French patient in Chad (Bessaud, Peyrefitte, Pastorino, Gravier, et al., 2006).

SFV-Gluc: L10 strain isolated from female mosquitoes (*Aedes abnormalis* Theobald) caught in Bwamba, Uganda in 1942.

For all Gluc virus stocks, Gluc reporter gene was inserted into the CHIKV, ONNV and SFV isolates at a site between the nonstructural and structural genes. The CMV promoter was placed upstream of viral sequence and an extra copy of the 26S subgenomic promoter was placed downstream of structural region to ensure expression of the structural proteins.

2.2. Methods

2.2.1. Cloning of alphavirus nsP2 for *Escherichia coli* expression

Codon-optimized genes encoding the full-length wildtype nsP2 protein from CHIKV-SGP11 isolate and VEEV Trinidad Donkey strain were purchased from Genescript. ONNV-IMTSSA/5163 replicon was from Professor Andres Merits from University of Tartu. Fragments encoding residues 2-795 (CHIKV-vc038) and 469-791 (CHIKV-vc045) of CHIKV-SGP11 were amplified and cloned into the pNIC-CH2 expression vector (Novagen), generating proteins with C-terminal hexahistidine tag. Fragments encoding residues 4-470 (CHIKV-vc039) of CHIKV-SGP11 were amplified and cloned into the pNIC28-Bsa4 expression vector (Novagen), generating proteins with N-terminal hexahistidine tag. Fragment encoding residues 471-792 (ONNV-vc002) of ONNV nsP2 was amplified from ONNV-IMTSSA/5163 replicon and cloned into pNIC-CH2 expression vector (Novagen), generating proteins with C-terminal hexahistidine tag. Fragments encoding residues 1-796 of ONNV nsP2 (ONNV nsP2FL) amplified from ONNV-IMTSSA/5163 replicon and residues 1-794 of VEEV (VEEV nsP2FL) amplified from VEEV Trinidad Donkey strain were cloned into pET28b and pET16b expression vectors (Novagen) respectively, generating proteins with N-terminal hexahistidine tag. All constructs are listed in Appendix 1. All plasmids were transformed into *Rossetta* (DE3) strain of *E. coli*.

2.2.2.Expression and purification of nsP2 proteins

A single transformed colony was inoculated in 50 mL LB broth, supplemented with 50 µg/mL kanamycin and 34 µg/mL chloramphenicol, and grown at 37°C overnight. Next, 10 mL of bacteria suspension was transferred to 1 L of LB broth supplemented with 50 µg/mL kanamycin and 34 µg/mL chloramphenicol and allowed to grow to an OD₆₀₀ of 0.8 units. Protein expression was induced with 0.4 mM IPTG. Following 18 h incubation at 16°C, bacterial cells were harvested via centrifugation (4000 x g, 4°C, 15 min) and cell pellets were stored at -80°C for further use.

Cell pellet was resuspended in ice cold lysis buffer containing 50 mM HEPES pH 7.4, 500 mM NaCl, 5% (v/v) glycerol, 1% Triton-X and 2 mM dithiothreitol (DTT) and lysed by sonication on ice (pulses; 5 s ON, 10 s OFF). The cell lysate was centrifuged (20,000 x g, 4°C, 45 min) to separate the supernatant from cell debris. The supernatant was filtered through a 0.45 µm Minisart syringe filter, and the lysate was incubated with 10 mL of pre-equilibrated Ni-NTA agarose (Qiagen) with agitation at 4°C for 1 h. The bound nsP2 protein was eluted with 300 mM imidazole in elution buffer containing 50 mM HEPES (pH7.4), 500 mM NaCl, 5% (v/v) glycerol and 5 mM DTT.

nsP2 protein fractions were pooled, diluted 5-folds with buffer containing 50 mM HEPES (pH 7.4), 5% (v/v) glycerol and 2 mM DTT, and then loaded onto a pre-equilibrated HiTrap SP column. The nsP2 protein was eluted using a salt gradient (from 0 to 1 M NaCl). The nsP2 protein were concentrated up to 5 mL, using a Vivaspin ultracentrifugal filter with MWCO 30 kDa (Sartorius), and further purified via size exclusion

chromatography (HiLoad® 16/600 Superdex® 200 prep grade, GE Healthcare). Following elution with 50 mM HEPES (pH 7.4), 50 mM Arginine, 300 mM NaCl, 5% (v/v) glycerol and 2 mM DTT, the samples containing pure nsP2 protein (0.75 mg per liter LB medium) were pooled and concentrated to ~9 mg/mL using a Vivaspin ultracentrifugal filter with MWCO 30 kDa (Sartorius). The ONNV and VEEV nsP2 proteins were expressed and purified by a similar method. The purity and activity of nsP2 proteins were verified by SDS-PAGE, NTPase and fluorophore-based protease activity assays as described below.

2.2.3. SDS-PAGE

12% SDS-PAGE gels were used to assess the purity of protein samples. Samples were collected at each step of the purification process and were mixed with 2X Laemlli gel loading buffer and heated at 95°C for 10 min. The denatured samples were then loaded to gel and run at 200 V for 40 min. The gel was stained with Coomassie Brilliant Blue solution.

2.2.4. NTPase activity assay

NTPase assay was performed by using the Malachite Green phosphate assay kit (BioAssay Systems) according to the manufacturer's protocol. In brief, 1 nM of nsP2 protein was pre-incubated with 1 μ M compound at 22°C for 10 min in NTPase reaction buffer in a final volume of 80 μ L. Then 1 mM ATP substrate was added and further incubated at 37°C for 40 min. After incubation, 20 μ L of Reagent A was added, incubated for 5 min at

RT and spectrophotometric measurement at 620 nm was carried out using Infinite M200 microplate reader (Tecan).

2.2.5. Cell-free FRET-based protease assay

The continuous protease assay was performed using a FRET-based peptide substrate representing the nsP1/nsP2 cleavage sites of either the CHIKV or VEEV P1234 polyprotein. The peptide substrates were chemically linked to the FRET pair 4-[[4-(dimethylamino) phenyl]-azo] benzoic acid (DABCYL) and 5-[(2'-aminoethyl)-amino] naphthalenesulfonic acid (EDANS) at the amino and carboxy termini, respectively. The cleavage of peptide substrates in the presence of protease leads to dissociation of FRET pair and allows for the detection of relative protease activity by monitoring fluorescence of EDANS. In a 100 μ L reaction, 0.5 μ M of nsP2 protein was mixed in protease reaction buffer and then peptide substrate was added to a final concentration of 1 μ M. Continuous FRET measurement was carried out with an excitation wavelength of 340 nm and emission wavelength of 490 nm at 30°C for 1 hr at 1 min intervals using a Infinite M200 microplate reader (Tecan).

2.2.6. Fluorescence Extinction Coefficient determination

In a 100 μ L reaction, 10 μ M of nsP2 protein was mixed with varying concentrations of peptide substrate (0.2 μ M, 0.5 μ M, 1 μ M and 2 μ M) in protease reaction buffer. Continuous FRET measurement was carried out with an excitation wavelength of 340 nm and emission wavelength of 490 nm at 30°C for 2 hr at 5 min intervals using a Infinite M200 microplate

reader (Tecan). The curve of RFU, obtained from varied substrate concentrations was plotted as a function of time. The slope of the curve was calculated and assigned as fluorescence extinction coefficient (FEC) for the peptide substrate.

2.2.7. Measurement of steady-state kinetic parameters

Michaelis-Menten constant (K_m) and V_{max} of peptide substrates for CHIKV, ONNV and VEEV were determined in the cell-free FRET-based protease assay. The peptide substrates were serially diluted two-fold for up to 8 times, where concentrations tested ranged from 20 μ M to 0.3125 μ M. To measure K_m of substrates, 1 μ L of serially diluted substrates was added to their respective wells containing 1 μ L nsP2 protein in 99 μ L reaction buffer. Continuous FRET measurement was carried out with an excitation wavelength of 340 nm and emission wavelength of 490 nm at 30°C for 1 hr using a Infinite M200 microplate reader (Tecan). All data points were performed in duplicates and the experiments were performed twice. The initial reaction velocity (V_i) was determined by monitoring the change in the RFU over time. The amount of product formed over time was calculated using the FEC. The V_i obtained was plotted against substrate concentration $[S]$, and the data were fitted to the Michaelis-Menten equation: $[V_0 = (V_{max} \times [S]) / (K_m + [S])]$ using Graphpad prism 7 software for the determination of the kinetic parameters including V_{max} and K_m values. The K_{cat} was calculated by dividing V_{max} with the nsP2 concentration.

2.2.8. Compound screening on CHIKV nsP2 protease

The screening of a library containing 1000 compounds was performed using the established cell-free FRET based protease assay as described above. Briefly, in a flat-bottomed black 96-well microplate (Greiner), 0.5 μ M of CHIKV nsP2FL protein and 100 μ M compound were mixed in protease reaction buffer, and pre-incubated for 10 min at 22°C. Then CHIKV peptide substrate was added to a final concentration of 1 μ M and continuous FRET measurement was carried out with an excitation wavelength of 340 nm and emission wavelength of 490 nm at 30°C for 1 hr using a Infinite M200 microplate reader (Tecan). The percent activity was calculated using the formula: % activity = $(\text{RFU}_{\text{sample}} - \mu\text{RFU}_{\text{DMSO}}) \times 100\%$, where μ represents the mean value, and the $\text{RFU}_{\text{sample}}$, RFU_{DMSO} , are the measured RFU of the CHIKV-infected HEK293T cells treated with test compounds, 1% DMSO vehicle (negative control) respectively.

2.2.9. Thermal shift assay

A 25 μ L reaction mix comprising of 0.8 μ g (0.8 μ M) protein, with or without 0 – 2.5 mM compound and SYPRO Orange dye (Invitrogen) in 1X assay buffer (50 mM HEPES (pH 7.4), 50 mM Arginine, 300 mM NaCl, 5% (v/v) glycerol, 2 mM DTT) was prepared in a 96-well white opaque plates (Bio-Rad). The plate was sealed and centrifuged briefly before heating at 0.5°C increments from 20°C to 85°C using the CFX96 Touch Real-Time PCR Detection System (Bio-Rad). Fluorescence was detected at excitation and emission wavelengths of 485 nm and 625 nm respectively,

and signals were recorded as RFU with respect to temperature. The derivatives ($-dRFU/dT$) were plotted using Graphpad Prism 7.0 software.

2.2.10. IC₅₀ determination

The half-maximal inhibitory concentrations (IC₅₀) of promising compounds was determined by subjecting them to dose response testings. In a 96-well black plate (Bio-Rad), 1 μ M of CHIKV nsP2 protein and serially diluted compounds (7-point, two-fold serially diluted compounds from 0-100 μ M) were mixed in protease reaction buffer and pre-incubated for 10 min at 22°C. Proteolytic reactions were initiated with the addition of 1 μ M peptide substrates and continuous FRET measurement was carried out with an excitation wavelength of 340 nm and emission wavelength of 490 nm at 30°C for 1 hr at 1 min intervals using a Infinite M200 microplate reader (Tecan). Data was fitted to the four parameter equation and IC₅₀ curves were plotted using Graphpad Prism 7.0 software. All data points were measured in duplicates.

2.2.11. *In vitro* antiviral assay

HEK293T cells were seeded at 30,000 cells/well in 96-well tissue culture plates (Greiner). After 2 hrs adhesion, serially diluted compounds were added to the cells and the cells were infected with Gluc-expressing viruses at a MOI of 0.1. After 16 hr of incubation at 37°C in 5% CO₂, 50 μ L of culture supernatant was used to assess viral replication through measuring Gluc activity using the luminometer function on a Promega GloMax Multi Detection plate reader. Each concentration of the hit

compound was tested in triplicate. Percent infection inhibition was determined by normalizing the percentage of infected cells in compound treated cells with mock treated (1% DMSO) and virus infected cells. The EC_{50} , defined as the effective concentration resulting in a 50% inhibition of virus infection, was used to evaluate compound activity.

2.2.12. Cell viability assay

In parallel, mock-infected HEK293T cells were treated with varying concentration of the compounds (0-100 μ M) and incubated at 37°C in 5% CO_2 . At 16 hr posttreatment, cells were lysed, and cell viability was measured using a Cell Titer Glo cytotoxicity kit (Promega) according to the manufacturer's instructions. The concentration of compound that reduced the viability of HEK293T cells to 50% of the control was estimated as the 50% cytotoxic concentration (CC_{50}).

2.2.13. Murine model of CHIKV-induced joint footpad inflammation

3-week old male C57BL/6 mice were infected with SGP11 CHIKV by subcutaneous inoculation of 10^6 PFU of CHIKV virions (in a final volume of 30 μ L PBS) in the ventral side of the right hind footpad, towards the ankle. Animals were treated with 10 mg/kg of compounds or 10% DMSO daily by intraperitoneal injection from 1 to 6 dpi. Disease progression in terms of footpad swelling and viremia were monitored on a daily basis over the next two weeks. Disease score based on the degree of joint footpad swelling around the peri-metatarsal area of the infected footpad was measured. Basically, height and width of the footpad were measured

using a pair of Vernier calipers and the degree of swelling was expressed as the relative folds increase in the footpad size (height X width) compared to the pre-infected state using the formula – $[(\phi - \psi)] / \psi \times 100$, where ϕ refers to the daily size of the footpad and ψ refers to the footpad size at the pre-infected state.

Viremia was determined from the 10 μ L of blood obtained from the tail vein. Briefly, 10 μ L of blood was obtained via tail-snip and mixed thoroughly with 10 μ L of Citrate-phosphate-dextrose solution (Sigma-Aldrich) in 120 μ L of PBS to prevent blood from clotting. Viral RNA was then extracted and quantified with qRT-PCR as described below.

2.2.14. Viral RNA extraction

Viral RNA was extracted from 140 μ L of samples with QIAmp Viral RNA Mini Kit (QIAGEN), following manufacturer's protocol.

2.2.15. CHIKV viral RNA quantification

Viral load or viremia was quantified via quantitative reverse transcription PCR (qRT-PCR) utilizing the QuantiTect Probe RT-PCR Kit (QIAGEN) in a final reaction volume of 12.5 μ L. All reactions were performed using the 7900HT Fast Real-Time PCR System machine (Applied Biosciences) with the following thermal cycling conditions: (1) Reverse transcription at 50°C for 10 min, (2) PCR activation at 95°C for 5 min and (3) 45 cycles of initial denaturation at 94°C for 15 sec followed by a combined annealing and extension step at 55°C for 1 min. Amount of CHIKV-specific negative sense nsP1 (Plaskon et al., 2009) RNA transcripts was detected and

absolute quantity was estimated from a standard curve generated from serially diluted synthetic CHIKV nsP1 RNA transcripts. Limit of detection of the assay is 10^2 RNA copies/ μ L. Sequences of primers and probes used are listed in Appendix 2.

2.2.16. Protein crystallization screen

Crystallization screenings were performed with screening buffers from various commercial manufacturers (INDEX, PEGlon; QIAGEN, JCSG Plus, Morpheus; Molecular Dimensions). Pure recombinant protein was concentrated up to different concentrations e.g. 10 mg/mL or 25 mg/mL and then used for crystallization set up using the mosquito® HTS (TTP Labtech) liquid handler. 100 nL of protein was mixed with either 50 nL, 100 nL or 200 nL of screening buffer using the sitting drop vapor diffusion method and incubated at 20°C. Those drops were analyzed at regular intervals for the appearance of any crystal growth under brightfield microscope.

ONNV-vc002 crystal form 3 was grown using the hanging drop vapor diffusion method at 20°C. The protein solution (26 mg/mL) was mixed in a 1:1 and 2:1 ratio with crystallization buffer on a clean cover slip, inverted and placed over a well of a siliconized 24-well plate containing 1 mL of crystallization buffer. The crystallization buffer was composed of 1.1 M sodium malonate pH 7.0, 0.1 M HEPES pH 7.0 and 0.5% (v/v) Jeffamine ED-2001 pH 7.0. The plate was incubated at 20°C until crystal growth was observed.

2.2.17. Crystal soaking and freezing

Prior to data collection, crystals were looped and transferred to the corresponding crystallization buffer solution that is supplemented with 20% (v/v) glycerol as cryo-protectant, followed by flash-freezing in liquid nitrogen.

2.2.18. X-ray data collection and structure refinement

Frozen crystals were diffracted and the X-ray diffraction data were collected on the beamlines PXII at Swiss Light Source (Paul Scherrer Institut, Switzerland) and MXII at Australian Synchrotron (Melbourne, Australia). The structure of ONNV nsP2 protease was solved by molecular replacement with the program PHASER (Storoni et al., 2004) using CHIKV nsP2 protease structure (PDB ID: 3TRK) as the search model. Structure refinement was carried out using macromolecular refinement packages such as REFMAC5 (Murshudov et al., 2011) and BUSTER (Blanc et al., 2004), and the structure was checked and manually improved in Coot (Emsley & Cowtan, 2004). Figures were prepared using the Pymol software package (Schrodinger, 2015).

2.2.19. Small angled X-ray scattering and data analysis

X-ray scattering data were collected, according to standard procedures, on the Bruker NANOSTAR SAXS instrument equipped with a MetalJet x-ray source and Vantec 2000 detector system. The x-ray radiation was generated from a liquid gallium alloy with microfocus x-ray source ($K = 1.3414 \text{ \AA}$ with a potential of 70 kV and a current of 2.857 mA). The

scattering patterns from CHIKV-vc038 at protein concentrations of 20 $\mu\text{g/mL}$ and 21 $\mu\text{g/mL}$ were measured at 15 °C with a sample volume of 40 μL in a vacuum tight quartz capillary. The buffer containing 20 mM HEPES (pH 7.4), 150 mM NaCl, 2 mM DTT and 5% (v/v) glycerol was measured twice. For each sample, a total of six measurements at 5-min intervals were recorded. The data were normalized to the intensity of the transmitted beam, and the scattering of the buffer was subtracted as background. The experimental data obtained for the two concentrations were analysed for aggregation and folding state using Guinier and Kratky plots respectively. The molar mass was evaluated by the Porod volume and radius of gyration (R_g) using the Guinier approximation. All data-processing steps were performed with the program package PRIMUS (Konarev et al., 2001). The distance distribution function, $p(r)$, of the particle was computed by the GNOM indirect transform package (Svergun, 1994), which also gives the maximal particle diameter, D_{max} . Two low-resolution models for the CHIKV-vc038 molecule were built using the programs DAMMIN (Svergun, 1992) and GASBOR (Svergun et al., 2001). A dimer model was generated using CORAL (Petoukhov et al., 2012) allowing for the flexibility for the linker between the N-terminal helicase and C-terminal protease domains.

2.2.20. Negative staining and electron microscopy image analysis

CHIKV nsP2FL protein was diluted to a final concentration of 20 $\mu\text{g/mL}$ in buffer containing 20 mM HEPES (pH 7.4), 150 mM NaCl, 2 mM DTT and 5% (v/v) glycerol. Copper coated carbon EM grids (200 mesh) were first

glow-discharged for 60 sec (PDC-32G, HARRICK). 4 μ L of diluted protein sample was applied onto the glow-discharged carbon surface for 1 min at room temperature and dry blotted with whatman paper. Following blotting, 4 μ L of freshly prepared 2% (wt/v) uranyl acetate solution was applied onto the grid for 1 min. Uranyl acetate was blotted off the grid and the grid was allowed to air dry for 10 min before data collection or for grid storage. Electron microscopy was performed using a T12 cryo-electron microscope (FEI) operating at 120 keV equipped with FEI Eagle 4k x 4K CCD camera. Images were collected at a defocus of -1.2 μ m to -1.6 μ m and at a dose of 25.6 e-/Å², at magnification of 60 000 (2.11 Å/pixel). A total of 1000 particles was selected for 2D projection analysis. Particles were picked and processed using EMAN2 (Tang et al., 2007).

2.2.21. Statistical analyses

Statistical analyses were performed using Prism 7 (GraphPad Software). Comparisons between the different groups in both *in vitro* and *in vivo* experiments were performed using either two-way ANOVA or Mann Whitney U test. P-values considered statistically significant are represented with * for $P < 0.05$, ** for $P < 0.01$, and *** for $P < 0.001$.

CHAPTER 3. Functional and structural characterization of alphavirus nsP2

3.1. Introduction

3.1.1. Importance of N-terminal nsP2 in 2/3 site processing

Alphavirus nsP2 is the largest (90 kDa) multidomain protein consisting of the N-terminal helicase and C-terminal protease domains, whereby they possess multiple enzymatic activities that are essential for virus replication. One of the key events of the infection process is ns polyprotein processing by nsP2 protease, where it is directly linked to the regulation of the virus replication cycle. The nsP2 protease recognizes and cleaves all three cleavage sites, located between nsP1 and nsP2 (1/2 site), nsP2 and nsP3 (2/3 site), and nsP3 and nsP4 (3/4 site), at different efficiencies. In the early infection stage, ns polyprotein is cleaved in *trans* at the 3/4 site, then by an in *cis* cleavage of 1/2 site and rapidly followed by an in *trans* cleavage of 2/3 site (Jose et al., 2009). The processing of the 2/3 site is extremely important during this infection stage as it triggers the cessation of negative strand RNA synthesis (Lemm et al., 1998, 1994). However, the sequential order and exact timing of processing events are not solely dependent on nsP2 protease but a result from the activity of a complex machinery of replication. For example, the N-terminal domain of nsP2 was revealed to be an essential player in 2/3 site processing as it functions as an "activator" (Lulla et al., 2012, 2006; Rausalu et al., 2016). It should also be noted that the RNA helicase activity of the N-terminal domain of nsP2 required the presence of the C-terminal protease domain (Das et al., 2014). These findings led to the hypothesis that the internal

architecture of nsP2 assumes various conformational arrangements to not only ensure protein-protein interactions during ns polyprotein processing but also to modulate the functional cross-talk between the nsP2 domains.

At present, the crystal structures of the individual domains are available but the conformational arrangement of the full length nsP2 is still lacking. Therefore, structural studies are needed to understand the molecular basis for *cis* cleavage activity and to present firm structural evidence of interactions between nsP2 domains. This could reveal the dynamic properties of the nsP2 molecule and also provide insights into interdomain interactions that could be associated to its various enzymatic activities during ns polyprotein processing and RNA replication. Apart from understanding these processes, the elucidation of the structure of nsP2 molecule could also provide an alternative target for antiviral drug discovery and improve development of antivirals against CHIKV and other pathogenic alphaviruses in the future.

3.1.2. Importance of revealing nsP2 structure of closely-related alphaviruses

The crystal structures of CHIKV, SINV and VEEV nsP2 proteases are available and they are often used as models for designing therapeutics that target the catalytic dyad and substrate binding site. Despite sharing low sequence identity (Figure 3.1A), as mentioned in the introduction section, their catalytic site residues remain conserved and their structure is similar (Figure 3.1B). However, it appears that amino acid differences

in the loop located close to the catalytic residues led to structural and surface charge potential differences (Figure 3.1C). Obviously, these notable differences shown in the substrate binding region contributed to the substrate recognition variances observed between the three alphaviruses. Interestingly, the structure of ONNV nsP2 protease, which is the most closely related to CHIKV, is not yet resolved. The nsP2 proteases from the two alphaviruses recognize the same substrate sequences located along their ns polyprotein. While it is tempting to conclude that they will display high structural similarities due to their high sequence identity and recognition of similar substrate sequences, it cannot be excluded that amino acid differences surrounding the substrate binding region could impact their substrate affinities. Structural comparison studies could therefore reveal any divergences that could give rise to differences in ns polyprotein processing rates and consequently disease outcomes. Given that ONNV is emerging as a clinically important alphavirus, it is imperative that efforts are also focused on revealing its nsP2 protease structure. This knowledge will be invaluable and can be useful in the future for designing new drug candidates targeting nsP2 protease.

3.1.3. Structural biology techniques to elucidate viral protein structures

All of the structures of the viral proteins described above were solved by X-ray crystallography, which is the most commonly used method for obtaining high-resolution structural information about biological macromolecules. Although it is a powerful technique, the method required large quantity of highly pure, mono-dispersed protein solution and the growth of good crystals that can be analysed using X-ray diffraction. Aside from X-ray crystallography, other structural methods such as nuclear magnetic resonance, high-resolution cryo-electron microscopy (EM) and small-angle X-ray scattering (SAXS) can be used. Furthermore, integrative structure determination approach, which involved the simultaneous use of complementary techniques, is often used now to achieve a complete and accurate structure. When combined with biophysical, biochemical analysis and cellular studies, they could give a more holistic view of a protein's function and mechanism. As an example, Luo et al. combined the use of X-ray crystallography and SAXS in solution to resolve the structure of DENV NS3 and to assess the quaternary conformations that are likely to be present in solution for the NS3 molecule (Luo et al., 2008).

3.1.4. Study approach

The focus of this chapter is to resolve the 3D structures of full length CHIKV nsP2 and ONNV nsP2 protease. Multiple CHIKV nsP2 constructs, which include full length, protease and helicase fragments, and other

constructs of closely related alphaviruses (ONNV, RRV, and VEEV), were cloned, expressed and purified. The proteins of the individual domains are included in the study to serve as positive controls for the functional assays. Biochemical assays such as cell-free FRET-based protease and NTPase activity assays were developed and used to validate the activities of the purified proteins. Structural determination techniques such as negative stain Electron Microscopy (EM) and Small-angle X-ray scattering (SAXS) were employed to elucidate the global architecture of nsP2 and X-ray crystallography was used to reveal the high resolution structure of ONNV nsP2 protease.

3.2. Results

3.2.1. Generation of nsP2 protein constructs

To obtain better characterization of nsP2 and increase the success rate of crystallization, multiple protein constructs were submitted to NTU's Protein Production Platform (PPP) for cloning and protein expression. In total, 52 CHIKV, six ONNV, seven RRV and six VEEV constructs with different start and stop positions and His6 tag orientations were submitted and subjected to small scale protein expression (Figure 3.2). 13 out of 53 CHIKV and 5 out of 6 ONNV constructs were soluble (Figure 3.2A-B). None of the RRV and VEEV constructs were soluble enough to be scaled up for purification (Figure 3.2C-D), and thus were excluded from the study. Here, CHIKV-vc038 (full-length nsP2, encompassing amino acids 2-795), CHIKV-vc039 (helicase domain alone, encompassing amino acids 4-470) and CHIKV-vc045 (protease domain alone, encompassing amino acids 469-791) constructs were chosen (Appendix 1) and subjected to a 2-step purification involving Ni-NTA batch purification and size-exclusion chromatography. The proteins were eluted in final buffer containing 20 mM HEPES (pH 7.4), 150 mM NaCl, 2 mM DTT and 5% (v/v) glycerol. CHIKV-vc038, which is 90 kDa in size, was purified through a HiLoad® Superdex 16/60 200 column and it eluted at 82 mL as a single peak. CHIKV-vc039, which is 55 kDa in size, and CHIKV-vc045, which is 39 kDa in size, were purified through a HiLoad® Superdex 16/60 75 column and they eluted at 56 mL and 65 mL respectively, as single peaks. These results indicated that all protein samples purified as monomers (Figure 3.3).

A

B

C

D

Figure 3.2. nsP2 protein constructs. (A) 52 CHIKV, (B) six ONNV, (C) seven RRV and (D) six VEEV constructs were designed. Abbreviation: CHIKV, Chikungunya virus; ONNV, O'nyong'nyong virus; RRV, Ross River virus; VEEV, Venezuelan Equine Encephalitis virus; His6, histidine-tag; TEV, Tobacco Etch Virus protease recognition site (amino acid sequence: SSGVDLG TENKYFQ\ S). Legend: Red, no visible expression; Yellow, target protein present but at lower amounts; Orange, target protein present but at medium levels; Green, target protein present and significantly stronger than background.

Figure 3.3. Size exclusion profiles of CHIKV-vc038, CHIKV-vc039, CHIKV-vc045. (A) CHIKV-vc038 was purified through a HiLoad® Superdex 16/60 200 column and eluted as a single peak at ~82 mL, estimated to be 90 kDa in size. (B) CHIKV-vc049 and (C) CHIKV-vc045 were purified through a HiLoad® Superdex 16/60 75 column and eluted as single peaks at 56 mL and 65 mL respectively. They were estimated to be 55 kDa and 39 kDa in size respectively.

3.2.2. Validation of nsP2 protease and ATPase activities of purified nsP2 proteins

Before subjecting the proteins to structural analyses, protease and ATPase activity assays were performed to verify that they are functional, which will further indicate that the proteins are properly folded and are suitable for structural studies. Firstly, to assess the proteolytic activity of CHIKV-vc038, CHIKV-vc039, and CHIKV-vc045, a cell-free FRET-based protease assay was set up and optimized for this purpose. This assay is a highly sensitive and quantitative fluorometric assay that has been widely used in the measurement of protease activity for a variety of enzymes such as human immunodeficiency virus-1 aspartic protease, Hepatitis C (HCV) NS3 serine protease and other alphavirus proteases (Aggarwal et al., 2014, 2015; Matayoshi et al., 1990; Saha et al., 2018; Taliani et al., 1996). The synthetic substrate was designed based on the nsP1/nsP2 junction sequence of CHIKV, where an 8-residue peptide is flanked by the fluorophore-quencher FRET pair, DABCYL and EDANS, at the N- and C-terminus respectively (Figure 3.4A, Appendix 2). At close proximity, the EDANS fluorophore is quenched since its emission spectrum (E_m) is completely overlapped by the absorbance (Abs) spectrum of DABCYL (Figure 3.4B). In the presence of active protease, the substrate is cleaved at the scissile bond between Ala-Gly, separating the fluorophore-quencher pair and resulting in the increased fluorescence signal. The assay was performed in 100 μ L reaction volumes containing 1 μ M of protein and 1 μ M of substrate in HEPES buffer, and the fluorescence was monitored over 60 min at 30°C. Given that CHIKV-vc038 and CHIKV-

vc045 both contain the protease domain, they demonstrated efficient cleavage of the substrate peptide. Interestingly, it was observed that the CHIKV-vc038 full length nsP2 protein was twice more active than CHIKV-vc045 in cleaving peptide substrate (Figure 3.4C). As expected, CHIKV-vc039 did not display any protease activity since it is lacking the protease domain. Furthermore, the addition of CHIKV-vc039 to CHIKV-vc045 did not enhance the protease activity, indicating that the protease activity of CHIKV nsP2 is not dependent on the presence of the helicase domain.

Secondly, to assess the helicase function of the purified nsP2 proteins, a biochemical analysis of the NTPase activity of helicase was performed. Nucleotide triphosphate (NTP) hydrolysis is known to provide energy required for the motor activity and translocation of helicases on single stranded RNA, and this activity can be quantified using a sensitive colorimetric assay. The assay measures the total amount of orthophosphate released, based on the green complex of phosphomolybdate and malachite green. The assay was performed in 100 μ L reaction volumes, using standard 96-well plate. Adenosine triphosphatase (ATPase) activity of the CHIKV constructs was carried out at 1 nM protein concentration, 1 mM ATP concentration and at different $MgCl_2$ concentrations (0–2 mM). As expected, those constructs that include the helicase domain (CHIKV-vc038 and CHIKV-vc039) showed an increase in the released phosphate when 2 mM $MgCl_2$ was added (Figure 3.5). This observation correlates with previous reports where NTPase activity was demonstrated to be completely dependent on Mg^{2+}

ions (Karpe et al., 2011; Karpe & Lole, 2010). Overall, all constructs examined were active and suitable for structural analyses.

Figure 3.4. Schematic representation of mechanism of EDANS-DABCYL FRET pair, using CHIKV peptide substrate as a representative. (A) The CHIKV peptide substrate is flanked by the donor-quencher FRET pair, DABCYL and EDANS, at the N- and C-terminus respectively. In the absence of CHIKV nsP2 protease, the donor-quencher pair are in close proximity to one another. The energy emitted from EDANS is quenched by DABCYL, resulting in low or no fluorescence. When they are separated upon substrate cleavage by nsP2 protease (cleavage site indicated by red arrow), the fluorescence will increase, and the fluorescence intensity change can be detected continuously and directly. **(B)** Representation of the spectral properties of EDANS-DABCYL pair. The absorbance wavelength (Abs) of DABCYL overlaps the emission (Em) spectra of EDANS. **(C)** Substrate cleavage by the various CHIKV nsP2 proteins. 1 μ M of protein was incubated with 1 μ M of peptide substrate and FRET was continuously measured over 60 min using the excitation wavelength of 340 nm and emission wavelength of 490 nm. All data points were measured in duplicates.

Figure 3.5. ATPase activity of CHIKV nsP2 helicase is dependent on the presence of MgCl₂. CHIKV nsP2 proteins (1 nM) were incubated in a 100 μ L reaction mixture containing 50 mM MOPS at pH 7.25, 0–2 mM MgCl₂ and 1 mM ATP, at 37°C for 30 min. Released orthophosphate was quantitated as described in the experimental procedures. Data are representative of two independent experiments and presented as mean $\pm$ SD. Statistical analyses were performed using two-way ANOVA with Bonferroni posttest (***) $p < 0.001$.

3.2.3. Low-resolution structure of CHIKV-vc038

Having shown that the purified proteins are active, crystallization screens were performed for CHIKV-vc038 (full length nsP2). However, attempts to obtain diffraction-quality crystals were not successful. Given that the protein consists of two domains and is flanked by a flexible linker, it is likely non-rigid and could be the reason for the failed crystallization attempts. Instead, an alternative approach for CHIKV-vc038 was pursued. Since the protein's overall molecular size is ~90 kDa and is within the range of small angle X-ray scattering (SAXS) and negative stain electron microscopy (EM), both techniques were used to assess its overall structure.

Although SAXS techniques produce low-resolution structures, they can give insights into the quaternary conformations that are likely to be present in solution for the CHIKV-vc038 molecule. SAXS patterns from CHIKV-vc038 were recorded and processed as described in the Methods section. Briefly, a monodispersed solution of CHIKV-vc038 was prepared by diluting the concentrated protein to approximately 20 µg/mL, followed by exposure with in-house high energy X-rays. The X-rays' scattering patterns were recorded and radially averaged to obtain a 1-dimensional plot (1D) that contained information on the intensity of scattered light with respect to the scattering angle. As buffer molecules also scatter light, the contributions from these molecules were subtracted to obtain a scattering pattern of CHIKV-vc038. The 1D scattering plot was converted to the electron pair-distance distribution plot ($P(r)$ plot), which provided the radius of gyration and maximum particle dimension of CHIKV-vc038. This

plot was used as the input file for the *ab initio* modeling packages (i.e., GASBOR and CORAL) to obtain low-resolution structure of CHIKV-vc038.

SAXS patterns of CHIKV-vc038 recorded at 20 and 21 $\mu\text{g/mL}$, representing two independent experiments, showed similar final composite scattering curves (Figure 3.6A). Furthermore, the six datasets measured for each sample were similar, indicating the absence of radiation damage. The Guinier plots at low angles were linear and revealed good data quality with no indication of protein aggregation (Figure 3.6B). The radius of gyration (R_g) and maximum particle diameter (D_{max}) values were 35 Å and 80.88 Å respectively, suggesting that the molecule was rather elongated. In addition, the Kratky plot showed a bell-shaped peak at low angles indicating a well-folded protein (Figure 3.6C). The estimated molecular mass calculated from R_g value and Porod volume was 91.7 kDa and, which indicated that the protein existed as a monomer in solution. A qualitative analysis of the distance distribution function suggested that CHIKV-vc038 consisted of a major part yielding a principal maximum in the $p(r)$ function at around 40 Å (Figure 3.6D). *Ab initio* reconstruction of the gross shape of CHIKV-vc038 from the scattering pattern using the shape determination program DAMMIN and the dummy residue modelling program GASBOR yielded similar results (Figure 3.6E). The rigid body model, which was computed using the program CORAL, was arbitrarily constructed based on the 3D structures and average sizes of the helicase and protease domains. Analysis of the aligned *ab initio* and rigid body model demonstrated that the helicase and protease domains are far apart from each other (Figure 3.7).

Apart from discerning the global architecture of CHIKV-vc038 through SAXS, it was important to elucidate the structure at the atomic level so that we can fully understand the intra-molecular domain organisation. Thus, CHIKV-vc038 was subsequently subjected to negative stain EM. This is a powerful tool in EM as it gives a higher image contrast and allows much smaller molecules (< 250 kDa) to be visualized. Moreover, negative staining tends to induce one or a few preferred orientations of the molecules on the carbon support film and this simplifies the image classification process. Briefly, diluted CHIKV-vc038 (~20 µg/mL) was adsorbed onto a glow-charged carbon-coated EM grid, dry-blotted and stained with 2% (w/v) uranyl acetate solution. The sample was then imaged using Tecnai T12 electron microscope. Images obtained showed weak contrast between the background and the particle, since stain clouds around the molecules were not as evident (Figure 3.8). Attempts were made to improve the image contrast, such as modifying the adsorption time and staining time, but they did not yield better micrographs. Despite the poor image quality, approximately 1000 single particles were picked and two-dimensional (2D) class averaging was performed through the EMAN2 software. Classification of particles selected from the micrographs revealed that the molecule adsorbed to the grid in random orientations and it was difficult to distinguish between molecules in different orientations and in different conformations. Thus, it is essentially impossible to obtain a reliable 3D reconstruction of CHIKV-vc038 using this set of negative stain images. Nonetheless, the basic configuration of the particle based on the 2D class averages showed two

distinct blobs, which were likely to be that of the two domains seen from the side view. The particle was estimated to be approximately 9 nm in diameter, likely representing single molecules of the nsP2 protein. The diameter of the particle was close to the D_{max} value obtained from SAXS (80.88 Å), indicating that the data from SAXS and negative stain EM were constant.

Figure 3.6. Small Angle X-ray Scattering of CHIKV-vc038. (A) Experimental scattering data (I_q vs q) for CHIKV-vc038. Illustration of sample and background raw data for CHIKV-vc038. (B) Guinier plot ($\log(I_q)$ vs q^2), indicating the linear region for determination of the radius of gyration using data at low-scattering angle (highlighted in yellow). (C) Kratky plot suggests that CHIKV-vc038 is well folded. (D) The electron pair-distance distribution curve $P(r)$ determined from the scattering data suggest an elongated shape of CHIKV-vc038. (E) Comparison of DAMMIN (left) and GASBOR (right) models. All models were visualized using the software Pymol.

Figure 3.7. Low resolution model of CHIKV-vc038 obtained from SAXS data using a combination of *ab initio* calculations and rigid body modeling. Superposition of the *ab initio* bead model of CHIKV-vc038 derived from solution X-ray scattering (grey spheres; computed by using the program GASBOR) with the rigid body model of CHIKV helicase and protease crystal structures (N-terminal helicase is colored green, C-terminal protease is colored yellow and the flexible linker region is colored magenta; computed by using the program CORAL with input PDB accession numbers 6JIM and 3TRK for CHIKV helicase and CHIKV protease respectively). The rigid body model was arbitrarily constructed and then spatially aligned to the bead model, according to the shapes and average sizes of the helicase and protease domains. The model was generated using the software Pymol.

Figure 3.8. Single particle analysis of CHIKV-vc038. Representative negative stain micrograph of CHIKV-vc038. Scale bar at the bottom: 100 nm. Some of the resulting 2D class averages obtained from ~1000 particles (white circles) are displayed in the inset below the micrograph. The side length of the individual panels is 10 nm. The single particles were picked and 2D class averages were performed using the EMAN2 software. The particle has an average size of 9 nm.

3.2.4. Purification, crystallization and data collection of ONNV-vc002

The ONNV-vc002 (protease domain), encompassing amino acids 471-792, was purified to homogeneity by Ni-NTA affinity chromatography and size-exclusion chromatography (Figures 3.9A-B). The molecular weight of the purified ONNV-vc002 was calculated from the standard graph of protein molecular weight markers, which were ran on Superdex 75 gel-filtration column. The protein, which eluted at 66 mL, was determined to be ~39 kDa, indicating that it existed as a monomer in solution. Additionally, a single band of ~39 kDa corresponding to the molecular weight of ONNV-vc002 was observed on SDS-PAGE (Figure 3.9B (insert)). The purified ONNV-vc002 was concentrated to 26 mg/mL and its activity was tested prior to structural analysis. A FRET-based protease assay was performed on the purified protein, using CHIKV peptide as substrate. Indeed, ONNV-vc002 was active, although it showed slightly lower levels of protease activity to that of CHIKV-vc045 (Figure 3.10). This observation may likely be attributed to the quality of the protein batch. Next, ONNV-vc002 was subjected to crystallization screen using the sitting drop diffusion method. Diffraction-quality crystals were obtained in five different conditions at 20°C between 3 to 7 days, with the crystallization conditions listed in Table 3.1. These protein crystals reached to a size of between 5 to 50 μm . For crystal form 3, the crystal were reproduced in the hanging drop vapor diffusion method. Nonetheless, crystals from all five conditions were sent for data collection (Figures 3.11A-E).

Crystal form 2 gave twinned datasets despite data being collected at two different beamlines (Figure 3.11B). This is likely an intrinsic characteristic of the crystal itself. Crystal forms 4 and 5 diffracted at more than 4 Å; thus, data was not collected. In crystal form 3 (Figure 3.11C), a 98.3% complete dataset was collected on the PXII Beamline at Swiss Light Source and structure was determined at a resolution of 2.92 Å. ONNV-vc002 crystal form 3 belongs to the orthorhombic space group $P12_11$ with unit cell parameters $a = 140.58$, $b = 109.59$, $c = 141.75$, and $\alpha = 90^\circ$, $\beta = 104.65^\circ$, $\gamma = 90^\circ$. The structure was built up with eight molecules per asymmetric unit (ASU). The complete data collection statistics and corresponding refinement statistics are summarized in Table 3.2. The 3D structure of ONNV-vc002 was determined using molecular replacement method by taking the crystal structure of the single subunit of CHIKV nsP2pro as a search model (PDB ID: 3TRK). The final model refined to an R_{factor} and R_{free} of 20.1% and 23.5%, respectively, and contained 2552 residues and 703 water molecules. In the final model, the first and last clearly visible residues were Phe471 and Tyr789, respectively.

Figure 3.9. SDS PAGE and gel filtration profiles of ONNV-vc002. (A) Visualization of Coomassie-stained SDS PAGE of eluted fractions from Ni-NTA beads. (B) Eluted fractions from Nickel bead batch purification was injected into a HiLoad® Superdex 16/60 75 column and eluted as a single symmetrical peak at 66 mL, which corresponded to ~39 kDa in size. The size-exclusion chromatography result suggested the monomeric nature of the protein and it purified to homogeneity. The protein was > 95% pure, as determined through visualization of Coomassie-stained SDS PAGE (right insert).

Figure 3.10. Comparison of protease activities between purified ONNV-vc002 and CHIKV-vc045. The protease assay was performed in 100 μ L reaction volumes containing 1 μ M of either ONNV-vc002 or CHIKV-vc045 and 1 μ M of peptide substrate in HEPES buffer. FRET measurements were carried out after 60 min incubation at 30°C. The protease activity was calculated using the formula: Protease activity = $RFU_{\text{sample}} - RFU_0$, where RFU_{sample} is the measured RFU of reactions with protein and RFU_0 is the measured RFU of the reaction without protein. The experiment was performed twice, each having duplicates. Data is presented as mean $\pm$ SD and representative of two independent experiments.

Table 3.1. Crystallization conditions of ONNV-vc002

Crystal form	Protein purification buffer	Crystallization condition
1	HEPES buffer [#]	0.8 M Succinic acid pH 7.0
2	HEPES buffer	2.8 M Sodium acetate trihydrate pH 7.0
3	HEPES buffer	1.1 M Sodium malonate pH 7.0, 0.1 M HEPES pH 7.0, 0.5% v/v Jeffamine ED-2001 pH 7.0
4	HEPES buffer	1.0 M Succinic acid pH 7.0, 0.1 M HEPES pH 7.0, 1% PEG monomethyl ether 2000
5	HEPES buffer	1.0 M Sodium citrate tribasic dihydrate, 0.1 M Sodium cacodylate pH 6.5

[#] HEPES buffer: 20 mM HEPES, 300 mM NaCl, 5% (v/v) glycerol and 2mM TCEP, pH 7.5

Figure 3.11. Crystal forms obtained from screening of ONNV-vc002. Crystal (A) form 1, (B) form 2, (C) form 3, (D) form 4 and (E) form 5. 26 mg/mL of purified ONNV-vc002 was mixed in a 1:1 and 2:1 ratio with crystallization buffers of multiple screening kits. Drops were set up using the sitting drop diffusion method and incubated at 20°C between 3 to 7 days. The crystallization buffer conditions of the respective crystals are summarized in Table 3.1.

Table 3.2. Data collection and refinement statistics of crystal form 3 of ONNV-vc002

Parameters	Value for ONNV-vc002
Data collection statistics	
X-ray source	SLS PXII beamline
Wavelength (Å)	1.00
Space group	P12 ₁ 1
Cell dimensions	
a, b, c, (Å)	140.58, 109.59, 141.75
α , β , γ (°)	90.00, 104.65, 90.00
Resolution range (Å)	49.57 – 2.92
Completeness (%)	98.3
R _{merge} (%) (last shell)	18.8 (83.7)
Mean $I/\sigma(I)$ (last shell)	4.57 (0.92)
No. of observed reflections	259244
No. of unique reflections	167056
No. of molecules per asymmetric unit	8
Matthew's coefficient (Å ³ /Da)	2.93
Solvent content (%)	58
Refinement statistics	
No. of residues	2552
No. of water molecules	703
R _{factor} (%)	20.1
R _{free} (%)	23.5
Root-mean-square deviation	
Bond length (Å)	0.01
Bond angle (Å)	1.10
Ramachandran plot (%)	
Favored region	91.78
Allowed region	8.22
Outlier region	0

3.2.5. Crystal structure of ONNV-vc002

The overall structure of each monomer of ONNV-vc002 is similar to the crystal structures of nsP2 protease reported from CHIKV, SINV and VEEV (Figure 3.12). Each monomer of ONNV-vc002 structure consists of 319 residues starting from Phe471 to Tyr789 and it revealed two subdomains which are connected through a flexible linker: an N-terminal protease subdomain possessing catalytic residues and a C-terminal subdomain resembling SAM-dependent MTase (Figure 3.13A). The N-terminal protease subdomain consists of 123 residues from Phe471 to Gln593 and contains seven α -helices that is flanked by two sets of short β -hairpins. Whereas the larger C-terminal MTase-like subdomain extends from Met624 to Tyr789 and contains five α -helices and seven β -strands aligned with each other. The two subdomains are connected through a random coil spanning from Asn594 to Arg623. The residues of the catalytic dyad, Cys478 and His548, are located near the interface of the two subdomains, with the residue Cys478 positioned at the N-terminus of α 1-helix and the second residue His548 located in the loop region between β 1 and β 2-strands. Interestingly, structural comparisons of the ONNV-vc002 active site residues with that of CHIKV, SINV and VEEV nsP2 proteases (PDB IDs: 4ZTB, 4GUA and 2HWK respectively) showed they are positioned in a similar conformation to CHIKV (Figure 3.13B) but are slightly different to that of SINV and VEEV (Figures 3.13C-D). The difference in amino acid composition and/or in length of the loop likely contributed to the slight variations in conformations.

Previous molecular docking and simulation studies of VEEV nsP2 protease with peptide substrates revealed a substrate binding region, which is a deep and pronounced groove traversing the active site that can accommodate a five residue long peptide (Russo et al., 2010). Likewise, structural analysis of ONNV-vc002 revealed that the substrate binding cleft is present at the interface of the two subdomains (Figure 3.13E). A surface view around the active site showed that the Asn547 located on the loop between β 1 and β 2-strands, together with Leu670 found on the loop between β 7-strand and α 9-helix of the MTase-like subdomain forms a tunnel, which is consistent with that seen in CHIKV nsP2 protease (Figure 3.13E). It is not surprising to observe this structural resemblance since Asn547 and Leu670, together with those residues within the substrate binding site, are conserved for the two alphaviruses. Also, given that Asn547 and Leu670 in CHIKV nsP2 protease were proposed to participate in substrate binding and entry respectively, they are likely to have the same function in ONNV-vc002 and it is possible that their substrate affinities do not vary too.

Figure 3.12. Superposition of ONNV-vc002 with CHIKV, SINV and VEEV nsP2 proteases. ONNV-vc002 is colored in yellow, CHIKV nsP2 protease (PDB ID: 4ZTB) is colored in magenta, SINV nsP2 protease (PDB ID: 4GUA) is colored in pale blue and VEEV nsP2 protease (PDB ID: 2HWK) is colored in cyan. The superposition is done using the software Pymol.

Figure 3.13. Crystal structure of ONNV-vc002. (A) The N-terminal protease subdomain, in yellow, contains both the catalytic dyad Cys478 and His548, colored in cyan. The C-terminal MTase-like subdomain is colored in green and the linker connecting the two subdomains is colored in red. All structural elements are labelled from N-to-C terminus. Structural comparisons of the active site residues of ONNV-vc002 with (B) CHIKV nsP2 protease (PDB ID: 4ZTB), (C) SINV nsP2 protease (PDB ID: 4GUA) and (D) VEEV nsP2 protease (PDB ID: 2HWK). ONNV-vc002 is in yellow, CHIKV nsP2 protease is in magenta, SINV nsP2 protease is in pale blue and VEEV nsP2 protease is in cyan. The catalytic residues are shown as sticks. (E) Surface view of ONNV-vc002. Zoom panel shows the closed tunnel, colored in magenta, formed by Asn547 located on the loop between $\beta 1$ and $\beta 2$ strands and Leu670 of the MTase-like subdomain. The structural elements and electrostatic surface were generated using the software Pymol.

3.3. Discussion

Dissecting the interactions between the helicase and protease of alphavirus nsP2 would be informative to elucidate how their interdomain interactions would impact ns polyprotein processing and viral replication. The solved structure could possibly facilitate drug discovery and design targeting the nsP2.

The elucidation of low-resolution structure of the CHIKV full-length nsP2 from this study provides new insights into understanding the versatile and multi-functional roles played by nsP2 in RNA replication and polyprotein processing. It opens up new possibilities for drug discovery and design for antiviral therapeutic intervention. Biochemical studies demonstrated that the full-length nsP2 was more efficient than the truncated protein in hydrolysing the peptide substrate, despite exhibiting comparable ATPase activities. Additionally, the helicase and protease domains could not elicit additive or synergistic enzymatic effects on one another. These observations not only suggested that the domains function independently but may also imply that for the protease to function efficiently, it was necessary for the domains to be connected and exist in close proximity. The flexible linker present between the two domains of the full-length nsP2 may likely play a pivotal role by giving the domains a level of freedom to move relative to each other. This could allow the protein to adopt a conformational arrangement where the distance between the domains was optimal for promoting efficient proteolysis of the substrate. The similar levels of ATPase activity observed for the full length and truncated proteins also seemed to suggest that the helicase

may already exist in its most stable state to efficiently perform its helicase function.

The *ab initio* model from SAXS data combined with the high-resolution rigid structures of helicase and protease revealed that the CHIKV nsP2 demonstrated an elongated overall conformation. The model proposed the absence of intra-molecular interactions between the two domains. This may imply that the RNA-binding pocket of the helicase and the catalytic cavity of the protease domain remained unobstructed in solution and possibly allowed RNA and peptide substrates to access their respective binding sites. The extended conformation could also allow both NTP hydrolysis and proteolysis to function simultaneously. In addition, interdomain flexibility could also be conferred by the linker region. This flexibility allowed functional conformational plasticity of the helicase and protease domains as well as accommodate other host and viral partners in the replication complex. I regardless, the full length CHIKV nsP2 model defined in this Chapter provides an overview of its overall architecture and could be useful as a reference model for future high-resolution structural determination.

As crystallization efforts on full length CHIKV nsP2 remained unsuccessful, attempts were made to reveal the atomic structure through negative stain EM instead. However, the weak contrast combined with the random adsorption of the molecule to the grid made it impossible to obtain a reliable 3D reconstruction. Nevertheless, the 2D class averages displayed an extended conformation which was similar to that observed

in the proposed *ab initio* model. Additionally, the average particle size conformed with that observed in SAXS.

Apart from attempting to resolve the structure of CHIKV full length nsP2, the structure of ONNV nsP2 protease was elucidated. This is the first reported structure of an ONNV viral protein. The N-terminal subdomain, catalytic residues, the substrate binding cleft and C-terminal MTase-like subdomain were distinguishable from the structure. Furthermore, the protease activity together with the secondary structural elements (Figure 3.14) and substrate binding cleft structure were demonstrated to be highly similar to that of CHIKV nsP2 protease. Despite their similarities, this information remains valuable as they further support our knowledge of alphavirus nsP2 proteases.

CHAPTER 4. Identification of small molecules as alphavirus nsP2 protease inhibitors

4.1. Introduction

One of the most attractive antiviral targets is the CHIKV nsP2 protease because (i) viral polyprotein cleavage during replication is required for the maturation of the virus particle and (ii) the protease is one of the most conserved viral protein among the other alphaviruses. Therefore, the likelihood of a single compound to possess pan-alphavirus anti-protease activity is higher than compounds targeting other viral proteins. The CHIKV protease activity resides in the C-terminal of nsP2 (Das et al., 2014), where it regulates the proteolytic processing of the four viral nsP cleavage sites through a cysteine protease mechanism of action (Pastorino et al., 2008). The protease activity of nsP2 is specific for substrates with certain target sequences and is essential for alphavirus replication, making nsP2 an attractive target for the development of therapeutics for alphavirus infection (Chung et al., 2014; Singh et al., 2018).

With the nsP2 protease structure being solved recently, molecular docking and computer-aided drug design have been employed as tools to investigate inhibitory compounds that can potentially target the protease's active site with better specificity and efficacy (Bassetto et al., 2013; El-Labbad et al., 2015; Das et al., 2016; Hu et al., 2016). Combined with the development of a cell-free protease assay, it has facilitated the discovery of peptidomimetics (El-Labbad et al., 2015; Singh et al., 2018) and non-peptidic inhibitors (Das *et al.*, 2016) targeting the active site,

which have been described to display decent inhibitory activity against CHIKV nsP2 protease *in vitro*. However, these compounds have not yet been proven to be potent CHIKV replication inhibitors.

4.1.1. Fragment to lead optimization

Since the first report of fragment-based lead discovery (FBLD) some 20 years ago (Shuker et al., 1996), there are now many fragment-derived compounds in clinical trial phase and two approved drugs that are in use to treat cancer patients (Bollag et al., 2012; Souers et al., 2013). The main feature of fragment methods is to firstly identify low molecular weight (MW) compounds or fragments (average MW 200 Da) that bind to a protein of interest. A diverse library of 1-2000 fragments is usually sufficient to identify some hits. The small size of the fragments is also an advantage as they are more likely to identify functional motifs that are found in small or shallow binding pockets within the target. Although they may be weak binders, they may give a starting point for identifying fragments more suitable for optimization. Then, the fragment hits can be further improved to become higher affinity compounds that can be used as a lead for drug discovery (Figure 4.1) (Lamoree & Hubbard, 2017).

In the early stage of FBLD, fragments that bind to a protein target need to be identified through some form of screening system. Many fragment screening approaches can be applied and the two most widely used are nuclear magnetic resonance (NMR), and X-ray crystallography. These methods are highly sensitive as they can detect fragments in the high affinity range (nM to low μ M). However, both systems require high

quantities of stable protein. Additionally, in the case of X-ray chromatography, the main requirement is for the protein to form a robust crystal that is suitable for soaking with high concentrations of fragment. Nonetheless, solving the crystal structures will not only identify fragment hits but simultaneously yield information about the binding site, which is hugely advantageous. More recently, the development of highly sensitive and robust biochemical assays has made this a favourable system, since it needs much less protein and can successfully report fragment binding and activity.

After identifying hits through one of the above-mentioned methods and they are fully characterized, fragments undergo optimization strategies which heavily relies on medicinal chemistry and the use of available fragment-bound protein crystal structures. Three optimization approaches can be applied to the fragment core, either through fragment linking of two hits, structure-guided fragment growing or merging of multiple hits together (Figures 4.2A-C). The successfully optimized fragment can then be used to probe their effect on the disease or as a lead for drug discovery. One successful example of a direct-acting antiviral candidate that underwent FBLD and is currently in Phase III clinical trial is Dasabuvir (ABT-333), which is a Hepatitis C virus (HCV) NS5B polymerase inhibitor (El Kassas et al., 2017; Shuster et al., 2019). High-throughput screening of fragment libraries for the HCV polymerase using established biochemical assays was central to the identification of the aryl dihydrouracil fragment (Liu et al., 2012). And recently, the fragment-based approach led to the identification of protease inhibitors

against West Nile Virus NS2B-NS3 protease (Schöne et al., 2017).

Although this study is still in its nascent stage, the viral inhibitors identified have the potential to be developed into a drug candidate.

Figure 4.1. Fragment to lead discovery process. A schematic representation of the five main components of a fragment-based discovery platform – fragment libraries, target enablement, fragment screening, generating a model for fragment binding and fragment to lead optimization. Adapted from (Lamoree & Hubbard, 2017).

Figure 4.2. Fragment optimization strategies. A schematic representation of the three main approaches in fragment to lead optimization. **(A)** In fragment linking, two hit fragments that bind to adjacent sites of the target are individually optimized to improve binding and are then linked together to generate a new set of fragments. **(B)** In fragment growing, the initial fragment is optimized by structure-based design to introduce the desired selectivity and drug-like properties. **(C)** Fragment merging involves combining information of multiple hits and structure-based design. Adapted from (Lamoree & Hubbard, 2017).

4.1.2. Study Approach

CHIKV full length nsP2 protein was chosen as the target for the screening of nsP2 protease inhibitors. A cell-free FRET-based protease assay, similar to that employed in sub-chapter 3.2.2, was used as a biochemical screening tool against CHIKV nsP2 protein. Notably, similar protease assays have been used for protease inhibitor validation of alphavirus proteases (Aggarwal et al., 2014, 2015; Singh et al., 2018). Hits were subsequently validated against ONNV and VEEV nsP2 proteases for cross inhibition and were further tested against multiple CHIKV isolates and other closely-related alphaviruses (ONNV and SFV), where their ability to inhibit viral replication were examined. Due to the absence of Gluc-tagged VEEV, Gluc-tagged SFV was included in the cell-based virus inhibition study instead. The use of ONNV and SFV would cross validate if hits can potentially perform as a pan-antiviral inhibitor. The biochemical assay together with the established cell-based virus inhibition assay were used to 1) assess the protease and viral replication inhibition potential of the hits, and 2) determine their mode of inhibition. In addition, co-crystallization strategy was employed in an attempt to determine the nsP2 protease-compound complexed structure and to understand their interaction.

4.2. Results

4.2.1. Expression and purification of CHIKV, ONNV and VEEV

nsP2FL proteins

Full length nsP2 of CHIKV, ONNV and VEEV were amplified from codon-optimized CHIKV-SGP11, ONNV-IMTSSA/5163 and VEEV genes using the primers listed in Appendix 2. The amplified CHIKV, ONNV and VEEV fragments were cloned into pNIC28-Bsa4, pET28b and peT16b protein expression vectors respectively (Appendix 1). They were transformed into Rosetta (DE3) cells and expressed in *E. coli*. For CHIKV nsP2FL alone, a three-step purification involving 1) Ni-NTA batch purification, 2) cation exchange chromatography and 3) size-exclusion chromatography was performed in order to obtain highly pure, homogenous and active nsP2 that were suitable for fragment screening. In the first batch-purification step, CHIKV nsP2 full length (nsP2FL) was eluted at 100 mM Imidazole with some protein contaminants observed (Figure 4.3A). The three eluted fractions were combined and loaded onto the cation exchange column. A single symmetrical peak appeared at ~280 mM NaCl concentration (Figure 4.3B). The eluted fractions were combined and loaded onto the size-exclusion column. A single symmetrical peak appeared at retention volume of 76 mL, which corresponded to ~90 kDa (Figure 4.3C). Fractions from the single peak were combined and concentrated to ~9 mg/mL. For ONNV and VEEV nsP2FL proteins, a two-step purification involving 1) Ni-NTA batch purification and 2) size-exclusion chromatography were performed. Both nsP2FL proteins were eluted at 100 mM Imidazole with some protein contaminants observed (Figures

4.4A and 4.5A). Single symmetrical peaks appeared at retention volumes 80 mL (~90 kDa) (Figure 4.4B) and 76 mL (~90 kDa) (Figure 4.5B) for the respective ONNV and VEEV nsP2FL proteins.

Figure 4.3. SDS PAGE, cation exchange and gel filtration profiles of CHIKV nsP2FL. (A) Visualization of Coomassie-stained SDS PAGE of eluted fractions from Ni-NTA beads. (B) The eluted fractions from Ni-NTA beads were diluted to reduce the NaCl concentration to ~50 μ M and then injected into SP HiTrap column for cation exchange. The protein eluted as a single peak at ~280 mM NaCl. (C) Eluted fractions from cation exchange was injected into a HiLoad® Superdex 16/60 200 column and eluted as a single symmetrical peak at 76 mL, which corresponded to ~90 kDa in size. The size-exclusion chromatography result suggested the monomeric nature of the protein and it purified to homogeneity. The protein was > 95% pure, as determined through visualization of Coomassie-stained SDS PAGE (right insert).

Figure 4.4. SDS PAGE and gel filtration profiles of ONNV nsP2FL. (A) Visualization of Coomassie-stained SDS PAGE of eluted fractions from Ni-NTA beads. **(B)** Eluted fractions from Nickel bead batch purification was injected into a HiLoad® Superdex 16/60 200 column and eluted as a single symmetrical peak at 80 mL, which corresponded to ~90 kDa in size. The size-exclusion chromatography result suggested the monomeric nature of the protein and it purified to homogeneity. The protein was > 95% pure, as determined through visualization of Coomassie-stained SDS PAGE (right insert).

Figure 4.5. SDS PAGE and gel filtration profiles of VEEV nsP2FL. (A) Visualization of Coomassie-stained SDS PAGE of eluted fractions from Ni-NTA beads. **(B)** Eluted fractions from Nickel bead batch purification was injected into a HiLoad® Superdex 16/60 200 column and eluted as a single symmetrical peak at 76 mL, which corresponded to ~90 kDa in size. The size-exclusion chromatography result suggested the monomeric nature of the protein and it purified to homogeneity. The protein was > 90% pure, as determined through visualization of Coomassie-stained SDS PAGE (right insert).

4.2.2. Determination of Fluorescence Extinction Coefficient of CHIKV and VEEV nsP1/2 peptide substrates

The concentration of the fluorophore cannot be calculated from the measured fluorescence emission since fluorescence intensities vary with fluorophore concentration. Thus, it is necessary to determine the fluorescence extinction coefficients (FECs) of CHIKV and VEEV nsP1/2 peptide substrates by constructing calibration curves for each of them. The fluorophore concentration in the protease reaction could then be determined by extrapolating the emission measurements to the calibration curves. The cell-free FRET-based protease assay setup was similar to the one described in Chapter 3 (Figure 3.2A), but with some modifications to the protein concentration, the peptide substrate concentrations and overall reaction time. Briefly, 10 μM of nsP2 was added to multiple peptide substrate concentrations (0.2 μM , 0.5 μM , 1 μM and 2 μM) in protease reaction buffer and the reactions were continuously monitored over 2 hrs, or until the RFU at each substrate concentration reached a plateau (Figures 4.6A and 4.6C). Then, the RFU values from the plateau region of the curve was plotted against substrate concentration to give a linear regression line, $Y = aX + c$. The slope of the line, a , was assigned as the FEC. It was determined that CHIKV and VEEV peptide substrates have FEC values 27541 RFU/ μM and 35544 RFU/ μM respectively (Figures 4.6B and 4.6D).

Figure 4.6. Dose-dependent curves and linear regression lines of CHIKV and VEEV peptide substrates. (A, C) Dose-dependent curves of CHIKV and VEEV nsP1/2 peptide substrates respectively. Briefly, 10 μ M of either CHIKV nsP2FL or VEEV nsP2FL was added to their respective peptide substrates at multiple concentrations (0.2 μ M, 0.5 μ M, 1 μ M and 2 μ M) in protease reaction buffer and the reactions were continuously monitored over 2 hrs, or until the RFU at each substrate concentration reached a plateau. The protease activity was measured relative to the signal at time 0s. (B, D) Linear regression to determine the fluorescence extinction coefficient (FEC) of CHIKV and VEEV nsP1/2 substrates. The RFU values from the plateau region of the curve was plotted against substrate concentration to give a linear regression line, $Y = aX + c$. The slope of the line, a , was assigned as FEC for the peptide substrate, with CHIKV and VEEV having 27541 RFU/ μ M and 35544 RFU/ μ M respectively.

4.2.3. Fragment-based screening against CHIKV nsP2FL

In this work, a cell-free FRET-based protease assay, similar to that used in Chapter 3, was employed to screen a fragment library against CHIKV nsP2FL protein for potential anti-protease inhibitors. Prior to fragment screening, the specificity and reproducibility of the cell-free protease system was tested with several reported nsP2 protease inhibitors and alphavirus replication inhibitors. Briefly, 0.5 μ M of nsP2 protein was pre-treated with 200 μ M of test compounds for 10 min at 22°C, before CHIKV peptide substrate was added and its fluorescence intensity monitored over time. Viral replication inhibitors such as Berberine, Obatoclax, Prostratin and Phorbol 12-myristate 13-acetate (PMA) were included as test compounds. When compared to DMSO control, they showed either no effect on protease activity or some levels of protease inhibition (Figure 4.7A). The results for Prostratin and PMA were expected as both were reported to inhibit CHIKV by specifically targeting the host factor Protein Kinase C (Bourjot et al., 2012; Nothias-Scaglia et al., 2015). It was interesting that Berberine and Obatoclax, which were postulated to inhibit MAPK signaling pathways and alphavirus membrane fusion respectively, could inhibit nsP2 protease activity to some extent (Figure 4.7A). This could however be an artifact caused by compound aggregation. nsP2 protease inhibitors such as Compound 1, Compound 1C, Compound 1D, Compound 9 and E64d (a generic cysteine protease inhibitor reported to inhibit VEEV nsP2 protease) that were reported to specifically target nsP2 protease were also included as test compounds. Overall, they demonstrated varying potency levels against the nsP2 protease (Figure

4.7B), which corroborated with the reports where these compounds were developed (Das et al., 2016; Hu et al., 2016). In addition, dose-dependent inhibition of E64d could also be clearly demonstrated (Figure 4.7C). Altogether, the results suggested that the assay is specific and robust.

Next, the fragment screening was performed against a library of 1000 structurally and chemically diverse fragments that were 'rule of three' compliant and generally used in the preliminary phase of drug discovery (Schöne et al., 2017; Screen et al., 2008; Silvestre et al., 2013). E64d, a generic cysteine protease inhibitor reported to inhibit VEEV nsP2 protease (Hu et al., 2016), and Pep-I, which was identified as a CHIKV substrate peptidomimetic inhibitor (Singh et al., 2018), served as positive reference compounds. The library was screened at a maximum concentration of 100 μ M in 1% DMSO against CHIKV nsP2FL protein. The measured RFU from each compound treatment was normalized against DMSO-treated and then reported as percent protease activity. The threshold for positive hits was determined as those less than 50% protease activity. Figure 4.8 showed an overview of the results of the fragment screening against CHIKV nsP2FL. A total of 29 compounds showed more than 50% inhibition in the single-concentration screening and were selected as primary hits, giving a hit rate of 2.9% (Figure 4.8). These were selected for further concentration-dependent analysis, where compounds were tested across 8 dilution points. Among the 29 tested, nine compounds displayed notable dose-dependent inhibition (compounds **1-9**; Figure 4.9). The nine hits were structurally diverse and were functionalized with pyridyl, hydroxyl, or amine groups. They were

classified into three categories, based on their core structures: a pyridine, a 5-membered heterocyclic and benzene derivatives (Table 4.1). In particular, four compounds (Compounds **1**, **2**, **4** and **8**) showed significant dose-dependent inhibition of nsP2 protease activity (Figure 4.9). Interestingly, Compound **2** performed better than Pep-1 and was comparable to E64d.

Figure 4.7. Reproducibility assessment of cell-free FRET based protease assay. Compounds known to either inhibit (A) alphavirus replication *in vitro* and *in vivo*, or (B) protease function were assessed. (C) Dose-dependent inhibition of E64d. 0.5 μ M of nsP2 protein was pre-treated with 200 μ M of test compounds for 10 min measured for 60 min at 5 min intervals. The protease activity was calculated using the formula: protease activity = $RFU_{\text{sample}} - RFU_{\text{DMSO}}$, where RFU_{sample} and RFU_{DMSO} were the measured RFU of the treated with E64d and 1% DMSO control respectively.

Figure 4.8. Primary screening of fragment library. A collection of 1000 compounds from Maybridge Ro3 fragment library, divided into 13 sets (represented on the X-axis), were screened at a 100 μ M concentration against recombinant CHIKV nsP2FL protein. nsP2FL protein was pre-treated with compounds at 22°C for 10 min, and then incubated at 30°C for 1 hr. The percentage of compound-treated nsP2FL protein was normalized with DMSO-treated to determine the percentage of protease activity. Compounds that reduce the nsP2 protease's activity to > 50% (i.e. values that fall below the horizontal dotted line, which represents > 50% residual protease activity) were selected for further dose-dependent analysis.

Table 4.1. Chemical structures of compounds 1-9

Compound	Structure
Pyridine derivatives	
1	
2	
3	
4	
5-membered heterocyclic derivatives	
5	
6	
7	
Benzene derivative	
8	
9	

Figure 4.9. Dose-dependent analysis of Compounds 1-9. CHIKV nsP2FL protein was treated with compounds that were diluted 2-folds (0-100 μM) in 100% DMSO and incubated at 30°C for 1 hr. Dotted lines indicate 50% protease activity and is a measurement for the test compounds' IC₅₀ values. Data is presented as mean ± SEM and representative of three independent experiments. Statistical significance was analysed by two-way ANOVA test and Dunnett's posttest (* P<0.05; ** P<0.01; *** P<0.001).

4.2.4. NTPase activity not affected by compound treatment

The CHIKV nsP2FL protein includes the N-terminal helicase domain, which possess an NTPase function. Hence, it would be important to establish if the hits selected from the fragment screening could inhibit the NTPase activity by preventing ATP hydrolysis. To demonstrate the effects of the nine hits on the NTPase function of CHIKV nsP2FL, NTPase assay was performed. 1 nM of CHIKV nsP2FL protein was pre-incubated with 1 μ M compound at 22°C for 10 min in NTPase reaction buffer, followed by the addition of 1 mM ATP substrate and further incubated at 37°C for 40 min. Interestingly, all nine hits, together with the reference compounds, did not appear to be able to decrease the amount of ATP hydrolysed since they were shown to be comparable to that of DMSO-treated (Figure 4.10). This result demonstrated that the hits did not disrupt the NTPase activity of the helicase domain and further suggested they are specific for the nsP2 protease.

Figure 4.10. nsP2 ATPase activity in the presence of 100 μ M compounds. CHIKV nsP2FL protein was incubated with compounds in a 100 μ L reaction at 37°C for 40 min. Released inorganic phosphate was quantitated as described in the experimental procedures. Dotted lines indicate 100% ATPase activity. Data is presented as mean $\pm$ SD and representative of two independent experiments.

4.2.5. Anti-protease activity of CHIKV primary hits are dose-dependent

Having shown that the nine best hits have protease inhibitory potential, they were re-validated using freshly prepared compounds and treated in a dose-dependent manner, to determine their IC_{50} values. All nine hits were tested at different compound concentrations (1 μ M – 100 μ M) against either 0.5 μ M CHIKV nsP2FL or 1 μ M CHIKV nsP2 protease (nsP2pro) proteins. Among these, compound **1** was shown to moderately inhibit CHIKV nsP2FL only (IC_{50} = 195.3 μ M), whereas compound **2** clearly exhibited dose-dependent inhibition against both CHIKV nsP2FL and CHIKV nsP2pro proteins tested (Figure 4.11A). Interestingly, compound **2** was twice as effective in inhibiting CHIKV nsP2pro (IC_{50} = 17.65 ± 1.12 μ M) as compared to CHIKV nsP2FL (IC_{50} = 42.05 ± 1.07 μ M) (Figure 4.11A, Table 4.2). Intriguingly, it displayed comparable inhibition to E64d and was 1.5 times more effective than Pep-1, despite its small size (Figure 4.11A, Table 4.2). The retesting of the remaining 7 compounds against CHIKV nsP2FL were found to be mostly ineffective (Figure 4.11B).

4.2.6. Similar inhibitory effect on nsP2 protease of ONNV and VEEV

Compounds **1** and **2** were then further assessed for their inhibitory activity against nsP2 protease of other alphaviruses, namely ONNV and VEEV, which have 89.5% and 40.8% sequence identity to CHIKV full length nsP2 respectively (Appendix 6). Purified nsP2FL proteins of ONNV and VEEV, and ONNV nsP2pro were treated and evaluated through the FRET-based

assay, where alphavirus-specific peptide substrates were used. None of the VEEV nsP2pro constructs were soluble, and thus excluded from the evaluation (Figure 3.2D). Compound **1** moderately inhibited both CHIKV and ONNV nsP2FL but was not effective against CHIKV nsP2pro, ONNV nsP2pro and VEEV nsP2FL. Interestingly, Compound **2** inhibited ONNV nsP2FL protease's activity 1.5 times more than that of CHIKV nsP2FL (Figure 4.11C, Table 4.2). Moreover, compound **2** was found to be the most effective in inhibiting VEEV nsP2FL ($IC_{50} = 7.948 \pm 1.12 \mu M$) and it was observed to be twice more potent than E64d ($IC_{50} = 17.29 \pm 1.33 \mu M$) (Figure 4.11C, Table 4.2). Despite its small size, Compound **2** consistently displayed comparable IC_{50} values to peptide-like compounds E64d and Pep-1, which are 3 to 4 times larger in size. Collectively, these data indicated that Compound **2** was the most potent out of the two compounds tested and it displayed significant inhibition against CHIKV, ONNV and VEEV nsP2 proteases. Thus, Compound **2** has the potential to be developed further into a more potent molecule that could be effective against multiple alphaviruses.

Figure 4.11. Dose-dependent nsP2 protease inhibition of CHIKV, ONNV and VEEV by hit compounds. (A) Comparison between CHIKV nsP2FL and nsP2pro protease activities in the presence of either hit compounds 1, 2, E64d or Pep-1. CHIKV nsP2FL was treated with compounds that were diluted 2-folds

(0-100 μ M) in 100% DMSO for 10 min at room temperature and then incubated with CHIKV fluorogenic substrate 30°C for 1 hr. Data is presented as mean $\pm$ SEM and representative of two independent experiments. **(B)** Dose-dependent inhibition of CHIKV nsP2FL by Compounds **3-9**. Data is presented as mean $\pm$ SD and representative of two independent experiments. **(C)** ONNV nsP2FL, ONNV nsP2pro and VEEV nsP2FL protease activities in the presence of either hit compounds **1**, **2** or E64d. Proteins were treated with compounds that were diluted 2-folds (0-100 μ M) in 100% DMSO for 10 min at room temperature and then incubated with their respective fluorogenic substrates at 30°C for 1 hr. FRET signal was measured at an excitation wavelength of 340 nm and emission wavelength of 490 nm, and the data was fitted to the four parameter equation. IC₅₀ curves were plotted using Graphpad Prism 7.0 software. Red arrow represents the cleavage site between the nsP1/2 sequences of the CHIKV, ONNV and VEEV substrates. Data is presented as mean $\pm$ SEM and representative of two independent experiments.

Table 4.2. Effect of hit compounds on nsP2 protease of CHIKV, ONNV and VEEV, where their IC₅₀ values were determined.

Compound	IC ₅₀ (μM) ^a				
	CHIKV nsP2FL	CHIKV nsP2pro	ONNV nsP2FL	ONNV nsP2pro	VEEV nsP2FL
1	> 100	n.i	> 100	n.i	n.i
2	42.05 ± 1.07	17.65 ± 1.12	12.89 ± 1.14	18.27 ± 1.18	7.948 ± 1.12
2-1	n.i	n.d	n.d	n.d	n.d
2-2	n.i	n.d	n.d	n.d	n.d
Pep-1	52.13 ± 1.12	n.d	n.d	n.d	n.d
E64d	22.89 ± 1.07	32.81 ± 1.13	8.847 ± 1.09	30.89 ± 1.11	17.29 ± 1.33

^aTo determine the IC₅₀ of an inhibitor, protease assay was performed using standard conditions as described in Materials and Methods. The IC₅₀ values were determined from the curves generated from the data that were fitted to the four parameter equation found in Graphpad Prism 7.0 software. Data were expressed as the mean of measurements ± SD of two independent experiments, with 2 replicates each (n=4). n.i: no inhibition, n.d: not determined.

4.2.7. Effect of combinatorial treatment on protease inhibition

There is also a possibility that Compounds **1** and **2** could elicit either an additive or synergistic inhibitory effect on one another. To test the effects of combinatorial treatment, CHIKV nsP2FL protein was treated with equimolar concentrations of compounds and the protease activity was measured. The result showed that combinatorial treatment of Compounds **1** and **2** did not enhance protease inhibition, indicating that linking the two fragments together in the future would likely not improve its overall potency (Figure 4.12). Additionally, the combinatorial effect of Compound **2** and E64d was evaluated, since both demonstrated significant inhibition individually. However, there was no improvement in protease inhibition either. This suggested that they could not elicit additive or synergistic effects on one another, but most importantly they did not antagonistically alter each other's inhibitory mechanism.

Figure 4.12. Effect of combinatorial treatment on protease activity. CHIKV nsP2FL protein was either single- or co-treated with Compounds 1, 2 and E64d. Prior to compound treatment, all compounds were diluted first. For a single treatment set up, the compounds were diluted 2-folds (0-100 µM) in 100% DMSO. For a combinatorial treatment set up, compounds were mixed in a 1:1 ratio before they were diluted 2-folds (0-100 µM) in 100% DMSO. 1 µM of CHIKV nsP2FL was treated with the diluted compounds for 10 min at 22°C and then incubated with 1 µM fluorogenic substrate at 30°C for 1 hr. Data are expressed as the mean of measurements ± SD of two independent experiments.

4.2.8. Amine at meta position of Compound 2 is crucial to its bioactivity

Within the fragment library, there are also two other compounds that share similar scaffold structure to compound **2**, namely Compounds **2-1** and **2-2**. These two compounds possess an amine group at para and methylamino at meta positions of the aromatic ring respectively (Figure. 4.13A). To evaluate their inhibitory effect on protease activity, CHIKV nsP2FL protein was treated with either Compound **2-1** or Compound **2-2** in a dose-dependent manner. Interestingly, both compounds were evaluated to be inefficient in inhibiting nsP2 protease activity as compared to Compound **2** (Figure 4.13B), suggesting that the amine group at the meta position of the aromatic ring of Compound **2** was crucial for its bioactivity.

Figure 4.13. Compounds 2-1 and 2-2 do not inhibit CHIKV nsP2 protease activity. (A) Compounds 2-1 and 2-2 share similar scaffold structures to Compound 2. (B) Comparison of protease inhibition by Compounds 2, 2-1 and 2-2. 1 μM CHIKV nsP2FL protein was treated with either Compounds 2, 2-1 or 2-2, that has been diluted 2-folds (0-100 μM) in 100% DMSO, for 10 min at 22°C and then incubated with 1 μM of fluorogenic substrate at 30°C for 1 hr. Data is presented as mean ± SD and representative of two independent experiments.

4.2.9. Kinetic analysis of inhibition by Compound 2

Multiple biochemical experiments were performed to characterize the binding affinity of the most potent compound **2** for CHIKV protease. Firstly, isothermal titration calorimetry was performed using 50 μM of either CHIKV nsP2FL or CHIKV nsP2pro, with agitation at 100 rpm. However, the protein precipitated prior to the addition of the test compound. Decreasing the protein concentration to 20 μM , agitation speed to 50 rpm and temperature to 20°C did not lessen the precipitation too. Thus, this system proved to be incompatible. Secondly, microscale thermophoresis was performed using 0.5 μM of CHIKV nsP2FL and varying Compound **2** concentrations (0–8 mM). However, decreasing fluorescence was observed with increasing ligand concentration (Figure 4.14). This is typical of either quenching of the fluorescence intensity of the protein upon ligand binding or nonspecific adsorption. Although 0.05% Tween-20 was added to the buffer in an attempt to reduce nonspecific absorption, the differences in fluorescence intensity persisted. Thus the data could not be used to extract affinities.

Since the above-mentioned systems were deemed unsuitable for studying protein-ligand kinetics, the FRET-based protease assay was used instead. Although this assay did not allow the study of physical interaction of the compound with the protein, it could determine the mechanism of inhibition. In this assay, six substrate concentrations ranging from 0.3125 μM to 10 μM were used to treat either nsP2FL or nsP2pro. In the absence of Compound **2** treatment, CHIKV nsP2FL showed a two fold-lower $K_{\text{cat}}/K_{\text{m}}$ than CHIKV nsP2pro while VEEV

nsP2FL was two and three times more active than ONNV nsP2FL and CHIKV nsP2FL respectively (Table 4.3). When the enzymes were treated with increasing concentrations of Compound **2** (25-100 μ M for CHIKV and ONNV; 5-20 μ M for VEEV), there was a concomitant decrease in the apparent $V_{\max}$ values, while the values of K_m remained constant (Figure 4.15A-C, Table 4.3). At the highest compound concentration tested against CHIKV nsP2FL, CHIKV nsP2pro and ONNV nsP2FL, their catalytic efficiencies (K_{cat}/K_m values) reduced five-fold (Table 4.3). Although the highest compound concentration tested against VEEV nsP2FL was five times lower than those tested against CHIKV and ONNV, its catalytic efficiency reduced three-fold (Table 4.3). Overall, Lineweaver-Burk plots revealed that Compound **2** demonstrated a non-competitive mode of inhibition against all viral proteases tested (Figure 4.15D-F). This suggested that the compound did not interfere with the binding of peptide substrate to the catalytic site of the enzyme (since K_m was unaffected) but influenced the progression to reaction products (since $V_{\max}$ was decreased with increasing concentration of compound).

8 mM
Compound 2

0.04 mM
Compound 2

Figure 4.14. MST capillary scan during binding affinity test. 0.5 μ M of CHIKV nsP2FL was treated with Compound 2 that was diluted 2-folds (0.04 – 8 mM) in protease reaction buffer (20 mM HEPES (pH 7.2), 2 mM DTT). The samples were measured using the Monolith.NT115, at 10% LED-Power.

Table 4.3. Kinetic parameters of nsP2FL and nsP2pro in the presence of Compound 2.

Condition	K_{cat} (10^{-4} s^{-1})	K_m (μM)	$\frac{K_{cat}}{K_m}$ ($10^{-4} \text{ s}^{-1} \cdot \mu\text{M}^{-1}$)
No treatment			
CHIKV nsP2FL	9.54 ± 0.91	16.3 ± 2.25	0.58
CHIKV nsP2pro	38.79 ± 14.56	41.74 ± 17.79	0.93
ONNV nsP2FL	24.7 ± 1.62	34.1 ± 3.23	0.724
VEEV nsP2FL	22.26 ± 5.03	16.16 ± 4.91	1.38
Compound 2 treatment^a			
CHIKV nsP2FL + 100 μM	1.86 ± 0.57	12.8 ± 6.11	0.145
CHIKV nsP2pro + 100 μM	7.89 ± 3.89	35.26 ± 20.24	0.22
ONNV nsP2FL + 100 μM	7.30 ± 1.45	36.0 ± 10.2	0.203
VEEV nsP2FL + 20 μM	10.24 ± 3.07	14.06 ± 5.87	0.557

The kinetic parameters were determined using progress curve analysis. The data were calculated by Graphpad Prism 7.0. The kinetic results are represented as mean $\pm$ S.D. for at least two independent experiments.

^aThe kinetic results of the highest Compound 2 concentrations tested are represented here.

Figure 4.15. Enzyme inhibition kinetics of Compound 2 against alphavirus nsP2 protease. Cell-free FRET-based protease assay was used to understand the mechanism of inhibition against CHIKV, ONNV and VEEV nsP2FL proteins by increasing the concentrations of peptide

substrates respectively. **(A-C)** The maximal velocity ($V_{\max}$) achieved by enzymes in the presence of increasing concentrations of Compound **2**, from 0.3125 μM to 10 μM . At every compound concentration tested, the initial velocity (V_i) was plotted against the substrate concentration ($[S]$). Data is expressed as the mean of measurements $\pm$ SD of two independent experiments, with 2 replicates each ($n=4$). The data was fitted to the Michaelis–Menten equation: $[V_0 = (V_{\max} \times [S]) / (K_m + [S])]$ using Graphpad prism 7 software. **(D-F)** Lineweaver-Burk plots for noncompetitive inhibition of Compound **2**. The reciprocal values of V_i was plotted against the reciprocal values of the concentration of substrates. The units for V_i are $\mu\text{M/s}$; the units for K_m and substrate concentrations are micromolar concentrations. The experiments were repeated twice with similar results. Results of a representative experiment are shown.

4.2.10. Compound 2 destabilizes CHIKV nsP2 protease

Next, a thermal shift experiment was used as a biophysical technique to evaluate the binding mechanism of compound **2** on CHIKV nsP2pro. This fluorescence-based assay could monitor fragment-induced changes in the protein T_m (Lo et al., 2004), where it is often used as a first screen in a screening cascade for fragments (Wan et al., 2009). Sypro Orange, a dye sensitive to its chemical environment, is added to a fragment-protein solution, where it will preferentially bind to partially denatured “molten globule” state of the protein and its fluorescence increases. By measuring the fluorescence emission over a range of temperatures, it is possible to determine the protein T_m (Figure 4.16). When a compound binds to the nsP2pro, the enzyme may undergo a conformational change that either stabilizes or destabilizes its tertiary structure. Such stabilization could be measured by monitoring changes in nsP2pro T_m in the presence of compounds. Herein, an attempt was made to follow compound-induced changes in T_m by adding Sypro Orange to nsP2pro-compound solutions.

Before initiating the compound screen, the T_m of CHIKV nsP2pro and the optimal protein-to-dye ratio was determined, by titrating various concentrations of protein against multiple Sypro Orange concentrations. The melting profile of CHIKV nsP2pro was monophasic and its T_m was observed to be 45°C (Figure 4.17A). In addition, the optimal protein-to-dye ratio was determined at 0.8 µg protein: 8X Sypro Orange. To assess the binding of E64d, Compound **1** and **2** to CHIKV nsP2pro, a wide range of compound concentrations (1.9 µM to 1.25 mM) were added to the protein. E64d and Compound **1** had minimal effect on CHIKV nsP2pro

(Figures 4.17B-C). The result for E64d was surprising as the relatively potent peptidomimetic inhibitor was shown in crystal structure to bind to the protease's catalytic core (Hu et al., 2016). This was contrary to expectations as potent compounds typically induce a significant positive change in T_m (Wan et al., 2009). It may be possible that peptide binding could instead induce minimal remodeling of the protease hydrophobic groove as compared to chemical inhibitors, thereby causing minimal unfolding transition. Intriguingly, Compound **2** shifted the melting profile to the left in a concentration-dependent manner and gave the highest negative change in T_m value (-6.8°C) at the highest concentration tested (Figure 4.17D). At inhibitory concentration ($\text{IC}_{50} = 17.65 \mu\text{M}$), CHIKV nsP2pro T_m was extrapolated to be $\sim 43^{\circ}\text{C}$ (Figure 4.17D). The 2°C decrease in T_m suggested that Compound **2** binding destabilized nsP2 protease and disrupted the hydrophobic interactions between the substrate and the catalytic pocket, thereby preventing the formation of stable nsP2pro-substrate complex. It is plausible that the compound interacts with distinct sites other than the substrate binding cleft where E64d was shown to bind to and may result in the disordered conformation.

Figure 4.16. Thermal shift assay schematic. An example of a typical melt curve obtained from a TSA experiment. This curve is characteristic of classic "two-state" protein unfolding, where the sample population transitions from folded to denatured without detectable partially-folded intermediates. Adapted from (Bruce et al., 2019).

Figure 4.17. The effect of compounds on melting temperature (T_m) of CHIKV nsP2pro measured by thermal shift assay. (A) Titration of CHIKV nsP2pro and Sypro Orange® dye. Unfolding transition of CHIKV nsP2pro in the presence of (B) E64d, (C) Compound 1 and (D) Compound 2. Solutions containing 0.8 µg of CHIKV nsP2pro and 0.8X Sypro Orange dye in 50 mM Hepes pH 7.5, 300 mM NaCl, 1 mM DTT were incubated with the respective compounds at varying concentrations (from 1.9 µM to 1.25 mM) at 25°C for 2 min, then scanned between 25°C and 85°C. Thermal denaturation (1°C/min) was performed in a CFX quantitative real-time PCR instrument (Biorad). Integrated intensity from reactions was fitted to Boltzmann equation and are represented as solid lines.

4.2.11. *In vitro* antiviral efficacy against CHIKV isolates

The antiviral activities of the nine primary hits (Table 4.1) selected from the CHIKV nsP2 protease screening were then evaluated by measuring their inhibitory property and toxicity at varying concentrations (0.01 μM – 1000 μM) against four *Gaussia* luciferase (Gluc)-tagged CHIKV isolates (i.e., SGP11, CNR20235, LR2006-OPY1 and IND91) in human embryonic kidney 293T (HEK293T) cells. These CHIKV isolates belonging to ECSA lineage were included in this study because virulence levels between isolates were known to vary (Chen et al., 2013), despite having identical protease sequences (Appendix 7). The EC_{50} and CC_{50} values were determined for the nine primary hits and three reference compounds, namely E64d, Pep-1 and Ribavirin. Eight of the hits showed varying inhibition of CHIKV-SGP11 isolate in a dose-dependent manner (Figures 4.18A-B, Table 4.4). Although Compound **9** displayed the lowest effective dose (EC_{50} = 20.4 μM) out of the nine hits tested (Figure 4.18B), it was demonstrated to be the most toxic to cells (CC_{50} = 57 μM) suggesting that the observed antiviral effect could be due to toxicity (Table 4.4). Compound **1** had the best effective dose (EC_{50} = 69.08 μM), followed by Compound **5** (EC_{50} = 167.3 μM) and they both appeared to have no cytotoxicity at the tested concentrations (CC_{50} > 1000 μM) (Figure 4.18A, Table 4.4). Interestingly, Compound **1** was as potent as Pep-1 and E64d (EC_{50} = 78.54 μM and 76.64 μM respectively) (Table 4.4). Interestingly, Compound **2** was only moderately inhibitory (EC_{50} = 251.7 μM), which was four times less efficient than Compound **1**. The remaining five compounds were demonstrated to be either less inhibitory or not effective

at all (Figure 4.18B). Furthermore, all nine hits were less potent against CNR20235, LR2006-OPY1, and IND91 CHIKV isolates, with EC_{50} values averaging two times higher as compared to that against CHIKV-SGP11 (Table 4.4). The compound sensitivity differences observed between the CHIKV isolates are likely attributed to variances in their virulence levels.

In order to examine the potential inhibitory effects of these compounds on a more relevant cell line, the screens of Compounds **1**, **2** and reference compound Ribavirin against CHIKV-SGP11 were performed in BJ cells, a human skin fibroblast cell line. The effective doses of Compounds **1** and **2** were approximately three- and one-fold less effective respectively as compared to when tested in HEK293T cells (Figure 4.19A, Table 4.5). Despite showing very efficient viral inhibition in HEK293T cells ($EC_{50} = 1.13 \mu\text{M}$), Ribavirin was also not effective in inhibiting viral replication in BJ cells ($EC_{50} = 334.7 \mu\text{M}$). This was not surprising since Franco et al. demonstrated that the effectiveness of antivirals varied between the type of host cell lines used (Franco et al., 2018). Nonetheless, Compounds **1** and **2** were non-toxic to the cells even at very high doses (1000 μM) as compared to Ribavirin (Figure 4.19B, Table 4.5).

4.2.12. Antiviral efficacy on other pathogenic alphaviruses

The hit compounds were further screened against IMTSSA/5163-Gluc (ONNV) and SFV-Gluc (SFV) in order to determine if they possess inhibitory activity against other pathogenic alphaviruses. SFV was included instead of VEEV at this juncture due to the unavailability of Gluc-

tagged VEEV. The potency of the hit compounds was evaluated by dose-response analysis in HEK293T cells and compared to Ribavirin's antiviral activity. These compounds showed varying degrees of antiviral activity across the two alphaviruses tested and were collectively less effective. In ONNV infections, Compounds **1**, **2** and **5** were moderately potent with EC₅₀ values of 96.83 µM, 209.5 µM and 129.4 µM respectively (Figure 4.18A, Table 4.4), which were similar to that of CHIKV-SGP11-infection in HEK293T cells. This observation was expected as ONNV is the most closely related to CHIKV. However, it is noteworthy that the compounds were overall 100-fold less potent than Ribavirin. All three compounds were also 2 to 4 folds less effective against SFV than against CHIKV-SGP11 (Figure 4.18A, Table 4.4).

Figure 4.18. Antiviral activities of compounds against several alphaviruses in HEK293T cells. (A) Efficacy of Compounds 1, 2 and 5 against four CHIKV isolates, ONNV and SFV. (B) Efficacy of Compounds 3, 4, 6, 7, 8, and 9, E64d, Pep-1 and Ribavirin against CHIKV-SGP11, ONNV and SFV. HEK293T cells were infected with Gluc-tagged viruses at a MOI of 0.1 and simultaneously treated with different concentrations of the compounds. Viral replication was determined at 16 hr post-infection by measuring Gluc luminescence. Experiments were carried out in triplicates and repeated twice for confirmation. Data presented as mean of measurement $\pm$ SEM (n=9).

Table 4.4. Inhibitory profile of compounds against multiple luciferase-tagged CHIKV isolates, ONNV and SFV in HEK293T cells.

<u>Compound</u>	<u>EC₅₀ (mM)^a</u>						
	<u>CHIKV Isolates</u>				<u>Other alphaviruses</u>		<u>CC₅₀ (mM)^b</u>
	<u>SGP11</u>	<u>CNR20235</u>	<u>LR2006-OPY1</u>	<u>IND91</u>	<u>ONNV</u>	<u>SFV</u>	
Ribavirin	1.13	ND ^c	ND ^c	ND ^c	1.39	8.62	> 1000
Pep-1	78.54	ND ^c	ND ^c	ND ^c	10.16	193.8	> 1000
E64d	76.64	ND ^c	ND ^c	ND ^c	50.23	193.2	> 1000
1	69.08	248.1	187	213.3	96.83	234.7	> 1000
2	251.7	> 500	434.3	484.1	209.5	328.9	> 1000
3	312.9	> 500	> 500	> 500	258.4	286.5	> 1000
4	> 500	> 1000	> 1000	> 1000	> 500	> 1000	> 1000
5	167.3	162.4	203.3	162.3	129.4	144.3	> 1000
6	178	196.1	255.3	194.2	177.2	205.7	> 1000
7	184	1872.5	460	325.5	233.4	482.2	> 1000
8	251.4	> 1000	> 1000	> 1000	> 500	> 1000	> 1000
9	20.4	25.63	35.86	33.22	16.71	21.2	57

^aEC₅₀ values of hit compounds were determined from dose response curve analyses, where data are presented as mean ± SEM of three independent experiments.

^bCC₅₀ values of hit compounds, a measurement for their cellular cytotoxicity, were determined from dose response curve analyses, where data is presented as mean ± SEM of three independent experiments.

^cND: Not determined

Table 4.5. Inhibition of CHIKV-SGP011-Gluc infection by different compounds in HEK293T and human skin fibroblast BJ cells.

<u>Compound</u>	<u>HEK293T cells</u>			<u>BJ cells</u>		
	<u>EC₅₀</u> <u>(mM)^a</u>	<u>CC₅₀</u> <u>(mM)^b</u>	<u>SI^c</u>	<u>EC₅₀</u> <u>(mM)^a</u>	<u>CC₅₀</u> <u>(mM)^b</u>	<u>SI^c</u>
Ribavirin	1.13	> 1000	> 500	334.7	> 1000	2.99
E64d	76.64	> 1000	13	226.3	> 1000	4.42
1	69.08	> 1000	14.5	196.6	> 1000	5.09
2	251.7	> 1000	4.38	288.5	> 1000	3.47

^aEC₅₀ values of hit compounds were determined from dose response curve analyses, where data are presented as mean ± SEM of three independent experiments.

^bCC₅₀ values of the compounds were determined from dose response curve analyses, where data is presented as mean ± SEM of three independent experiments.

^cSI is determined by the formula: CC₅₀/EC₅₀.

^dND: Not determined

Figure 4.19. Inhibition of CHIKV infection by Compounds 1, 2 and Ribavirin in human skin fibroblast BJ cells. (A) BJ cells were infected with Gluc-tagged CHIKV-SGP11 virus at MOI 1 and subsequently treated with different concentrations of the compounds for 16 hr. Supernatant was collected posttreatment and analysed for dose-dependent inhibition of CHIKV replication. **(B)** BJ cells were similarly treated with compounds for 16 hr, lysed and cell viability was measured to determine cytotoxicity of compounds. The experiments were carried out in triplicates and repeated twice for confirmation. Data is expressed as the mean of measurements $\pm$ SEM (n=9).

4.2.13. Efficacy of compounds *in vivo*

The availability of CHIKV mouse models that could recapitulate viremia and joint pathology allows for the assessment of rheumatic manifestations similar to that seen in infected humans (Gardner et al., 2010; Her et al., 2015; Lum et al., 2013; Morrison et al., 2011; Teo et al., 2013). The mouse model involved the inoculation of CHIKV at the ventral site of the footpad of wild-type (WT) C57BL/6, 3-weeks old animals, where they displayed a measurable, but self-limiting peri-metatarsal joint footpad swelling and viremic phase (Teo et al., 2012). This model has been proven to be suitable for proof-of-efficacy testings of potential CHIKV drugs. Fingolimod, an agonist of sphingosine 1-phosphate receptor, was recently shown to abrogate CHIKV-induced joint pathology (Teo et al., 2017). Furthermore, a combination therapy involving CTLA4-Ig and a neutralizing anti-CHIKV monoclonal antibody was revealed to be effective in ameliorating acute CHIKV arthritis in mice (Miner et al., 2017).

To assess whether the compounds were able to limit the symptoms of CHIKV-induced disease, the established model was adopted. In this *in vivo* experiment, Ribavirin, a synthetic guanosine analog with a broad-spectrum of antiviral activity, was included as a positive control since studies have demonstrated that Ribavirin therapy in CHIKV patients led to faster resolution of joint and soft tissue manifestations (Ravichandran & Manian, 2008), and could considerably reduce viral load *in vivo* (Rothan et al., 2016, 2015). Briefly, four-week old WT C57BL/6 male mice were inoculated at the joint footpad with CHIKV-SGP11 and treated from 1 to 6 days post infection (dpi) by administering intraperitoneally with 10 mg/kg

of Compound **1**, Compound **2**, Ribavirin, or vehicle control. The treatment window from 1 to 6 dpi was selected because this represents the time point between the first and second peak of clinically apparent footpad swelling in the CHIKV mouse model and it mimics a typical antiviral prescription. Given that the pharmacokinetics of all compounds tested were not known, the dose was chosen based on a study that used Ribavirin at a concentration which could protect mice against lethal influenza A infection (Rowe et al., 2010). Animals were followed over the course of 14 days of infection, whereby disease score (as a measure of footpad swelling) and viremia were monitored daily.

Intriguingly, Ribavirin treatment did not control joint inflammation in mice. Instead, footpad swelling was aggravated at 2 dpi and the inflammation was prolonged from 7 to 10 dpi (Figure 4.20A). Although Ribavirin treatment could not protect mice against CHIKV infection, a slight reduction in viral load was observed from 2 dpi, suggesting that Ribavirin could reduce virus replication (Figure 4.20B). Despite showing a slightly reduced viral load, Ribavirin was not effective as a therapeutic at the dose given since it could not alleviate footpad swelling. Treatment with either Compound **1** or Compound **2** did not show any significant differences in footpad swelling when compared with DMSO-treated group. However, like Ribavirin, both compounds could slightly reduce viral load. Despite exhibiting either moderate viral replication inhibition and good protease inhibition *in vitro*, Compounds **1** and **2** were demonstrated to lack therapeutics at the dose and treatment regime tested, as they could not protect mice against CHIKV infection.

Figure 4.20. Screen of candidate compounds for treatment of acute CHIKV arthritis in mice. Time-course dependent infection kinetics were performed *in vivo* with WT C57BL/6, four-weeks old male mice. Mice (n = 8 animals per group) were infected via footpad inoculation with 10⁶ PFU of CHIKV-SGP11. These mice were treated on 1 to 6 dpi by administering intraperitoneally with 5 mg/kg of Compound 1, Compound 2, Ribavirin, or vehicle control. Daily assessment of **(A)** disease score and **(B)** viremia were performed. Data was representative of two independent experiments and are presented as mean $\pm$ SD. Statistical analyses were performed by Mann-Whitney U two-tailed test (**P < 0.01, ***P < 0.001). Treatment periods are indicated by shaded region in the graphs.

4.2.14. Crystallization attempts of CHIKV protease-Compound 2 complex

Although the specificity, potency and mechanism of action of Compound 2 were determined, it is imperative that the binding of the compound is visualized in 3D structure. The structural information could reveal the unique interactions within the protein-compound complex and can be used to facilitate the development of the compound into more potent lead candidates. In this study, I attempted to co-crystallize CHIKV-vc045 with Compound 2. Prior to crystallization screening, CHIKV-vc045 and Compound 2 were incubated at a 1:1.3 or 1:2 molar ratio at 22°C for 10 min. The complexed solutions were centrifuged to remove debris and was subsequently screened at a 1:1, 1:2 and 2:1 ratio using commercial crystallization kits through sitting drop vapor diffusion at 20°C. In parallel, apo-protein was screened too. In the 1:2 complexed solution, heavy precipitates were observed in more than 50% of screening conditions. For the 1:1.3 complexed protein, microcrystals appeared at Day 11 in JCSG+ Condition 12 crystallization buffer containing 0.04 M potassium phosphate monobasic, 16% (w/v) PEG 8000 and 20% (v/v) glycerol (Figure 4.21A) but heavy protein precipitation was observed for the apo-protein in the same crystallization buffer condition. The microcrystals of the complexed protein were reproduced through hanging drop vapor diffusion and was further optimized through multiple rounds of microseeding. Briefly, the microcrystals were diluted 20-folds for up to six times and was screened at a 1:1 and 1:2 ratio in the crystallization buffer. Bigger crystal clusters were observed at 7 days and they were sent for

data collection (Figure 4.21B). However, they did not diffract. Alternatively, attempts were made to soak apo ONNV nsP2 crystals (from Sub-chapter 3.2.4) with Compound **2**. Crystal form 3 (Figure 3.11C) was soaked with either 1 mM, 2 mM or 10 mM of Compound **2** at 20°C for 1 hr, 2 hrs and overnight. Overnight soaking for all conditions and soaking in 10 mM compound at 1 hr and 2 hrs caused the crystals to dissolve into the crystallization solution. Crystals soaked in either 1 mM or 2 mM compound at shorter incubation times did not display any signs of cracking or dissolution and were sent for data collection. Crystals soaked in 1 mM compound diffracted at 3 Å and datasets were collected and analysed. However, no compound was found in the structure. Failure to obtain a compound-bound structure could be attributed to the inaccessibility of the binding pocket in the protease crystal. Although crystals soaked in 2 mM compound diffracted at 3.5 Å, the dataset was not collected as it would be impossible to view the small compound at such low resolution.

Figure 4.21. Crystals obtained from screening of CHIKVpro-Compound 2 complex. (A) Microcrystals at Day 11 of crystal screening. Complexed protein solution was screened at 1:1, 1:2 and 2:1 ratio in Condition D12 of JCSG+ crystallization kit (1.1 M sodium malonate pH 7.0, 0.1 M HEPES pH 7.0 and 0.5% (v/v) Jeffamine ED-2001 pH 7.0) through the sitting drop vapor diffusion method and was incubated at 20°C. (B) Crystal clusters at Day 7 of microseeding. Microseeds from the initial screening drop (Figure 4.20A) was diluted 20-folds for up to six times in JCSG+ Condition D12 crystallization buffer. The crystals were grown using the hanging drop vapor diffusion method at 20°C.

4.3. Discussion

The debilitating disease caused by CHIKV together with the absence of FDA-approved drugs to treat arthritogenic CHIKV infections highlight an imperative need to search for effective therapeutics. Infected patients are often prescribed with generic nonsteroidal anti-inflammatory drugs to control inflammation (Weaver & Lecuit, 2015). However, these drugs do not specifically target the primary pathogenic pathways in virus-induced inflammation, thus limiting their efficacy in disease management.

With the recent availability of the crystal structure of alphavirus nsP2 protease, target-based modeling is frequently used and has resulted in the design and development of many small molecules that were predicted to interact with the active site of nsP2 protease (Singh et al., 2012, 2018; Das et al., 2016; Hu et al., 2016). In recent years, FBLD has also facilitated the discovery of several potent inhibitors. Although initial hits from fragment screening are usually weak binders, they can be optimized to become high affinity compounds. Inhibitors such as IMP-1088 (a rhinovirus capsid assembly blocker) (Bell et al., 2018) and AT7519 (an inhibitor of cyclin dependent kinases 1 and 2) (Murray & Rees, 2009), were successfully developed through fragment linking of weak-binding fragments.

In this study, I described the use of a fragment-based approach to identify novel ligands that have the ability to inhibit the enzymatic activity of CHIKV nsP2 protease. For the primary screen, a fragment-based library, similar to that used in the discovery of novel bidentate West Nile virus NS2B-NS3 protease inhibitors (Schöne et al., 2017), was chosen

and fragments were selected based on their ability to inhibit CHIKV nsP2 protease. Out of the nine hits that emerged from the two-phase screening process, Compound **2** consistently showed the best inhibitory activity and specificity for CHIKV nsP2 protease (Figure 4.11A). Interestingly, fragments with similar scaffold structure to Compound **2** (Compounds **2-1** and **2-2**) were found to be ineffective, suggesting that the amine group at the meta position of the aromatic ring of Compound **2** is crucial for its bioactivity (Figure 4.13). Moreover, Compound **2** was demonstrated to be effective in inhibiting the nsP2 protease activities of ONNV and VEEV (Figure 4.11B). Its IC_{50} value was comparable to that of E64d, and surprisingly more potent than the peptidomimetic inhibitor, Pep-1 (Table 4.2). Taken together, it was evident that Compound **2** possessed pan-alphavirus anti-nsP2 protease activity. Despite its submicromolar potency, Compound **2** had the potential to be further developed into a more potent protease inhibitor. Interestingly, Compound **1** could moderately inhibit nsP2FL, but not nsP2pro. This suggested that the compound did not bind to protease's catalytic site but targeted either a part of the protease or the flexible linker of the nsP2FL, which altered the spatial arrangement of the domains and consequently had a slight inhibitory effect on the protease's function.

Next, the substrate cleavage characterization of nsP2FL from CHIKV, ONNV and VEEV showed that the enzymes have large differences in their binding affinity and catalytic efficiency for their respective nsP1/2 peptide substrates (Table 4.3). The catalytic efficiency of CHIKV nsP2pro for nsP1/2 substrate was similar to previous reports

(Saisawang et al., 2015a; Saisawang et al., 2015b), However, contrary result was observed for CHIKV nsP2FL as it was found to be less efficient than CHIKV nsP2pro (Saisawang et al., 2015b). Interestingly, VEEV nsP2FL was most efficient in cleaving its nsP1/2 substrate, as illustrated by the high K_{cat}/K_m value (Table 4.3). Nonetheless, these variances could contribute to the IC_{50} differences observed upon Compound **2** treatment. In fact, VEEV protease was evaluated to be the most sensitive to the inhibition by Compound **2**, since it required 1.6 to 5 times lower compound concentration to achieve similar fold-reduction to that of CHIKV and ONNV nsP2 proteases (Table 4.3). Regardless of their differences, Compound **2** demonstrated non-competitive inhibition against all enzymes tested (Figures 4.14D-F, Table 4.3). This suggested that the compound did not interfere with the binding of the protease to the respective peptide substrates (constant K_m) but influenced the progression of the reaction products. In addition, the negative change in T_m of CHIKV nsP2 protease upon Compound **2** binding, suggesting it destabilized the protein (Figure 4.17D). Compounds of such nature were usually dismissed as potential inhibitors because they were assumed to be non-specific. However, studies have shown that these compounds could selectively inhibit protein function (Askin et al., 2018; Zhao et al., 2016), which could be a possibility for Compound **2** too.

Despite its potency and specificity to nsP2 protease, Compound **2** was not as effective in inhibiting CHIKV infection *in vitro*. Several other compounds analyzed in this study (most notably compounds **1** and **5**) inhibited CHIKV and ONNV replications with much lower EC_{50} values than

could be deduced from their ability to inhibit the protease activity of nsP2. It should be noted that these compounds displayed varying viral inhibition efficacies against the different CHIKV isolates (CNR20235, LR2006-OPY1, and IND91), where differences in virulence levels among the CHIKV isolates could have contributed to the differences in compound sensitivity observed (Chen et al., 2013). The differences in binding mechanism may also justify why Compound **1** was more efficient than Compound **2** in inhibiting virus replication, despite the former showing little protease inhibition. Compound **1** was speculated to bind to a region of nsP2FL that could alter the spatial arrangement of the helicase and protease domains. Although the alteration minimally impacted the protease's function, which was substantiated by the little nsP2FL inhibition, it was plausible that the resulting structure could prevent optimal assembly of the nsPs. This would have detrimental effect on virus replication since the formation of active replicase would be disrupted. Conversely, the non-competitive manner with which Compound **2** inhibited the protease would suggest that the compound elicited its inhibitory mechanism by binding to a region other than the protein's substrate recognition site. Since pure nsP2 protein was used in the cell-free protease assay, the region was likely exposed and eventually targeted. However, it could be buried within the replicase by the other nsPs during viral replication. This would prevent the compound from accessing the area and as a result cause little viral inhibition. It is plausible that the hydrophobic nature of Compound **2** could have prevented it from traversing the cellular membrane to reach the viral replication site.

Nonetheless, the possibility of off-target effects due to their small size should never be excluded.

Despite the positive outcome from the *in vitro* studies, Compounds **1** and **2** did not perform as well *in vivo*. Treatment with either compounds could not alleviate footpad swelling, although it could slightly reduce viral load. This lack of therapeutic efficacy may be attributed to the dosage regimen, which includes the route of drug administration, the amount to be administered each time, the dosage intervals and the duration of the treatment. Additionally, mice age could play a role since it has been shown to have an effect on viremia levels and virus persistence (Uhrlaub et al., 2016). Since these factors have not been assessed individually, the chosen experimental design may not be ideal for evaluating the compounds' therapeutic potential. Further experimental validation and more optimization is required before concluding that compounds **1** and **2** are ineffective.

The disconnect between protease inhibition and virus infection *in vitro* and *in vivo* exposed a major limitation of FBLD. Since the screening was performed against pure proteins, there was a high tendency to identify hits that interact with regions that were naturally unexposed *in vivo*. They may not work in desired cellular environments since they may either 1) form multiple interactions with either cellular or viral components, or 2) have difficulty accessing the binding site, which could be hidden by interacting proteins. Another drawback of fragment-based screening is that hits are weakly binding, which will make progressing them to potency difficult.

Although the outcome of the *in vivo* study was rather inconclusive, the small fragments could still be considered as selective anti-CHIKV agents since they demonstrated either decent levels of anti-protease activity or some degree of viral replication inhibition *in vitro*. Though weak in potency, these fragments may actually form high-quality interactions likely in shallow binding pockets and they can be readily optimized into potent inhibitors. The identification of these fragments provide a starting point for the development of novel alphavirus protease inhibitors. In addition, this would encourage the use of fragment-based screening, since the approach is not a popular option for searching novel antivirals. Overall, the results present a set of low molecular weight molecules shown to either inhibit alphavirus nsP2 protease activity or viral replication.

CHAPTER 5. Conclusions and future perspectives

In this thesis, it has been shown through an *ab initio* model that CHIKV nsP2 adopted an extended conformation, whereby the RNA-binding groove in the helicase and catalytic cavity of protease domain were proposed to be distant from each other. The model also suggested that there is absence of intra-molecular interactions between the domains, which likely allowed them to function independently. Likewise, 2D class averages obtained from EM data showed similar elongated conformation. Given its extended nature and the presence of a short linker between the two domains, the protein is possibly flexible and could have a significant role whereby it could allow the protein to undergo conformational changes to support their functions during the various replication stages. However, it remained unclear if these models accurately demonstrated the suggested flexible nature of the protein. Therefore, the 3D structure of CHIKV nsP2 would need to be resolved to validate the proposed resolution-limited models and it could also provide more information on the functional significance of the linker region.

In addition, it would be interesting to study if altering the flexibility of the linker could have any implications on the biological function of nsP2. This was previously demonstrated for DENV NS3, which was also an elongated multidomain protein that possessed both helicase and protease domains like alphavirus nsP2 (Luo et al., 2008). It was shown in 3D structures that a flexible interdomain linker between the domains allowed the NS3 protein to exist in two different conformations and by either mutating or extending the linker region, the enzymatic activities of

the helicase domain were reduced (Luo et al., 2010). These studies suggested that the flexible linker was required to allow the NS3 protein to adopt several different configurations during its role in polyprotein protein processing and viral genome replication. If the alphavirus nsP2 linker is validated to be structurally significant like that of DENV NS3, future research could also involve the search for peptides or small molecules to disrupt the degree of freedom of the overall molecular conformation of nsP2 as part of the drug discovery effort towards antiviral therapy. However, if crystallization of nsP2 remains futile because of the probably flexible nature of the protein, more efforts could be focused on optimizing the negative staining protocol used in preparing samples for EM. This could include the choice of negative stains (e.g. uranyl vs tungstate) and the approaches used (e.g. negative stain vs carbon sandwich), where they can improve the quality of the images by reducing artifacts caused by dehydration, adsorption and incomplete stain embedding. Although the current EM data remained indistinct, the 2D projections could still be useful as a reference to facilitate the identification and selection of more particles for 3D reconstruction in the future.

The elucidation of the crystal structure of ONNV nsP2 protease offered a new target for drug discovery and design for antiviral therapeutic intervention. Like all other alphaviruses, ONNV nsP2 protease was revealed to consist of the protease and MTase-like subdomains. In addition, the close structural resemblance of ONNV nsP2 protease to CHIKV was highly anticipated due to their high sequence identities and this was further substantiated from the 3D structure itself. Although, the

crystal structure of ONNV nsP2 protease additionally supported our knowledge of alphavirus nsP2 protease, however, the alphavirus nsP2 protease-substrate complex structure is still not defined. Future work could involve soaking non-hydrolyzable peptide substrates into preformed ONNV nsP2 protease crystals. Nevertheless, care needs to be taken when interpreting the soaked complex structure because it might not display a biologically relevant conformation that was induced in solution upon substrate binding. This was demonstrated in a recent report whereby the ligand-bacterial tRNA-binding protein complexed structures from ligand-soaked and co-crystallized crystals revealed completely different conformations (Ehrmann et al., 2017). Thus, there is also a possibility that large conformational arrangements are required during substrate binding to the recognition pocket of nsP2 protease and this may not be visible in the soaked complex. Ideally, co-crystallization of the protease with substrate is preferred as it is likely to reveal the conformation that is biologically relevant.

The identification of Compound **2** as a moderately potent, non-competitive alphavirus nsP2 protease inhibitor showed that the fragment has the potential to become a novel antiviral compound. The destabilizing effect of Compound **2** on the protease further suggested its binding disrupted the formation of a stable protease-substrate complex, thereby inhibiting protease function. It is therefore important to elucidate the fragment-bound nsP2 protease complex structure in the future so as to reveal its binding site. A detailed analysis of the 3D structure of the complex will provide invaluable interaction information, which could then

be used to evolve the fragment through structure-based design and medicinal chemistry into a high affinity lead. Unique interactions can be incorporated into the binding site, either by expanding with additional binding elements to the scaffold or by joining fragments together. Eventually, this higher affinity compounds can be used as a lead for the discovery of a novel antiviral therapeutic.

References

- Abdelnabi, R., Amrun, S. N., Ng, L. F. P., Leyssen, P., Neyts, J., & Delang, L. (2017). Protein kinases C as potential host targets for the inhibition of chikungunya virus replication. *Antiviral Research*, 139, 79–87.
- Aggarwal, M., Dhindwal, S., Kumar, P., Kuhn, R. J., & Tomar, S. (2014). trans-Protease Activity and Structural Insights into the Active Form of the Alphavirus Capsid Protease. *Journal of Virology*, 88(21), 12242–12253.
- Aggarwal, M., Sharma, R., Kumar, P., Parida, M., & Tomar, S. (2015). Kinetic characterization of trans-proteolytic activity of Chikungunya virus capsid protease and development of a FRET-based HTS assay. *Scientific Reports*, 5(September), 14753.
- Ahola, T., & Kääriäinen, L. (1995). Reaction in alphavirus mRNA capping: formation of a covalent complex of nonstructural protein nsP1 with 7-methyl-GMP. *Proceedings of the National Academy of Sciences of the United States of America*, 92(2), 507–511.
- Ahola, T., Kujala, P., Tuittila, M., Blom, T., Laakkonen, P., Hinkkanen, A., & Auvinen, P. (2000). Effects of palmitoylation of replicase protein nsP1 on alphavirus infection. *Journal of Virology*, 74(15), 6725–6733.
- Ahola, T., Laakkonen, P., Vihinen, H., & Kääriäinen, L. (1997). Critical residues of Semliki Forest virus RNA capping enzyme involved in methyltransferase and guanylyltransferase-like activities. *Journal of Virology*, 71(1), 392–397.

- Ahola, T., Lampio, A., Auvinen, P., & Kaariainen, L. (1999). Semliki Forest virus mRNA capping enzyme requires association with anionic membrane phospholipids for activity. *EMBO J*, 18(11), 3164–3172.
- Ali Ou Alla, S., & Combe, B. (2011). Arthritis after infection with Chikungunya virus. *Best Pract Res Clin Rheumatol*, 25(3), 337–346.
- Amraoui, F., & Failloux, A.-B. (2016). Chikungunya: an unexpected emergence in Europe. *Current Opinion in Virology*, 21, 146–150.
- Arankalle, V. A., Shrivastava, S., Cherian, S., Gunjekar, R. S., Walimbe, A. M., Jadhav, S. M., Sudeep, A. B., & Mishra, A. C. (2007). Genetic divergence of Chikungunya viruses in India (1963-2006) with special reference to the 2005-2006 explosive epidemic. *Journal of General Virology*, 88(7), 1967–1976.
- Askin, S., Bond, T. E. H., Sorenson, A. E., Moreau, M. J. J., Antony, H., Davis, R. A., & Schaeffer, P. M. (2018). Selective protein unfolding: a universal mechanism of action for the development of irreversible inhibitors. *Chemical Communications*, 1738–1741.
- Assunção-Miranda, I., Cruz-Oliveira, C., & Da Poian, A. T. (2013). Molecular Mechanisms Involved in the Pathogenesis of Alphavirus-Induced Arthritis. *BioMed Research International*, 2013, 1–11.
- Atasheva, S., Gorchakov, R., English, R., Frolov, I., & Frolova, E. (2007). Development of Sindbis Viruses Encoding nsP2/GFP Chimeric Proteins and Their Application for Studying nsP2 Functioning. *Journal of Virology*, 81(10), 5046–5057.

- Bakar, F. A., & Ng, L. F. P. (2018). Nonstructural proteins of alphavirus—potential targets for drug development. *Viruses*, 10(2), 1–15.
- Bassetto, M., De Burghgraeve, T., Delang, L., Massarotti, A., Coluccia, A., Zonta, N., Gatti, V., Colombano, G., Sorba, G., Silvestri, R., Tron, G. C., Neyts, J., Leyssen, P., & Brancale, A. (2013). Computer-aided identification, design and synthesis of a novel series of compounds with selective antiviral activity against chikungunya virus. *Antiviral Res*, 98(1), 12–18.
- Bell, A. S., Solari, R., Tate, E. W., Guedán, A., Hopkins-Navratilova, I., Wilkinson, A. J., Newman, J., Ritzefeld, M., Swieboda, D. P., Johnston, S. L., Leatherbarrow, R. J., Morales-Sanfrutos, J., Mousnier, A., Asfor, A. S., Hutton, J. A., Pérez-Dorado, I., Tuthill, T. J., Brannigan, J. A., & Robinson, S. W. (2018). Fragment-derived inhibitors of human N-myristoyltransferase block capsid assembly and replication of the common cold virus. *Nature Chemistry*, 10(6), 599–606.
- Bernard, E., Solignat, M., Gay, B., Chazal, N., Higgs, S., Devaux, C., & Briant, L. (2010). Endocytosis of chikungunya virus into mammalian cells: Role of clathrin and early endosomal compartments. *PLoS ONE*, 5(7).
- Bessaud, M., Peyrefitte, C. N., Pastorino, B. A. M., Gravier, P., Tock, F., Boete, F., Tolou, H. J., & Grandadam, M. (2006). O'nyong-nyong virus, Chad. *Emerging Infectious Diseases*, 12(8), 1248–1250.
- Bessaud, M., Peyrefitte, C. N., Pastorino, B. A. M., Tock, F., Merle, O.,

- Colpart, J. J., Dehecq, J. S., Girod, R., Jaffar-Bandjee, M. C., Glass, P. J., Parker, M., Tolou, H. J., & Grandadam, M. (2006). Chikungunya virus strains, reunion island outbreak. *Emerging Infectious Diseases*, 12(10), 1604–1606.
- Bhalla, N., Sun, C., Metthew Lam, L. K., Gardner, C. L., Ryman, K. D., & Klimstra, W. B. (2016). Host translation shutoff mediated by non-structural protein 2 is a critical factor in the antiviral state resistance of Venezuelan equine encephalitis virus. *Virology*, 496, 147–165.
- Blanc, E., Roversi, P., Vonnrhein, C., Flensburg, C., Lea, S. M., & Bricogne, G. (2004). Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT. *Acta Crystallographica Section D: Biological Crystallography*, 60(12 I), 2210–2221.
- Bollag, G., Tsai, J., Zhang, J., Zhang, C., Ibrahim, P., Nolop, K., & Hirth, P. (2012). Vemurafenib: the first drug approved for BRAF-mutant cancer. *Nature Reviews Drug Discovery*, 11(11), 873–886.
- Borgherini, G., Poubeau, P., Staikowsky, F., Lory, M., Le Moullec, N., Becquart, J. P., Wengling, C., Michault, A., & Paganin, F. (2007). Outbreak of chikungunya on Reunion Island: early clinical and laboratory features in 157 adult patients. *Clin Infect Dis*, 44(11), 1401–1407.
- Borgherini, G., Paganin, F., Cotte, L., Jossaume, A., Arvin-Berod, C., Michault, A., Poubeau, P., & Gouix, A. (2008). Persistent Arthralgia Associated with Chikungunya Virus: A Study of 88 Adult Patients on Reunion Island. *Clinical Infectious Diseases*, 47(4), 469–475.
- Bourjot, M., Leyssen, P., Eydoux, C., Guillemot, J. C., Canard, B.,

- Rasoanaivo, P., Guéritte, F., & Litaudon, M. (2012). Chemical constituents of *Anacolosa pervilleana* and their antiviral activities. *Fitoterapia*, 83(6), 1076–1080.
- Brault, A. C., Foy, B. D., Myles, K. M., Kelly, C. L. H., Higgs, S., Weaver, S. C., Olson, K. E., Miller, B. R., & Powers, A. M. (2004). Infection patterns of o'nyong nyong virus in the malaria-transmitting mosquito, *Anopheles gambiae*. *Insect Molecular Biology*, 13(6), 625–635.
- Breakwell, L., Dosenovic, P., Karlsson Hedestam, G. B., D'Amato, M., Liljestrom, P., Fazakerley, J., & McInerney, G. M. (2007). Semliki Forest virus nonstructural protein 2 is involved in suppression of the type I interferon response. *J Virol*, 81(16), 8677–8684.
- Brighton, S. W. (1984). Chloroquine phosphate treatment of chronic Chikungunya arthritis. An open pilot study. *South African Medical Journal = Suid-Afrikaanse Tydskrif Vir Geneeskunde*, 66(6), 217–218.
- Briolant, S., Garin, D., Scaramozzino, N., Jouan, A., & Crance, J. M. (2004). In vitro inhibition of Chikungunya and Semliki Forest viruses replication by antiviral compounds: synergistic effect of interferon-alpha and ribavirin combination. *Antiviral Res*, 61(2), 111–117.
- Bruce, D., Cardew, E., Freitag-Pohl, S., & Pohl, E. (2019). How to Stabilize Protein: Stability Screens for Thermal Shift Assays and Nano Differential Scanning Fluorimetry in the Virus-X Project. *Journal of Visualized Experiments*, (144).
- Byler, K. G., Collins, J. T., Ogungbe, I. V., & Setzer, W. N. (2016).

- Alphavirus protease inhibitors from natural sources: A homology modeling and molecular docking investigation. *Comput Biol Chem*, 64, 163–184.
- Carey, D. E. (1971). Chikungunya and dengue: a case of mistaken identity? *Journal of the History of Medicine and Allied Sciences*, 26(3), 243–262.
- Carissimo, G., & Ng, L. F. P. (2019). Understanding Molecular Pathogenesis with Chikungunya Virus Research Tools. *Current Topics in Microbiology and Immunology*.
- Chan, Y., Teo, T., Utt, A., Tan, J. J., Amrun, S. N., Abu Bakar, F., Yee, W., Becht, E., Lee, C. Y., Lee, B., Rajarethinam, R., Newell, E., Merits, A., Carissimo, G., Lum, F., & Ng, L. F. (2019). Mutating chikungunya virus non-structural protein produces potent live-attenuated vaccine candidate. *EMBO Molecular Medicine*, e10092.
- Chen, K., Kam, Y.-W., Lin, R. T., Ng, M. M.-L., Ng, L. F., & Chu, J. J. (2013). Comparative analysis of the genome sequences and replication profiles of chikungunya virus isolates within the East, Central and South African (ECSA) lineage. *Virology Journal*, 10(1), 169.
- Chen, M. W., Tan, Y. B., Zheng, J., Zhao, Y., Lim, B. T., Cornvik, T., Lescar, J., Ng, L. F. P., & Luo, D. (2017). Chikungunya virus nsP4 RNA-dependent RNA polymerase core domain displays detergent-sensitive primer extension and terminal adenylyltransferase activities. *Antiviral Research*, 143, 38–47.
- Cheng, R. H., Kuhn, R. J., Olson, N. H., Rossmann, M. G., Choi, H. K.,

- Smith, T. J., & Baker, T. S. (1995). Nucleocapsid and glycoprotein organization in an enveloped virus. *Cell*, 80(4), 621–630.
- Cheung, J., Franklin, M., Mancina, F., Rudolph, M., Cassidy, M., Gary, E., Burshteyn, F., & Love, J. (2011). Structure of the Chikungunya virus nsP2 protease. doi: 10.2210/pdb3TRK/pdb .
- Chung, D. H., Jonsson, C. B., Tower, N. A., Chu, Y. K., Sahin, E., Golden, J. E., Noah, J. W., Schroeder, C. E., Sotsky, J. B., Sosa, M. I., Cramer, D. E., McKellip, S. N., Rasmussen, L., White, E. L., Schmaljohn, C. S., Julander, J. G., Smith, J. M., Filone, C. M., Connor, J. H., Sakurai, Y., & Davey, R. A. (2014). Discovery of a Novel Compound with Anti-Venezuelan Equine Encephalitis Virus Activity That Targets the Nonstructural Protein 2. *PLoS Pathogens*, 10(6).
- Cross, R. K. (1983). Identification of a unique guanine-7-methyltransferase in Semliki Forest virus (SFV) infected cell extracts. *Virology*, 130(2), 452–463.
- Das, P. K., Merits, A., & Lulla, A. (2014). Functional Cross-talk between Distant Domains of Chikungunya Virus Non-structural Protein 2 Is Decisive for Its RNA-modulating Activity. *Journal of Biological Chemistry*, 289(9), 5635–5653.
- Das, P. K., Puusepp, L., Varghese, F. S., Utt, A., Ahola, T., Kananovich, D. G., Lopp, M., Merits, A., & Karelson, M. (2016). Design and validation of novel chikungunya virus protease inhibitors. *Antimicrobial Agents and Chemotherapy*, 60(12), AAC.01421-16.
- De Clercq, E., & Neyts, J. (2009). *Antiviral Agents Acting as DNA or*

- RNA Chain Terminators* (H.-G. Kräusslich & R. Bartenschlager, Eds.). In (pp. 53–84).
- De Curtis, I., & Simons, K. (1988). Dissection of Semliki Forest virus glycoprotein delivery from the trans-Golgi network to the cell surface in permeabilized BHK cells. *Proceedings of the National Academy of Sciences*, 85(21), 8052–8056.
- De Lamballerie, X., Ninove, L., & Charrel, R. N. (2009). Antiviral treatment of chikungunya virus infection. *Infectious Disorders Drug Targets*, 9(2), 101–104.
- Delang, L., Guerrero, N. S., Tas, A., Quérat, G., Pastorino, B., Froeyen, M., Dallmeier, K., Jochmans, D., Herdewijn, P., Bello, F., Snijder, E. J., de Lamballerie, X., Martina, B., Neyts, J., van Hemert, M. J., & Leyssen, P. (2014). Mutations in the chikungunya virus non-structural proteins cause resistance to favipiravir (T-705), a broad-spectrum antiviral. *Journal of Antimicrobial Chemotherapy*, 69(10), 2770–2784.
- Delogu, I., & de Lamballerie, X. (2011). Chikungunya disease and chloroquine treatment. *Journal of Medical Virology*, 83(6), 1058–1059.
- Delogu, I., Pastorino, B., Baronti, C., Nougairède, A., Bonnet, E., & de Lamballerie, X. (2011). In vitro antiviral activity of arbidol against Chikungunya virus and characteristics of a selected resistant mutant. *Antiviral Research*, 90(3), 99–107.
- Dudha, N., Rana, J., Rajasekharan, S., Gabrani, R., Gupta, A., Chaudhary, V. K., & Gupta, S. (2015). Host–pathogen interactome

- analysis of Chikungunya virus envelope proteins E1 and E2. *Virus Genes*, 50(2), 200–209.
- Ehrmann, F. R., Stojko, J., Metz, A., Debaene, F., Barandun, L. J., Heine, A., Diederich, F., Cianférani, S., Reuter, K., & Klebe, G. (2017). Soaking suggests “alternative facts”: Only cocrystallization discloses major ligand-induced interface rearrangements of a homodimeric tRNA-binding protein indicating a novel mode-of-inhibition. *PLoS ONE*, 12(4), 1–30.
- Ehteshami, M., Tao, S., Zandi, K., Hsiao, H. M., Jiang, Y., Hammond, E., Amblard, F., Russell, O. O., Merits, A., & Schinazi, R. F. (2017). Characterization of β -D-N4-hydroxycytidine as a novel inhibitor of Chikungunya virus. *Antimicrobial Agents and Chemotherapy*, 61(4), 1–6.
- El-Labbad, E. M., Ismail, M. A. H., Abou Ei Ella, D. A., Ahmed, M., Wang, F., Barakat, K. H., & Abouzid, K. A. M. (2015). Discovery of novel peptidomimetics as irreversible CHIKV NsP2 protease inhibitors using quantum mechanical-based ligand descriptors. *Chemical Biology and Drug Design*, 86(6), 1518–1527.
- El Kassas, M., Elbaz, T., Hafez, E., Wifi, M. N., & Esmat, G. (2017). Discovery and preclinical development of dasabuvir for the treatment of hepatitis C infection. *Expert Opinion on Drug Discovery*, 12(6), 635–642.
- Emsley, P., & Cowtan, K. (2004). Coot: Model-building tools for molecular graphics. *Acta Crystallographica Section D: Biological Crystallography*, 60(12 I), 2126–2132.

- Fox, J. M., Long, F., Edeling, M. A., Lin, H., Van Duijl-Richter, M. K. S., Fong, R. H., Kahle, K. M., Smit, J. M., Jin, J., Simmons, G., Doranz, B. J., Crowe, J. E., Fremont, D. H., Rossmann, M. G., & Diamond, M. S. (2015). Broadly Neutralizing Alphavirus Antibodies Bind an Epitope on E2 and Inhibit Entry and Egress. *Cell*, 163(5), 1095–1107.
- Franco, E. J., Rodriguez, J. L., Pomeroy, J. J., Hanrahan, K. C., & Brown, A. N. (2018). The effectiveness of antiviral agents with broad-spectrum activity against chikungunya virus varies between host cell lines. *Antiviral Chemistry and Chemotherapy*, 26, 1–7.
- Frolov, I., Garmashova, N., Atasheva, S., & Frolova, E. I. (2009). Random insertion mutagenesis of sindbis virus nonstructural protein 2 and selection of variants incapable of downregulating cellular transcription. *J Virol*, 83(18), 9031–9044.
- Fros, J. J., Major, L. D., Scholte, F. E. M., Gardner, J., van Hemert, M. J., Suhrbier, A., & Pijlman, G. P. (2015). Chikungunya virus non-structural protein 2-mediated host shut-off disables the unfolded protein response. *Journal of General Virology*, 96(3), 580–589.
- Furuta, Y., Gowen, B. B., Takahashi, K., Shiraki, K., Smee, D. F., & Barnard, D. L. (2013). Favipiravir (T-705), a novel viral RNA polymerase inhibitor. *Antiviral Research*, 100(2), 446–454.
- Gardner, J., Anraku, I., Le, T. T., Larcher, T., Major, L., Roques, P., Schroder, W. A., Higgs, S., & Suhrbier, A. (2010). Chikungunya Virus Arthritis in Adult Wild-Type Mice. *Journal of Virology*, 84(16), 8021–8032.

- Gérardin, P., Fianu, A., Malvy, D., Mussard, C., Boussaïd, K., Rollot, O., Michault, A., Gaüzere, B. A., Bréart, G., & Favier, F. (2011). Perceived morbidity and community burden after a Chikungunya outbreak: The TELECHIK survey, a population-based cohort study. *BMC Medicine*, 9(1), 5.
- Gomez De Cedrón, M., Ehsani, N., Mikkola, M. L., García, J. A., & Kääriäinen, L. (1999). RNA helicase activity of Semliki Forest virus replicase protein NSP2. *FEBS Letters*, 448(1), 19–22.
- Gorbalenya, A. E., Koonin, E. V., & Lai, M. M.-C. (1991). Putative papain-related thiol proteases of positive-strand RNA viruses Identification of rubi- and aphthovirus proteases and delineation of a novel conserved domain associated with proteases of rubi-, α - and coronaviruses. *FEBS Letters*, 288(1–2), 201–205.
- Gorchakov, R., Frolova, E., & Frolov, I. (2005). Inhibition of transcription and translation in Sindbis virus-infected cells. *J Virol*, 79(15), 9397–9409.
- Goyal, M., Chauhan, A., Goyal, V., Jaiswal, N., Singh, S., & Singh, M. (2018). Recent development in the strategies projected for chikungunya vaccine in humans. *Drug Design, Development and Therapy*, Volume 12, 4195–4206.
- Graham, F. L., Smiley, J., Russell, W. C., & Nairn, R. (1977). Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *The Journal of General Virology*, 36(1), 59–74.
- Hahn, Y. S., Strauss, E. G., & Strauss, J. H. (1989). Mapping of RNA-

- temperature-sensitive mutants of Sindbis virus: assignment of complementation groups A, B, and G to nonstructural proteins. *Journal of Virology*, 63(7), 3142–3150.
- Hardy, W. R., & Strauss, J. H. (1989). 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. *Journal of Virology*, 63(11), 4653–4664.
- Her, Z., Kam, Y.-W., Lin, R. T. P., & Ng, L. F. P. (2009). Chikungunya: a bending reality. *Microbes and Infection*, 11(14–15), 1165–1176.
- Her, Z., Teng, T., Tan, J. J., Teo, T., Kam, Y., Lum, F., Lee, W. W., Gabriel, C., Melchiotti, R., Andiappan, A. K., Lulla, V., Lulla, A., Win, M. K., Chow, A., Biswas, S. K., Leo, Y., Lecuit, M., Merits, A., Rénia, L., & Ng, L. F. (2015). Loss of TLR3 aggravates CHIKV replication and pathology due to an altered virus-specific neutralizing antibody response. *EMBO Molecular Medicine*, 7(1), 24–41.
- Hu, X., Compton, J. R., Leary, D. H., Olson, M. A., Lee, M. S., Cheung, J., Ye, W., Ferrer, M., Southall, N., Jadhav, A., Morazzani, E. M., Glass, P. J., Marugan, J., & Legler, P. M. (2016). Kinetic, Mutational, and Structural Studies of the Venezuelan Equine Encephalitis Virus Nonstructural Protein 2 Cysteine Protease. *Biochemistry*, 55(21), 3007–3019.
- Hyde, J. L., Gardner, C. L., Kimura, T., White, J. P., Liu, G., Trobaugh, D. W., Huang, C., Tonelli, M., Paessler, S., Takeda, K., Klimstra, W. B., Amarasinghe, G. K., & Diamond, M. S. (2014). A Viral RNA

Structural Element Alters Host Recognition of Nonself RNA.

Science, 343(6172), 783–787.

Jose, J., Snyder, J. E., & Kuhn, R. J. (2009). A structural and functional perspective of alphavirus replication and assembly. *Future Microbiology*, 4(7), 837–856.

Judith, D., Mostowy, S., Bourai, M., Gangneux, N., Lelek, M., Lucas-Hourani, M., Cayet, N., Jacob, Y., Prévost, M., Pierre, P., Tangy, F., Zimmer, C., Vidalain, P., Couderc, T., & Lecuit, M. (2013). Species-specific impact of the autophagy machinery on Chikungunya virus infection. *EMBO Reports*, 14(6), 534–544.

Kääriäinen, L., & Ahola, T. (2002). Functions of alphavirus nonstructural proteins in RNA replication. *Progress in Nucleic Acid Research and Molecular Biology*, 71, 187–222.

Karlas, A., Berre, S., Couderc, T., Varjak, M., Braun, P., Meyer, M., Gangneux, N., Karo-Astover, L., Weege, F., Raftery, M., Schönrich, G., Klemm, U., Wurzlbauer, A., Bracher, F., Merits, A., Meyer, T. F., & Lecuit, M. (2016). A human genome-wide loss-of-function screen identifies effective chikungunya antiviral drugs. *Nature Communications*, 7(1), 11320.

Karpe, Y. A., Aher, P. P., & Lole, K. S. (2011). NTPase and 5'-RNA Triphosphatase Activities of Chikungunya Virus nsP2 Protein. *PLoS ONE*, 6(7), e22336.

Karpe, Y. A., & Lole, K. S. (2010). NTPase and 5' to 3' RNA Duplex-Unwinding Activities of the Hepatitis E Virus Helicase Domain. *Journal of Virology*, 84(7), 3595–3602.

- Kaur, P., & Chu, J. J. H. (2013). Chikungunya virus: An update on antiviral development and challenges. *Drug Discovery Today*, 18(19–20), 969–983.
- Khan, A. H., Morita, K., del Carmen Parquet, M., Hasebe, F., Mathenge, E. G. M., & Igarashi, A. (2002). Complete nucleotide sequence of chikungunya virus and evidence for an internal polyadenylation site. *Journal of General Virology*, 83(12), 3075–3084.
- Kielian, M., & Rey, F. A. (2006). Virus membrane-fusion proteins: more than one way to make a hairpin. *Nat Rev Microbiol*, 4(1), 67–76.
- Kiwanuka, N., Sanders, E. J., Rwaguma, E. B., Kawamata, J., Ssengooba, F. P., Najjemba, R., Were, W. A., Lamunu, M., Bagambisa, G., Burkot, T. R., Dunster, L., Lutwama, J. J., Martin, D. A., Cropp, C. B., Karabatsos, N., Lanciotti, R. S., Tsai, T. F., & Campbell, G. L. (1999). O’Nyong-Nyong Fever in South-Central Uganda, 1996–1997: Clinical Features and Validation of a Clinical Case Definition for Surveillance Purposes. *Clinical Infectious Diseases*, 29(5), 1243–1250.
- Konarev, P. V., Petoukhov, M. V., & Svergun, D. I. (2001). MASSHA – a graphics system for rigid-body modelling of macromolecular complexes against solution scattering data. *Journal of Applied Crystallography*, 34(4), 527–532.
- Laakkonen, P., Ahola, T., & Kaariainen, L. (1996). The effects of palmitoylation on membrane association of Semliki forest virus RNA capping enzyme. *J Biol Chem*, 271(45), 28567–28571.
- Laakkonen, P., Hyvönen, M., Peränen, J., & Kääriäinen, L. (1994).

- Expression of Semliki Forest virus nsP1-specific methyltransferase in insect cells and in *Escherichia coli*. *Journal of Virology*, 68(11), 7418–7425.
- Lam, S., Chen, H., Chen, C. K., Min, N., & Chu, J. J. H. (2015). Antiviral phosphorodiamidate morpholino oligomers are protective against chikungunya virus infection on cell-based and murine models. *Scientific Reports*, 5(February), 1–14.
- Lamballerie, X. De, Boisson, V., Reynier, J.-C., Enault, S., Charrel, R. N., Flahault, A., Roques, P., & Grand, R. Le. (2008). On Chikungunya Acute Infection and Chloroquine Treatment. *Vector-Borne and Zoonotic Diseases*, 8(6), 837–840.
- Lamoree, B., & Hubbard, R. E. (2017). Current perspectives in fragment-based lead discovery (FBLD). *Essays in Biochemistry*, 61(5), 453–464.
- Lampio, A., Kilpeläinen, I., Pesonen, S., Karhi, K., Auvinen, P., Somerharju, P., & Kääriäinen, L. (2000). Membrane Binding Mechanism of an RNA Virus-capping Enzyme. *Journal of Biological Chemistry*, 275(48), 37853–37859.
- Lanciotti, R. S., Ludwig, M. L., Rwaguma, E. B., Lutwama, J. J., Kram, T. M., Karabatsos, N., Cropp, B. C., & Miller, B. R. (1998). Emergence of epidemic O'nyong-nyong fever in uganda after a 35-year absence: Genetic characterization of the virus. *Virology*, 252(1), 258–268.
- Lastarza, M. W., Grakoui, A., & Rice, C. M. (1994). Deletion and duplication mutations in the C-terminal nonconserved region of

- Sindbis virus nsP3: Effects on phosphorylation and on virus replication in vertebrate and invertebrate cells. *Virology*, Vol. 202, pp. 224–232.
- LaStarza, M. W., Lemm, J. A., & Rice, C. M. (1994). Genetic analysis of the nsP3 region of Sindbis virus: evidence for roles in minus-strand and subgenomic RNA synthesis. *Journal of Virology*, 68(9), 5781–5791.
- Law, Y.-S., Utt, A., Tan, Y. B., Zheng, J., Wang, S., Chen, M. W., Griffin, P. R., Merits, A., & Luo, D. (2019). Structural insights into RNA recognition by the Chikungunya virus nsP2 helicase. *Proceedings of the National Academy of Sciences*, (8), 201900656.
- Lemm, J. A., Bergqvist, A., Read, C. M., & Rice, C. M. (1998). Template-dependent initiation of Sindbis virus RNA replication in vitro. *J Virol*, 72(8), 6546–6553.
- Lemm, J. A., Rumenapf, T., Strauss, E. G., Strauss, J. H., & Rice, C. M. (1994). Polypeptide requirements for assembly of functional Sindbis virus replication complexes: a model for the temporal regulation of minus- and plus-strand RNA synthesis. *EMBO J*, 13(12), 2925–2934.
- Leparc-Goffart, I., Nougairede, A., Cassadou, S., Prat, C., & De Lamballerie, X. (2014). Chikungunya in the Americas. *The Lancet*, 383(9916), 514.
- Lescar, J., Roussel, A., Wien, M. W., Navaza, J., Fuller, S. D., Wengler, G., Wengler, G., & Rey, F. A. (2001). The Fusion glycoprotein shell of Semliki Forest virus: an icosahedral assembly primed for

- fusogenic activation at endosomal pH. *Cell*, 105(1), 137–148.
- Levinson, R. S., Strauss, J. H., & Strauss, E. G. (1990). Complete sequence of the genomic RNA of O’Nyong-Nyong virus and its use in the construction of alphavirus phylogenetic trees. *Virology*, 175(1), 110–123.
- Li, C., Guillén, J., Rabah, N., Blanjoie, A., Debart, F., Vasseur, J.-J., Canard, B., Decroly, E., & Coutard, B. (2015). mRNA Capping by Venezuelan Equine Encephalitis Virus nsP1: Functional Characterization and Implications for Antiviral Research. *Journal of Virology*, 89(16), 8292–8303.
- Liu, Y., Lim, B. H., Jiang, W. W., Flentge, C. A., Hutchinson, D. K., Madigan, D. L., Randolph, J. T., Wagner, R., Maring, C. J., Kati, W. M., & Molla, A. (2012). Identification of aryl dihydrouracil derivatives as palm initiation site inhibitors of HCV NS5B polymerase. *Bioorganic & Medicinal Chemistry Letters*, 22(11), 3747–3750.
- Liumbruno, G. M., Calteri, D., Petropulacos, K., Mattivi, A., Po, C., Macini, P., Tomasini, I., Zucchelli, P., Silvestri, A. R., Sambri, V., Pupella, S., Catalano, L., Piccinini, V., Calizzani, G., & Grazzini, G. (2008). The Chikungunya epidemic in Italy and its repercussion on the blood system. *Blood Transfusion*, 6(4), 199–210.
- Lo, M. C., Aulabaugh, A., Jin, G., Cowling, R., Bard, J., Malamas, M., & Ellestad, G. (2004). Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. *Analytical Biochemistry*, 332(1), 153–159.
- Lulla, A., Lulla, V., & Merits, A. (2012). Macromolecular Assembly-

- Driven Processing of the 2/3 Cleavage Site in the Alphavirus Replicase Polyprotein. *Journal of Virology*, 86(1), 553–565.
- Lulla, A., Lulla, V., Tints, K., Ahola, T., & Merits, A. (2006). Molecular Determinants of Substrate Specificity for Semliki Forest Virus Nonstructural Protease. *Journal of Virology*, 80(11), 5413–5422.
- Lulla, V., Karo-Astover, L., Rausalu, K., Saul, S., Merits, A., & Lulla, A. (2018). Timeliness of proteolytic events is prerequisite for efficient functioning of the alphaviral replicase. *Journal of Virology*, 92(14), e00151-18.
- Lum, F.-M., Teo, T.-H., Lee, W. W. L., Kam, Y.-W., Rénia, L., & Ng, L. F. P. (2013). An Essential Role of Antibodies in the Control of Chikungunya Virus Infection. *The Journal of Immunology*, 190(12), 6295–6302.
- Lumsden, W. H. R. (1955). *An epidemic of virus disease in the Southern Province, Tanganyika Territory, in 1952-1953*. 49(1), 33–57.
- Luo, D., Wei, N., Doan, D. N., Paradkar, P. N., Chong, Y., Davidson, A. D., Kotaka, M., Lescar, J., & Vasudevan, S. G. (2010). Flexibility between the protease and helicase domains of the dengue virus NS3 protein conferred by the linker region and its functional implications. *Journal of Biological Chemistry*, 285(24), 18817–18827.
- Luo, D., Xu, T., Hunke, C., Gruber, G., Vasudevan, S. G., & Lescar, J. (2008). Crystal Structure of the NS3 Protease-Helicase from Dengue Virus. *Journal of Virology*, 82(1), 173–183.
- Mancini, E. J., Clarke, M., Gowen, B. E., Rutten, T., & Fuller, S. D.

- (2000). Cryo-electron microscopy reveals the functional organization of an enveloped virus, Semliki Forest virus. *Mol Cell*, 5(2), 255–266.
- Manimunda, S. P., Vijayachari, P., Uppoor, R., Sugunan, A. P., Singh, S. S., Rai, S. K., Sudeep, A. B., Muruganandam, N., Chaitanya, I. K., & Guruprasad, D. R. (2010). Clinical progression of chikungunya fever during acute and chronic arthritic stages and the changes in joint morphology as revealed by imaging. *Trans R Soc Trop Med Hyg*, 104(6), 392–399.
- Marsh, M., & Helenius, A. (2006). Virus entry: open sesame. *Cell*, 124(4), 729–740.
- Matayoshi, E. D., Wang, G. T., Krafft, G. A., & Erickson, J. (1990). Novel fluorogenic substrates for assaying retroviral proteases by resonance energy transfer. *Science (New York, N.Y.)*, 247(4945), 954–958.
- Merits, A., Vasiljeva, L., Ahola, T., Kääriäinen, L., & Auvinen, P. (2001). Proteolytic processing of Semliki Forest virus-specific non-structural polyprotein by nsP2 protease. *The Journal of General Virology*, 82(Pt 4), 765–773.
- Miner, J. J., Cook, L. E., Hong, J. P., Smith, A. M., Richner, J. M., Shimak, R. M., Young, A. R., Monte, K., Poddar, S., Crowe, J. E., Lenschow, D. J., & Diamond, M. S. (2017). Therapy with CTLA4-Ig and an antiviral monoclonal antibody controls chikungunya virus arthritis. *Science Translational Medicine*, 9(375), 1–10.
- Morrison, T. E., Oko, L., Montgomery, S. A., Whitmore, A. C., Lotstein,

- A. R., Gunn, B. M., Elmore, S. A., & Heise, M. T. (2011). A mouse model of chikungunya virus-induced musculoskeletal inflammatory disease: Evidence of arthritis, tenosynovitis, myositis, and persistence. *American Journal of Pathology*, 178(1), 32–40.
- Mukhopadhyay, S., Zhang, W., Gabler, S., Chipman, P. R., Strauss, E. G., Strauss, J. H., Baker, T. S., Kuhn, R. J., & Rossmann, M. G. (2006). Mapping the structure and function of the E1 and E2 glycoproteins in alphaviruses. *Structure*, 14(1), 63–73.
- Murray, C. W., & Rees, D. C. (2009). The rise of fragment-based drug discovery. *Nature Chemistry*, 1(3), 187–192.
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., & Vagin, A. A. (2011). REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallographica. Section D, Biological Crystallography*, 67(Pt 4), 355–367.
- Myles, K. M., Kelly, C. L. H., Ledermann, J. P., & Powers, A. M. (2006). Effects of an Opal Termination Codon Preceding the nsP4 Gene Sequence in the O’Nyong-Nyong Virus Genome on *Anopheles gambiae* Infectivity. *Journal of Virology*, 80(10), 4992–4997.
- Narwal, M., Singh, H., Pratap, S., Malik, A., Kuhn, R. J., Kumar, P., & Tomar, S. (2018). Crystal structure of chikungunya virus nsP2 cysteine protease reveals a putative flexible loop blocking its active site. *International Journal of Biological Macromolecules*, 116(May), 451–462.
- Nguyen, P. T. V., Yu, H., & Keller, P. A. (2015). Identification of

- chikungunya virus nsP2 protease inhibitors using structure-base approaches. *Journal of Molecular Graphics & Modelling*, 57, 1–8.
- Ninla-aesong, P., Mitarnun, W., & Noipha, K. (2019). Proinflammatory Cytokines and Chemokines as Biomarkers of Persistent Arthralgia and Severe Disease After Chikungunya Virus Infection: A 5-Year Follow-Up Study in Southern Thailand. *Viral Immunology*, 32(10), 442–452.
- Nishikiori, M., Sugiyama, S., Xiang, H., Niiyama, M., Ishibashi, K., Inoue, T., Ishikawa, M., Matsumura, H., & Katoh, E. (2012). Crystal Structure of the Superfamily 1 Helicase from Tomato Mosaic Virus. *Journal of Virology*, 86(14), 7565–7576.
- Nothias-Scaglia, L. F., Pannecouque, C., Renucci, F., Delang, L., Neyts, J., Roussi, F., Costa, J., Leyssen, P., Litaudon, M., & Paolini, J. (2015). Antiviral Activity of Diterpene Esters on Chikungunya Virus and HIV Replication. *Journal of Natural Products*, 78(6), 1277–1283.
- Paessler, S., Rijnbrand, R., Stein, D. A., Ni, H., Yun, N. E., Dziuba, N., Borisevich, V., Seregin, A., Ma, Y., Blouch, R., Iversen, P. L., & Zacks, M. A. (2008). Inhibition of alphavirus infection in cell culture and in mice with antisense morpholino oligomers. *Virology*, 376(2), 357–370.
- Parashar, D., & Cherian, S. (2014). Antiviral perspectives for chikungunya virus. *BioMed Research International*, 2014.
- Paredes, A. M., Brown, D. T., Rothnagel, R., Chiu, W., Schoepp, R. J., Johnston, R. E., & Prasad, B. V. (1993). Three-dimensional

- structure of a membrane-containing virus. *Proc Natl Acad Sci U S A*, 90(19), 9095–9099.
- Partidos, C. D., Paykel, J., Weger, J., Borland, E. M., Powers, A. M., Seymour, R., Weaver, S. C., Stinchcomb, D. T., & Osorio, J. E. (2012). Cross-protective immunity against o'nyong-nyong virus afforded by a novel recombinant chikungunya vaccine. *Vaccine*, 30(31), 4638–4643.
- Pastorino, B. A. M., Peyrefitte, C. N., Almeras, L., Grandadam, M., Rolland, D., Tolou, H. J., & Bessaud, M. (2008). Expression and biochemical characterization of nsP2 cysteine protease of Chikungunya virus. *Virus Research*, 131(2), 293–298.
- Petoukhov, M. V., Franke, D., Shkumatov, A. V., Tria, G., Kikhney, A. G., Gajda, M., Gorba, C., Mertens, H. D. T., Konarev, P. V., & Svergun, D. I. (2012). New developments in the ATSAS program package for small-angle scattering data analysis. *Journal of Applied Crystallography*, 45(2), 342–350.
- Peyrefitte, C. N., Rousset, D., Pastorino, B. A. M., Pouillot, R., Bessaud, M., Tock, F., Mansaray, H., Merle, O. L., Pascual, A. M., Paupy, C., Vessiere, A., Imbert, P., Tchendjou, P., Durand, J.-P., Tolou, H. J., & Grandadam, M. (2007). Chikungunya Virus, Cameroon, 2006. *Emerging Infectious Diseases*, 13(5), 768–771.
- Pialoux, G., Gauzere, B. A., Jaureguiberry, S., & Strobel, M. (2007). Chikungunya, an epidemic arbovirosis. *Lancet Infect Dis*, 7(5), 319–327.
- Plaskon, N. E., Adelman, Z. N., & Myles, K. M. (2009). Accurate strand-

- specific quantification of viral RNA. *PLoS ONE*, 4(10).
- Powers, A. M., Brault, A. C., Tesh, R. B., & Weaver, S. C. (2000). Re-emergence of chikungunya and o'nyong-nyong viruses: Evidence for distinct geographical lineages and distant evolutionary relationships. *Journal of General Virology*, 81(2), 471–479.
- Ramakrishnan, C., Kutumbarao, N. H. V., Suhitha, S., & Velmurugan, D. (2017). Structure–function relationship of Chikungunya nsP2 protease: A comparative study with papain. *Chemical Biology and Drug Design*, 89(5), 772–782.
- Rathore, A. P. S., Ng, M.-L., & Vasudevan, S. G. (2013). Differential unfolded protein response during Chikungunya and Sindbis virus infection: CHIKV nsP4 suppresses eIF2 α phosphorylation. *Virology Journal*, 10(1), 36.
- Rausalu, K., Utt, A., Quirin, T., Varghese, F. S., Zusinaite, E., Das, P. K., Ahola, T., & Merits, A. (2016). Chikungunya virus infectivity, RNA replication and non-structural polyprotein processing depend on the nsP2 protease's active site cysteine residue. *Sci Rep*, 6, 37124.
- Ravichandran, R., & Manian, M. (2008). Ribavirin therapy for Chikungunya arthritis. *Journal of Infection in Developing Countries*, 2(2), 140–142.
- Rezza, G., Nicoletti, L., Angelini, R., Romi, R., Finarelli, A., Panning, M., Cordioli, P., Fortuna, C., Boros, S., Magurano, F., Silvi, G., Angelini, P., Dottori, M., Ciufolini, M., Majori, G., & Cassone, A. (2007). Infection with chikungunya virus in Italy: an outbreak in a temperate

- region. *Lancet*, 370(9602), 1840–1846.
- Rikkonen, M., Peränen, J., & Kääriäinen, L. (1994). ATPase and GTPase activities associated with Semliki Forest virus nonstructural protein nsP2. *Journal of Virology*, 68(9), 5804–5810.
- Robert, X., & Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research*, 42(W1), 320–324.
- Rothan, H. A., Bahrani, H., Abdulrahman, A. Y., Mohamed, Z., Teoh, T. C., Othman, S., Rashid, N. N., Rahman, N. A., & Yusof, R. (2016). Mefenamic acid in combination with ribavirin shows significant effects in reducing chikungunya virus infection in vitro and in vivo. *Antiviral Research*, 127, 50–56.
- Rothan, H. A., Bahrani, H., Mohamed, Z., Teoh, T. C., Shankar, E. M., Rahman, N. A., & Yusof, R. (2015). A Combination of Doxycycline and Ribavirin Alleviated Chikungunya Infection. *PLOS ONE*, 10(5), e0126360.
- Rowe, T., Banner, D., Farooqui, A., Ng, D. C. K., Kelvin, A. A., Rubino, S., Huang, S. S. H., Fang, Y., & Kelvin, D. J. (2010). In vivo ribavirin activity against severe pandemic H1N1 influenza A/Mexico/4108/2009. *Journal of General Virology*, 91(12), 2898–2906.
- Rubach, J. K., Wasik, B. R., Rupp, J. C., Kuhn, R. J., Hardy, R. W., & Smith, J. L. (2009). Characterization of purified Sindbis virus nsP4 RNA-dependent RNA polymerase activity in vitro. *Virology*, 384(1), 201–208.

- Rulli, N. E., Guglielmotti, A., Mangano, G., Rolph, M. S., Apicella, C., Zaid, A., Suhrbier, A., & Mahalingam, S. (2009). Amelioration of alphavirus-induced arthritis and myositis in a mouse model by treatment with bindarit, an inhibitor of monocyte chemotactic proteins. *Arthritis & Rheumatism*, 60(8), 2513–2523.
- Rulli, N. E., Rolph, M. S., Srikiatkachorn, A., Anantapreecha, S., Guglielmotti, A., & Mahalingam, S. (2011). Protection From Arthritis and Myositis in a Mouse Model of Acute Chikungunya Virus Disease by Bindarit, an Inhibitor of Monocyte Chemotactic Protein-1 Synthesis. *The Journal of Infectious Diseases*, 204(7), 1026–1030.
- Rupp, J. C., Jundt, N., & Hardy, R. W. (2011). Requirement for the Amino-Terminal Domain of Sindbis Virus nsP4 during Virus Infection. *Journal of Virology*, 85(7), 3449–3460.
- Rupp, J. C., Sokoloski, K. J., Gebhart, N. N., & Hardy, R. W. (2015). Alphavirus RNA synthesis and non-structural protein functions. *J Gen Virol*, 96(9), 2483–2500.
- Russo, A. T., Malmstrom, R. D., White, M. A., & Watowich, S. J. (2010). Structural basis for substrate specificity of alphavirus nsP2 proteases. *J Mol Graph Model*, 29(1), 46–53.
- Russo, A. T., White, M. A., & Watowich, S. J. (2006). The crystal structure of the Venezuelan equine encephalitis alphavirus nsP2 protease. *Structure (London, England : 1993)*, 14(9), 1449–1458.
- Rwaguma, E. B., Lutwama, J. J., Sempala, S. D., Kiwanuka, N., Kamugisha, J., Okware, S., Bagambisa, G., Lanciotti, R., Roehrig, J. T., & Gubler, D. J. (1997). Emergence of epidemic O'nyong-

- nyong fever in southwestern Uganda, after an absence of 35 years. *Emerging Infectious Diseases*, 3(1), 77.
- Ryman, K. D., & Klimstra, W. B. (2008). Host responses to alphavirus infection. *Immunological Reviews*, 225(1), 27–45.
- Saha, A., Acharya, B. N., Priya, R., Tripathi, N. K., Shrivastava, A., Rao, M. K., Kesari, P., Narwal, M., Tomar, S., Bhagyawant, S. S., Parida, M., & Dash, P. K. (2018). Development of nsP2 protease based cell free high throughput screening assay for evaluation of inhibitors against emerging Chikungunya virus. *Scientific Reports*, 8(1), 1–13.
- Saisawang, C., Saitornuang, S., Sillapee, P., Ubol, S., Smith, D. R., & Kettermann, A. J. (2015). Chikungunya nsP2 protease is not a papain-like cysteine protease and the catalytic dyad cysteine is interchangeable with a proximal serine. *Scientific Reports*, 5(1), 17125.
- Saisawang, C., Sillapee, P., Sinsirimongkol, K., Ubol, S., Smith, D. R., & Kettermann, A. J. (2015). Full length and protease domain activity of chikungunya virus nsP2 differ from other alphavirus nsP2 proteases in recognition of small peptide substrates. *Bioscience Reports*, 35(3), 1–9.
- Saxton-Shaw, K. D., Ledermann, J. P., Borland, E. M., Stovall, J. L., Mossel, E. C., Singh, A. J., Wilusz, J., & Powers, A. M. (2013). O'nyong nyong Virus Molecular Determinants of Unique Vector Specificity Reside in Non-Structural Protein 3. *PLoS Neglected Tropical Diseases*, 7(1), e1931.
- Schilte, C., Staikovsky, F., Couderc, T., Madec, Y., Carpentier, F.,

- Kassab, S., Albert, M. L., Lecuit, M., & Michault, A. (2013). Chikungunya Virus-associated Long-term Arthralgia: A 36-month Prospective Longitudinal Study. *PLoS Neglected Tropical Diseases*, 7(3).
- Schöne, T., Grimm, L. L., Sakai, N., Zhang, L., Hilgenfeld, R., & Peters, T. (2017). STD-NMR experiments identify a structural motif with novel second-site activity against West Nile virus NS2B-NS3 protease. *Antiviral Research*, 146, 174–183.
- Schrodinger, L. L. C. (2015). *The PyMOL Molecular Graphics System, Version 1.8*.
- Schuffenecker, I., Itman, I., Michault, A., Murri, S., Frangeul, L., Vaney, M. C., Lavenir, R., Pardigon, N., Reynes, J. M., Pettinelli, F., Biscornet, L., Diancourt, L., Michel, S., Duquerroy, S., Guigon, G., Frenkiel, M. P., Bréhin, A. C., Cubito, N., Desprès, P., Kunst, F., Rey, F. A., Zeller, H., & Brisse, S. (2006). Genome microevolution of chikungunya viruses causing the Indian Ocean outbreak. *PLoS Medicine*, 3(7), 1058–1070.
- Screen, H., Mueller, N. H., Pattabiraman, N., Ansarah-sobrinho, C., Viswanathan, P., Pierson, T. C., & Padmanabhan, R. (2008). Identification and Biochemical Characterization of Small-Molecule Inhibitors of West Nile Virus Serine Protease by a. *Antimicrobial Agents and Chemotherapy*, 52(9), 3385–3393.
- Shin, G., Yost, S. A., Miller, M. T., Elrod, E. J., Grakoui, A., & Marcotrigiano, J. (2012). Structural and functional insights into alphavirus polyprotein processing and pathogenesis. *Proceedings*

- of the National Academy of Sciences*, 109(41), 16534–16539.
- Shuker, S. B., Hajduk, P. J., Meadows, R. P., & Fesik, S. W. (1996).
Discovering High-Affinity Ligands for Proteins: SAR by NMR.
Science, 274(5292), 1531–1534.
- Shuster, D. L., Menon, R. M., Ding, B., Khatri, A., Li, H., Cohen, E.,
Jewett, M., Cohen, D. E., & Zha, J. (2019). Effects of chronic kidney
disease stage 4, end-stage renal disease, or dialysis on the plasma
concentrations of ombitasvir, paritaprevir, ritonavir, and dasabuvir in
patients with chronic HCV infection: pharmacokinetic analysis of the
phase 3 RUBY-I and R. *European Journal of Clinical
Pharmacology*, 75(2), 207–216.
- Silvestre, H. L., Blundell, T. L., Abell, C., & Ciulli, A. (2013). Integrated
biophysical approach to fragment screening and validation for
fragment-based lead discovery. *Proceedings of the National
Academy of Sciences*, 110(32), 12984–12989.
- Simizu, B., Yamamoto, K., Hashimoto, K., & Ogata, T. (1984). Structural
proteins of Chikungunya virus. *Journal of Virology*, 51(1), 254–258.
- Simon, F., Parola, P., Grandadam, M., Fourcade, S., Oliver, M.,
Brouqui, P., Hance, P., Kraemer, P., Ali Mohamed, A., de
Lamballerie, X., Charrel, R., & Tolou, H. (2007). Chikungunya
infection: an emerging rheumatism among travelers returned from
Indian Ocean islands. Report of 47 cases. *Medicine (Baltimore)*,
86(3), 123–137.
- Singh, H., Mudgal, R., Narwal, M., Kaur, R., Singh, V. A., Malik, A.,
Chaudhary, M., & Tomar, S. (2018). Chikungunya virus inhibition by

- peptidomimetic inhibitors targeting virus-specific cysteine protease. *Biochimie*, 149, 51–61.
- Singh, K. D., Kirubakaran, P., Nagarajan, S., Sakkiyah, S., Muthusamy, K., Velmurgan, D., & Jeyakanthan, J. (2012). Homology modeling, molecular dynamics, e-pharmacophore mapping and docking study of Chikungunya virus nsP2 protease. *Journal of Molecular Modeling*, 18(1), 39–51.
- Singleton, M. R., Dillingham, M. S., & Wigley, D. B. (2007). Structure and Mechanism of Helicases and Nucleic Acid Translocases. *Annual Review of Biochemistry*, 76(1), 23–50.
- Solignat, M., Gay, B., Higgs, S., Briant, L., & Devaux, C. (2009). Replication cycle of chikungunya: A re-emerging arbovirus. *Virology*, 393(2), 183–197.
- Souers, A. J., Levenson, J. D., Boghaert, E. R., Ackler, S. L., Catron, N. D., Chen, J., Dayton, B. D., Ding, H., Enschede, S. H., Fairbrother, W. J., Huang, D. C. S., Hymowitz, S. G., Jin, S., Khaw, S. L., Kovar, P. J., Lam, L. T., Lee, J., Maecker, H. L., Marsh, K. C., Mason, K. D., Mitten, M. J., Nimmer, P. M., Oleksijew, A., Park, C. H., Park, C.-M., Phillips, D. C., Roberts, A. W., Sampath, D., Seymour, J. F., Smith, M. L., Sullivan, G. M., Tahir, S. K., Tse, C., Wendt, M. D., Xiao, Y., Xue, J. C., Zhang, H., Humerickhouse, R. A., Rosenberg, S. H., & Elmore, S. W. (2013). ABT-199, a potent and selective BCL-2 inhibitor, achieves antitumor activity while sparing platelets. *Nature Medicine*, 19(2), 202–208.
- Sourisseau, M., Schilte, C., Casartelli, N., Trouillet, C., Guivel-

- Benhassine, F., Rudnicka, D., Sol-Foulon, N., Le Roux, K., Prevost, M. C., Fsihi, H., Frenkiel, M. P., Blanchet, F., Afonso, P. V., Ceccaldi, P. E., Ozden, S., Gessain, A., Schuffenecker, I., Verhasselt, B., Zamborlini, A., Saib, A., Rey, F. A., Arenzana-Seisdedos, F., Despres, P., Michault, A., Albert, M. L., & Schwartz, O. (2007). Characterization of reemerging chikungunya virus. *PLoS Pathog*, 3(6), e89.
- Spuul, P., Salonen, A., Merits, A., Jokitalo, E., Kaariainen, L., & Ahola, T. (2007). Role of the amphipathic peptide of Semliki forest virus replicase protein nsP1 in membrane association and virus replication. *J Virol*, 81(2), 872–883.
- Storoni, L. C., McCoy, A. J., & Read, R. J. (2004). Likelihood-enhanced fast rotation functions. *Acta Crystallographica Section D Biological Crystallography*, 60(3), 432–438.
- Strauss, J. H., & Strauss, E. G. (1994). The alphaviruses: gene expression, replication, and evolution. *Microbiol Rev*, 58(3), 491–562.
- Suhrbier, A., Jaffar-Bandjee, M. C., & Gasque, P. (2012). Arthritogenic alphaviruses-an overview. *Nature Reviews Rheumatology*, 8(7), 420–429.
- Svergun, D. I. (1992). Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *Journal of Applied Crystallography*, 25(4), 495–503.
- Svergun, D. I. (1994). Solution scattering from biopolymers: advanced contrast-variation data analysis. *Acta Crystallographica Section A*

- Foundations of Crystallography*, 50(3), 391–402.
- Svergun, D. I., Petoukhov, M. V., & Koch, M. H. J. (2001). Determination of Domain Structure of Proteins from X-Ray Solution Scattering. *Biophysical Journal*, 80(6), 2946–2953.
- Taliani, M., Bianchi, E., Narjes, F., Fossatelli, M., Urbani, A., Steinkühler, C., De Francesco, R., & Pessi, A. (1996). A continuous assay of hepatitis C virus protease based on resonance energy transfer depsipeptide substrates. *Analytical Biochemistry*, 240(1), 60–67.
- Tang, G., Peng, L., Baldwin, P. R., Mann, D. S., Jiang, W., Rees, I., & Ludtke, S. J. (2007). EMAN2: An extensible image processing suite for electron microscopy. *Journal of Structural Biology*, 157(1), 38–46.
- Teissier, E., Penin, F., & Pécheur, E. I. (2011). Targeting cell entry of enveloped viruses as an antiviral strategy. *Molecules*, 16(1), 221–250.
- Teo, T.-H., Lum, F.-M., Claser, C., Lulla, V., Lulla, A., Merits, A., Rénia, L., & Ng, L. F. P. (2013). A Pathogenic Role for CD4 + T Cells during Chikungunya Virus Infection in Mice . *The Journal of Immunology*, 190(1), 259–269.
- Teo, T. H., Chan, Y. H., Lee, W. W. L., Lum, F. M., Amrun, S. N., Her, Z., Rajarethinam, R., Merits, A., Röttschke, O., Rénia, L., & Ng, L. F. P. (2017). Fingolimod treatment abrogates chikungunya virus-induced arthralgia. *Science Translational Medicine*, 9(375), 1–12.
- Teo, T. H., Lum, F. M., Lee, W. W. L., & Ng, L. F. P. (2012). Mouse

- models for Chikungunya virus: Deciphering immune mechanisms responsible for disease and pathology. *Immunologic Research*, 53(1–3), 136–147.
- Thiberville, S. D., Moyen, N., Dupuis-Maguiraga, L., Nougairede, A., Gould, E. A., Roques, P., & de Lamballerie, X. (2013). Chikungunya fever: Epidemiology, clinical syndrome, pathogenesis and therapy. *Antiviral Research*, 99(3), 345–370.
- Thomas, S., Rai, J., John, L., Schaefer, S., Pützer, B. M., & Herchenröder, O. (2013). Chikungunya virus capsid protein contains nuclear import and export signals. *Virology Journal*, 10(1), 269.
- Tomar, S., Hardy, R. W., Smith, J. L., & Kuhn, R. J. (2006). Catalytic core of alphavirus nonstructural protein nsP4 possesses terminal adenylyltransferase activity. *Journal of Virology*, 80(20), 9962–9969.
- Tuittila, M., & Hinkkanen, A. E. (2003). Amino acid mutations in the replicase protein nsP3 of Semliki Forest virus cumulatively affect neurovirulence. *The Journal of General Virology*, 84(Pt 6), 1525–1533.
- Tuittila, M., Santagati, M. G., Røytta, M., Määttä, J. A., & Hinkkanen, A. E. (2000). Replicase complex genes of Semliki Forest virus confer lethal neurovirulence. *Journal of Virology*, 74(10), 4579–4589.
- Uhrlaub, J. L., Pulko, V., DeFilippis, V. R., Broeckel, R., Streblow, D. N., Coleman, G. D., Park, B. S., Lindo, J. F., Vickers, I., Anzinger, J. J., & Nikolich-Zugich, J. (2016). Dysregulated TGF- β Production

Underlies the Age-Related Vulnerability to Chikungunya Virus.

PLOS Pathogens, 12(10), e1005891.

Urakova, N., Kuznetsova, V., Crossman, D. K., Sokratian, A., Guthrie, D. B., Kolykhalov, A. A., Lockwood, M. A., Natchus, M. G., Crowley, M. R., Painter, G. R., Frolova, E. I., & Frolov, I. (2017). β -D-N(4)-hydroxycytidine is a potent anti-alphavirus compound that induces high level of mutations in viral genome. *Journal of Virology*.

Varghese, F. S., Kaukinen, P., Gläsker, S., Bepalov, M., Hanski, L., Wennerberg, K., Kümmerer, B. M., & Ahola, T. (2016). Discovery of berberine, abamectin and ivermectin as antivirals against chikungunya and other alphaviruses. *Antiviral Research*, 126, 117–124.

Varghese, F. S., Thaa, B., Amrun, S. N., Simarmata, D., Rausalu, K., Nyman, T. A., Merits, A., McInerney, G. M., Ng, L. F. P., & Ahola, T. (2016). The Antiviral Alkaloid Berberine Reduces Chikungunya Virus-Induced Mitogen-Activated Protein Kinase Signaling. *Journal of Virology*, 90(21), 9743–9757.

Vasiljeva, L., Merits, A., Auvinen, P., & Kaariainen, L. (2000). Identification of a novel function of the alphavirus capping apparatus. RNA 5'-triphosphatase activity of Nsp2. *J Biol Chem*, 275(23), 17281–17287.

Vasiljeva, L., Merits, A., Golubtsov, A., Sizemskaja, V., Kaariainen, L., & Ahola, T. (2003). Regulation of the sequential processing of Semliki Forest virus replicase polyprotein. *J Biol Chem*, 278(43), 41636–41645.

- Vasiljeva, L., Valmu, L., Kääriäinen, L., & Merits, A. (2001). Site-specific protease activity of the carboxyl-terminal domain of Semliki Forest virus replicase protein nsP2. *The Journal of Biological Chemistry*, 276(33), 30786–30793.
- Vazeille, M., Moutailler, S., Pages, F., Jarjaval, F., & Failloux, A. B. (2008). Introduction of *Aedes albopictus* in Gabon: What consequences for dengue and chikungunya transmission? *Tropical Medicine and International Health*, 13(9), 1176–1179.
- Velasco, J. M., Valderama, M. T., Lopez, M. N., Chua, D., Latog, R., Roque, V., Corpuz, J., Klungthong, C., Rodpradit, P., Hussem, K., Poolpanichupatam, Y., Macareo, L., Fernandez, S., & Yoon, I. K. (2015). Chikungunya virus infections among patients with dengue-like illness at a tertiary care hospital in the Philippines, 2012-2013. *American Journal of Tropical Medicine and Hygiene*, 93(6), 1318–1324.
- Vihinen, H., Ahola, T., Tuittila, M., Merits, A., & Kääriäinen, L. (2001). Elimination of Phosphorylation Sites of Semliki Forest Virus Replicase Protein nsP3. *Journal of Biological Chemistry*, 276(8), 5745–5752.
- Voss, J. E., Vaney, M.-C., Duquerroy, S., Vonnrhein, C., Girard-Blanc, C., Crublet, E., Thompson, A., Bricogne, G., & Rey, F. A. (2010). Glycoprotein organization of Chikungunya virus particles revealed by X-ray crystallography. *Nature*, 468(7324), 709–712.
- Wahid, B., Ali, A., Rafique, S., & Idrees, M. (2017). Global expansion of chikungunya virus: mapping the 64-year history. *International*

- Journal of Infectious Diseases*, 58, 69–76.
- Wan, K. F., Wang, S., Brown, C. J., Yu, V. C., Entzeroth, M., Lane, D. P., & Lee, M. A. (2009). Differential scanning fluorimetry as secondary screening platform for small molecule inhibitors of Bcl-XL. *Cell Cycle*, 8(23), 3943–3952.
- Wang, Y. F., Sawicki, S. G., & Sawicki, D. L. (1991). Sindbis virus nsP1 functions in negative-strand RNA synthesis. *Journal of Virology*, 65(2), 985–988.
- Wang, Y., Sawicki, S. G., & Sawicki, D. L. (1994). Alphavirus nsP3 Functions To Form Replication Complexes Transcribing Negative-Strand RNA. *J Virol*, 68(10), 6466–6475.
- Weaver, S. C. (2014). Arrival of Chikungunya Virus in the New World: Prospects for Spread and Impact on Public Health. *PLoS Neglected Tropical Diseases*, 8(6), 6–9.
- Weaver, S. C., & Forrester, N. L. (2015). Chikungunya: Evolutionary history and recent epidemic spread. *Antiviral Research*, 120, 32–39.
- Weaver, S. C., & Lecuit, M. (2015). Chikungunya virus and the global spread of a mosquito-borne disease. *New England Journal of Medicine*, 372(13), 1231–1239.
- Williams, M. C., Woodall, J. P., Corbet, P. S., & Gillett, J. D. (1965). O'nyong-nyong fever: An epidemic virus disease in East Africa VIII. Virus isolations from anopheles mosquitoes. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 59(3), 300–306.
- Yactayo, S., Staples, J. E., Millot, V., Cibrelus, L., & Ramon-Pardo, P. (2016). Epidemiology of Chikungunya in the Americas. *Journal of*

Infectious Diseases, 214(suppl 5), S441–S445.

- Young, A. R., Locke, M. C., Cook, L. E., Hiller, B. E., Zhang, R., Hedberg, M. L., Monte, K. J., Veis, D. J., Diamond, M. S., & Lenschow, D. J. (2019). Dermal and muscle fibroblasts and skeletal myofibers survive chikungunya virus infection and harbor persistent RNA. *PLOS Pathogens*, 15(8), e1007993.
- Zhang, R., Kim, A. S., Fox, J. M., Nair, S., Basore, K., Klimstra, W. B., Rimkunas, R., Fong, R. H., Lin, H., Poddar, S., Crowe, J. E., Doranz, B. J., Fremont, D. H., & Diamond, M. S. (2018). Mxra8 is a receptor for multiple arthritogenic alphaviruses. *Nature*, 557(7706), 570–574.
- Zhao, Y., Ren, J., Harlos, K., Jones, D. M., Zeltina, A., Bowden, T. A., Padilla-Parra, S., Fry, E. E., & Stuart, D. I. (2016). Toremifene interacts with and destabilizes the Ebola virus glycoprotein. *Nature*, 535(7610), 169–172.

Appendix

1. Constructs list

List of constructs for *E. coli* expression

Construct ID	Domains	Expression Vector	Molecular Size	Supplier
CHIKV-nsP2FL (2-795)	Helicase and protease	pNIC28-Bsa4	90 kDa	-
CHIKV-vc038 (2-795)	Helicase and protease	pNIC-CH2	90 kDa	PPP
CHIKV-vc039 (4-470)	Helicase	pNIC28-Bsa4	55 kDa	PPP
CHIKV-vc045 (469-791)	Protease	pNIC28-Bsa4	39 kDa	PPP
ONNV-nsP2FL (1-796)	Helicase and protease	pET28b	90 kDa	-
ONNV-vc002 (471-792)	Protease	pNIC-CH2	39 kDa	PPP
VEEV-nsP2FL (1-794)	Helicase and protease	pET16b	90 kDa	BioBasic

^aPPP; Protein Production Platform, NTU

2. Primers list

List of primers for cloning into expression plasmid

Gene	Forward Primer (5' → 3')	Reverse Primer (5' → 3')
ONNV nsP2 full length (1-796)	CCGCGCGGCAGCcatatgGGCATAGTCGAAAC A	GGTGGTGGTGctcgagTCACGCTCTAGTGGCCT G

Restriction digestion sites denoted in lowercase. NdeI added in forward primer and XhoI added to reverse primer.

List of CHIKV-specific primers for viral load quantification

CHIKV nsP1	
Forward Primer (5' → 3')	GGCAGTATCGTGAATTCGATGCGACACGGAGACGCCAACATT
Reverse Primer (5' → 3')	AATAAATCATAAGTCTGCTCTCTGTCTACATGA
Probe (5' → 3')	^a FAM-TGCTTACACACAGACGT-TAMRA ^b

^aReporter dye (FAM; 6-carboxyfluorescein) labelled nucleotide. ^bQuencher dye (TAMRA; 6-carboxy-tetramethyl-rhodamine) labelled nucleotide.

3. Peptide substrate list

Alphavirus	Peptide sequence	Supplier
CHIKV / ONNV	^a DABCYL-EDRAGA↓GI-EDANS ^b	BioBasic
VEEV	DABCYL-LQEAGA↓GSVETPE-EDANS	Mimotopes

^aReporter dye (DABCYL; 4-{[4-(dimethylamino) phenyl]-azo} benzoic acid) labelled peptide. ^bQuencher dye (EDANS; 5-[(2'-aminoethyl)-amino] naphthalenesulfonic acid) labelled peptide. The arrow indicates the location of the scissile bond, the cleavage site. Cleavage sites between the non-structural proteins contain a common motif, AG(A/C)↓(G/Y/A) (A. Lulla et al., 2006).

4. Sequence alignment of CHIKV, SINV and VEEV nsP2 proteases

```

#
CHIKV 1 QNKANVCWAKSLVPIL#ETAGIKLNDROWSOIQA#FKEDKAYSPEVALNEICTRMYGV#DLD
SINV 1 SCKTNVCWAKALEPILATAGIVLTGCOWSELFPQFVDDKPHSAIYALDVICIKFFGTHLT
VEEV 1 QNKANVCWAKALVPVLKTAGIDMTTEQWNTV-DYFETDKAHSAEIVLNQLCVRFFGLDLD

#
CHIKV 61 SGLFSKPLVSVYYA-----DNHWDNRPGGKMFGE#NPEAASILERKYPFTKGKNINKO
SINV 61 SGLFSKQSIPLTYHPADSARPV#AHWDNSPGTRKYGYDHAVAAELSTFPVFQLA-GKGTQ
VEEV 60 SGLFSAP#TVPLSIR-----NNHWDNSPSPNMYGLNKEVVRQLSRVYQ#LPRAVATGRV

CHIKV 114 ICVTTRRIEDFNPTTNIIPANRRRLPHSLVAEHRPVKGERMEWL#VNKINGHHVLLVSGCSL
SINV 120 LDLOTGRTRVVSQAQCNLVPVNRNLPHALVPEYKEKQPGPIKNFLN#QFHHHSILVVS#ETKI
VEEV 113 YDMNTGTLRN#YDPRINLVPVNRRLPHALVLHHNEHPQSD#FSFVSKLKGRTVLVV-GEKL

CHIKV 174 ALPTKRVTWVAPLGVRCADYTYNLELGLPATLGRYDLV#VINIHTPRIHHYQOCVDHAMK
SINV 180 EVPNKRTEWIAPLGIAGADKSYNLAFGFPQ-ARYDMVF#INIGTKYRNHHFQOCEDHAAT
VEEV 172 SVPGKMVDWLSDRPE--ATFRARLDLCTPGDVPKYDI#IFVNVRTPYKYHHYQOCEDHAMK

CHIKV 234 LQMLGGDSLRLKPGGSLIRAYGYADRTSERVICVLGRKFRSSRALKPPCVTSNTE#MEFF
SINV 239 LKTLRSR#ALNCLNPGGTLVVKSYGYADRNSEDVVTALARKFVRVSAARPECVSSNTE#MYL
VEEV 230 LSM#LTKKACLHLNPGGT#CVSIGYGYADRASESIIGATARQFKFSRVCKPKSSLEETE#VLF

CHIKV 294 LFSNFDNGR-RNFTTHVMNNQLNAA#FVGQATRAGC-
SINV 299 IFRLD#NSRTROFTPH#ELNCVISSVVECTREGVGA-
VEEV 290 VFIGYDRKA-RTHNPYKLSSTLTNIY#TGSR#LHEAGC

```

The amino acid sequences of CHIKV, SINV and VEEV nsP2 proteases were aligned with the ClustalOmega program. Residues in black boxes are strictly conserved and, similar residues are boxed in grey. The two conserved catalytic dyad for cysteine protease activities are indicated by the red #.

5. Sequence alignment of ONNV-vc002 with CHIKV nsP2 protease

Structural alignment of ONNV-vc002 with CHIKV nsP2 protease (PDB: 4ZTB). Residues in red boxes are strictly conserved and similar residues are boxed in yellow. Secondary structural elements from CHIKV protease domain are represented above the alignment. Sequence alignment was prepared using ClustalOmega and addition of structural elements was generated using ESPript 3 (Robert & Gouet, 2014).

6. Sequence alignment of CHIKV, ONNV and VEEV full length nsP2

CHIKV	1	GILETPRGAIKVTAQPTD	HVVGEYLVLSPTVLR	SQKLSLIHALAEQVK	TCTHSGRAGRY
ONNV	1	GIVETPRGAIKVTAQPSD	LVVGEYLVLPQAVLR	SQKLSLIHALAEQVK	TCTHSGRAGRY
VEEV	1	GSVETPRGLIKVTSYAGEDK	IGSYAVLSPOAVLKSEKL	SCIHPLAEQVIVITHSGRKGRY	
CHIKV	61	AVEAYDGRVLVPSGYAIS	PEDFQSLSESATMVYNEREF	VNRKLHHIAMHGPALNTDEESY	
ONNV	61	AVEAYDGRVLVPSGYAIP	QEDFQSLSESATMVYNEREF	VNRKLHHIAMHGPALNTDEESY	
VEEV	61	AVEPYHGRVVPVPEGHAIP	VQDFQALSSESATIVYNEREF	VNRVYLHHIATHGGALNTDEEYV	
CHIKV	121	ELVRAERTEHEYVDVDORR	CKKKEEAAGLVLVGDLTN	PPYHEFAYEGLKIRPACPYKIA	
ONNV	121	ELVRVEKTEHEYVDVDORR	CKKKEEATGLVLVGDLTS	PPYHEFAYEGLKIRPACPYKTA	
VEEV	121	KTVPKPEHDGEYLYDIDRK	QCVKKELVTGLGLTGELVD	PPFHEFAYESLRTRPAPYQVP	
CHIKV	181	VIGVFGVPGSGKS	SAIIKNLVTRQDLVTS	GKKENCQEI	TDVMRQRGLEISARTVDSLLL
ONNV	181	VIGVFGVPGSGKS	SAIIKNLVTRQDLVTS	GKKENCQEI	SDVMRQRKLEISARTVDSLLL
VEEV	181	TIGVYGVPGSGKSG	IIKSAVTKKDLV	VSAKKENC	AEIIRDVKKMKGLDVNARTVDSVLLN
CHIKV	241	GCNRPVDVLYVDEAFACHSGT	LLALIALVRPRQKV	VLCGDPKQCGFF	NMMQMKVNYNHN
ONNV	241	GCNKPVEVLYVDEAFACHSGT	LLALIAMVRPRQKV	VLCGDPKQCGFF	NMMQMKVNYNHN
VEEV	241	GCKHPVETLYVDEAFACHAGT	LRALIALIRPKK	-AVLCGDPKQCGFF	NMMCKLVHFNHEI
CHIKV	301	CTQVYHKSISRRCTLP	VTAIVSSLHYEGKMRTT	NEYNKPPIVVDTTG	SKTPD
ONNV	301	CTQVYHKSISRRCTLP	VTAIVSSLHYEGKMRTT	NEYNKPPIVVDTTG	ITKPE
VEEV	300	CTQVFHKSISRRCTKS	VTSSVSTL	FYDKMRTTNPKETKI	V
CHIKV	361	RGVVKQLQIDYRG	HEVMTAAASQGLTR	KGVYAVRQKVNEN	PLYASTSEHVN
ONNV	361	RGVVKQLQIDYRG	NEVMTAAASQGLTR	KGVYAVRQKVNEN	PLYAPTSEHVN
VEEV	360	RGVVKQLQIDYK	GNE	MTAAASQGLTR	KGVYAVR
CHIKV	421	L	VWKTLSGDPWIKTLQ	NPPKGNFKATIK	EWEVEHASIMAGICS
ONNV	421	L	VWKTLSGDPWIKTLQ	NPPKGD	FKATIK
VEEV	420	L	VWKTLAGDPWIKTL	TAKYPGNFTAT	IEWQAEHDAIMRH
CHIKV	481	K	SLVPIL	ETAGIKLND	RQWSQITQAFKED
ONNV	481	K	SLVPILD	ETAGIKLSD	RQWSQIVQAFKED
VEEV	480	K	ALVPVLK	ETAGIDMTTEQ	WNTV-DYFETDKAHS
CHIKV	541	S	VYYADNHWDNR	PGGKMGFNPEA	ASILERKYPFT
ONNV	541	S	VYYADNHWDNR	PGGKMGFNPEA	LMLEKKYPFT
VEEV	539	P	LSIRNHNWDNS	SPSPNMTGLN	KEVROLSRRYP
CHIKV	601	I	PANRRRLPHSLVAEHR	PKVGERME	VLVNKINGHHVLL
ONNV	601	I	PANRRRLPHSLVAEH	HSVRGERME	VLVNKISGHEML
VEEV	599	L	VPVNRRLPHALVL	HHNEHPQSD	FSSFVSKLKGRT
CHIKV	661	G	ADYTYNLELGLPAT	LGRYDLV	VINIHTPFRIHHY
ONNV	661	G	ADYTYNLELGLPAT	LGRYDLV	VINIHTPFRIHHY
VEEV	656	E	ATFRARLDLGI	PGDVPKYDI	IIFVNVRT
CHIKV	721	S	LLIRAYGYADRTSER	VICVLGRK	FRSSRALKPP
ONNV	721	S	LLIRAYGYADRTSER	VISVLGRK	FRSSRALKPP
VEEV	716	T	CVSIGYGYADRASE	SI	GAIARQFKFS
CHIKV	781	M	NNQLNAA	EVGA	-TRAGC
ONNV	781	M	NNQLNA	VYAGLA	-TRAGC
VEEV	776	L	SSTLTN	YTG	SRLHEAGC

The amino acid sequences of CHIKV, ONNV and VEEV full length nsP2 were aligned with the ClustalOmega program. Residues in black boxes are strictly conserved and, similar residues are boxed in grey. The two conserved catalytic dyad for cysteine protease activities are indicated by the red #.

7. Sequence alignment of SGP11, CNR20235, LR2006-OPY1 and IND91 nsP2 proteases

		#	
SGP11	MAGICSHQMTFDTFQNKANVCWAKSLVPILETAGIKLNDROWSQIIQAFKEDKAYSPEVA	60	
CNR20235	MAGICSHQMTFDTFQNKANVCWAKSLVPILETAGIKLNDROWSQIIQAFKEDKAYSPEVA	60	
LR2006-OPY1	MAGICSHQMTFDTFQNKANVCWAKSLVPILETAGIKLNDROWSQIIQAFKEDKAYSPEVA	60	
IND91	MAGICSHQMTFDTFQNKANVCWAKSLVPILETAGIKLNDROWSQIIQAFKEDKAYSPEVA	60	
	*****		
SGP11	LNEICTRMYGVDLDSGLFSKPSVSVYYADNHWDNRPGGKMF ~~GF~~ FNPEAASILERKYPFTKG	120	
CNR20235	LNEICTRMYGVDLDSGLFSKPLVSVYYADNHWDNRPGGKMF ~~GF~~ FNPEAASILERKYPFTKG	120	
LR2006-OPY1	LNEICTRMYGVDLDSGLFSKPLVSVYYADNHWDNRPGGKMF ~~GF~~ FNPEAASILERKYPFTKG	120	
IND91	LNEICTRMYGVDLDSGLFSKPLVSVYYADNHWDNRPGGKMF ~~GF~~ FNPEAASILERKYPFTKG	120	
	*****		
SGP11	KWNINKQICVTTRRIEDFNPTTNIIPANRRLPHSLVAEHRPVKGERMEWLVNKGHHVL	180	
CNR20235	KWNINKQICVTTRRIEDFNPTTNIIPANRRLPHSLVAEHRPVKGERMEWLVNKGHHVL	180	
LR2006-OPY1	KWNINKQICVTTRRIEDFNPTTNIIPANRRLPHSLVAEHRPVKGERMEWLVNKGHHVL	180	
IND91	KWNINKQICVTTRRIEDFNPTTNIIPANRRLPHSLVAEHRPVKGERMEWLVNKGHHVL	180	
	*****		
SGP11	LVSGYNLALPTKRVTWVAPLGVRGADYTYNLELGLPATLGRYDLVVINIHTPFRIHHYQQ	240	
CNR20235	LVSGYNLALPTKRVTWVAPLGVRGADYTYNLELGLPATLGRYDLVVINIHTPFRIHHYQQ	240	
LR2006-OPY1	LVSGYNLALPTKRVTWVAPLGVRGADYTYNLELGLPATLGRYDLVVINIHTPFRIHHYQQ	240	
IND91	LVSGYNLALPTKRVTWVAPLGVRGADYTYNLELGLPATLGRYDLVVINIHTPFRIHHYQQ	240	
	*****		
SGP11	CVDHAMKLQMLGGDSLRLKPGGSLLIRAYGYADRTSERVICVLGRKFRSSRALKPPCVT	300	
CNR20235	CVDHAMKLQMLGGDSLRLKPGGSLLIRAYGYADRTSERVICVLGRKFRSSRALKPPCVT	300	
LR2006-OPY1	CVDHAMKLQMLGGDSLRLKPGGSLLIRAYGYADRTSERVICVLGRKFRSSRALKPPCVT	300	
IND91	CVDHAMKLQMLGGDSLRLKPGGSLLIRAYGYADRTSERVICVLGRKFRSSRALKPPCVT	300	
	*****		
SGP11	SNTEMFFLFSNFDNGRRNFTTHVMNNQLNAAFVGQVTRAGC	341	
CNR20235	SNTEMFFLFSNFDNGRRNFTTHVMNNQLNAAFVGQVTRAGC	341	
LR2006-OPY1	SNTEMFFLFSNFDNGRRNFTTHVMNNQLNAAFVGQVTRAGC	341	
IND91	SNTEMFFLFSNFDNGRRNFTTHVMNNQLNAAFVGQVTRAGC	341	
	*****		

The amino acid sequences of CHIKV, ONNV and VEEV full length nsP2 were aligned with the ClustalOmega program. The two conserved catalytic dyad for cysteine protease activities are indicated by the red #.

8. Publication

viruses

Review

Nonstructural Proteins of Alphavirus—Potential Targets for Drug Development

Farhana Abu Bakar ^{1,2} and Lisa F. P. Ng ^{1,3,4,*}

¹ Singapore Immunology Network, Agency for Science, Technology and Research (A*STAR), Singapore 138648, Singapore; farhana_abubakar@immunol.a-star.edu.sg

² School of Biological Sciences, Nanyang Technological University, Singapore 636921, Singapore

³ Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore, Singapore 117596, Singapore

⁴ Institute of Infection and Global Health, University of Liverpool, Liverpool L69 7BE, UK

* Correspondence: lisa_ng@immunol.a-star.edu.sg; Tel.: +65-6407-0028

Received: 7 January 2018; Accepted: 6 February 2018; Published: 9 February 2018

Abstract: Alphaviruses are enveloped, positive single-stranded RNA viruses, typically transmitted by arthropods. They often cause arthralgia or encephalitic diseases in infected humans and there is currently no targeted antiviral treatment available. The re-emergence of alphaviruses in Asia, Europe, and the Americas over the last decade, including chikungunya and o'nyong'nyong viruses, have intensified the search for selective inhibitors. In this review, we highlight key molecular determinants within the alphavirus replication complex that have been identified as viral targets, focusing on their structure and functionality in viral dissemination. We also summarize recent structural data of these viral targets and discuss how these could serve as templates to facilitate structure-based drug design and development of small molecule inhibitors.

Keywords: alphavirus replication; nonstructural protein; protease inhibitors

1. Introduction

Alphaviruses belong to the *Togaviridae* family [1]. They are arboviruses that are transmitted to humans through the mosquito species *Aedes aegypti* and *Aedes albopictus*, where they cause various diseases and can be broadly divided into arthritogenic [2] and encephalitogenic [3]. Old World alphaviruses such as chikungunya virus (CHIKV), o'nyong'nyong virus (ONNV), and sindbis virus (SINV) are arthritogenic and commonly cause febrile illness accompanied by rash, polyarthralgia and chronic arthritis [4–8]. Infection with New World alphaviruses such as Eastern equine encephalitis (EEEV) and Venezuelan equine encephalitis virus (VEEV), which are less prevalent in humans, are mostly associated with neurological disease [9,10]. In 2005, the re-emergence of CHIKV in several Indian Ocean islands led to massive outbreaks, affecting one-third of the inhabitants in La-Réunion alone [11]. To date, millions of people have been infected in more than 40 countries in Asia, Europe and the Americas [12]. Despite the medical threat posed by CHIKV, there is currently no approved antiviral treatment or vaccine for CHIKV infection, as well as for the other less prevalent alphaviruses. Treatments are usually symptomatic, with administration of non-steroidal anti-inflammatory drugs and analgesic agents to control fever and severe joint pains.

Molecular Virology and Genome Organization

Alphaviruses are small, enveloped RNA viruses with a single-stranded, positive-sense RNA genome. The alphavirus genome is approximately 12 kb long and it consists of two open reading frames (ORFs): a 7 kb frame encoding the non-structural polyprotein and a 4 kb frame encoding the

structural polyprotein (Figure 1A) [1,13]. The non-structural polyprotein is cleaved into four different proteins (nsP1, nsP2, nsP3, and nsP4) which are necessary for the transcription and translation of viral mRNA inside the cytoplasm of host cells. Figure 1B describes the stages of nonstructural polyprotein processing. Translation of the viral RNA by host cell translational machinery produces two non-structural protein (nsP) precursors (P123 or P1234). P1234 is expressed as a read-through of an opal termination codon at the end of nsP3 (Figure 1A) [14]. These precursor polyproteins are cleaved by a carboxyl-terminal protease domain of nsP2 [15]. Following translation of P1234, cleavage at the P3/4 junction occurs either in *cis* or *trans*, followed by the P1/2 junction which occurs in *cis* only [16]. Both P123+nsP4 and nsP1+P23+nsP4, together with some cellular proteins, form the early replication complex (RC), which preferentially synthesizes negative strand viral RNA [17,18]. The final cleavage event at the P2/3 junction produces fully mature nsPs, which along with host cell proteins, forms the positive strand RC, switching RNA template for synthesis of positive-sense genomic (49S) and subgenomic (26S) RNAs. However, the correlation between P23 cleavage and the switch from negative- to positive-sense RNA production remains poorly understood.

Figure 1. (A) Schematic representation of the alphavirus genome showing the RNA sequence open reading frames (ORFs). The (*) indicates the position of opal termination codon; (B) schematic representation of non-structural polyprotein (nsP2) processing by nsP2 protease. Early processing of P1234 produces P123 and nsP4 which associate to form the early replication complex (RC), which performs negative-sense RNA synthesis. P123 is further processed to produce the individual nsPs, which associate to form mature RC that regulates positive-sense RNA synthesis and transcription of subgenomic 26S RNA.

The translation of 26S subgenomic positive sense RNA generates a single structural polyprotein, which is cleaved into five structural proteins: the Capsid (C), two major envelope glycoproteins E1 and E2, and two small cleavage products (E3, 6K) (Figure 1A). While the C protein is being autocatalytically cleaved off to encapsidate new positive sense RNA molecules, the envelope polyprotein precursor E3-E2-6k-E1 is translocated to the endoplasmic reticulum (ER). Host signalases process the polyprotein at the N- and C-terminal end of the 6k peptide, resulting in E3E2, 6k, and E1, where all anchored to the ER membrane. During export to the plasma membrane, the E3E2 precursor is cleaved by furin-like protease activity in the trans-Golgi system into E2 and E3 [19]. The nucleocapsid forms with the assembly of 120 dimers of the C protein, which buds at the cell membrane as spherical particles, acquiring a lipid envelope with embedded E1 and E2 glycoproteins [20,21]. Viral particles

exhibit 80 trimeric spikes composed of heterodimers of E1 and E2, with E2 glycoprotein facilitating binding of the virus to cell surface receptors [22–25]. Upon receptor binding, the virus particle enters the host cell via clathrin-dependent endocytosis [26]. The acidic environment of the endosome results in the E1-mediated fusion of the viral envelope and endosomal membrane, followed by the release of nucleocapsid and uncoating of the viral genome in the cytosol [27,28].

In this review, the organization of alphavirus RC focusing on the function and structure of nsP2 protease will be illustrated. These proteins play important roles in the various replication stages of the viral genome. Notably, the conserved architecture of the nsP2 protease across the New and Old World alphaviruses, as determined by the recently solved structures of VEEV, SINV and CHIKV nsP2 proteases, also makes it an ideal target for designing specific and pan-alphavirus protease inhibitors [29–32].

2. Roles and Function of Non-Structural Proteins

2.1. Non-Structural Proteins (nsPs)

Key advances have been made to understand the biological aspects and pathogenesis of alphaviruses, using mainly Semliki Forest virus (SFV) and SINV as prototypes [33,34]. Many of the functions of the nsPs have been characterized on the basis of sequence comparisons and biochemical assays (Table 1), and evidence suggests their intrinsic interactions are essential for the formation of functional RCs [18,35–37].

2.1.1. Non-Structural Protein 1 (nsP1)

The nsP1 protein is an mRNA capping enzyme that possesses both guanine-7-methyltransferase (MTase) and guanylyltransferase (GTase) activities, where they direct the methylation and capping of newly synthesized viral genomic and subgenomic RNAs [38–41]. The MTase motif in the N-terminal domain of nsP1 catalyzes the transfer of the methyl group from S-adenosylmethionine (AdoMet) to the N7 position of a GTP molecule (m^7Gppp). GTase then binds the m^7Gppp , forming a covalent link with a catalytic histidine (m^7Gp -GTase) and releasing PP_i . The GTase then transfers the m^7Gp molecule to the 5'-diphosphate RNA to create $m^7GpppNp$ -RNA [42]. The resulting cap structure is essential for viral mRNA translation and prevents the mRNA from being degraded by cellular 5' exonucleases. Following the N-terminal domain are features that allow the association of the nsP1 protein to cellular membranes. The presence of α -helical amphipathic loop and palmitoylation sites allow the nsP1 protein and nsP1-containing RC to anchor onto the plasma membrane, possibly through nsP1 interaction with the membrane's anionic phospholipids [43–47].

2.1.2. Non-Structural Protein 2 (nsP2)

The nsP2 protein possesses numerous enzymatic activities and functional roles. The N-terminal region contains a helicase domain that has seven signature motif of Superfamily 1 (SF1) helicases [48]. It functions as an RNA triphosphatase that performs the first of the viral RNA capping reactions [49,50]. It also functions as a nucleotide triphosphatase (NTPase), fueling the RNA helicase activity [50,51]. The C-terminal region of nsP2 contains a papain-like cysteine protease, which is responsible for processing the viral non-structural polyprotein (Figure 2A) [52–54]. The protease recognizes conserved motifs within the polyprotein (Figure 2B) [55]. This proteolytic function is highly regulated and is modulated by other domains of nsP2 [16]. The crystal structure has revealed two distinct domains. The N-terminal subdomain has a α/β -fold that is novel, unlike the structure of other known cysteine proteases. The C-terminal subdomain is an S-adenosyl-L-methionine-dependent (SAM) RNA methyltransferase domain with a classical MTase fold, but enzymatically non-functional [29,30]. Recent crystal structure of VEEV's enzyme surface has revealed that the predicted active site is in a major surface groove, which is likely to accommodate the substrate polyprotein to be cleaved [29,30]. The major enzyme groove may act as an enzyme mouth holding the protein to

be processed. The alphavirus nsP2 protein has also been described as a virulence factor responsible for the transcriptional and translational shutoff in infected host cells and the inhibition of interferon (IFN) mediated antiviral responses contributing to the controlling of translational machinery by viral factors [56–59].

Figure 2. (A) Illustration of superposed structures of papain and chikungunya virus (CHIKV) nsP2 proteases (Protein Data Bank (PDB) 9PAD and 3TRK respectively). The structure of papain protease is presented as a solid, blue ribbon. The structure of CHIKV nsP2 protease is presented as solid, green ribbon. The domain common for both proteases is highlighted in dark colors and is enlarged subsequently to show the conserved catalytic dyad; (B) alignment of nsP2 cleavage sites. The nomenclature of Berger and Schechter is used to identify residues on the amino (P1, P2, etc.) or carboxy (P1', P2', etc.) termini of the scissile bond. The arrow indicates the location of the cleavage site. Cleavage sites between the non-structural proteins contain a common motif, AG(A/C)↓(G/Y/A) [55]. Abbreviations: Chikungunya virus (CHIKV), o'nyong'nyong virus (ONNV), semliki forest virus (SFV), ross river virus (RRV), sindbis virus (SINV) and Venezuelan equine encephalitis virus (VEEV).

2.1.3. Non-Structural Protein 3 (nsP3)

The precise role(s) of alphavirus nsP3 protein in the RC is less clear. The nsP3 protein has three recognized domains: the N-terminal macrodomain with phosphatase activity and nucleic acid binding ability, the alphavirus unique domain (AUD) and the C-terminal hypervariable domain [60]. The crystal structures of the macrodomain from CHIKV and VEEV were found to closely resemble the homologous *Escherichia coli* domain [60]. To date, the most well-defined structural information available are the protease region of nsP2 and the folded N-terminal region of nsP3 (Table 1). It has been solved to 2.85 Å resolution and has a zinc coordination site within the AUD [32]. The hypervariable domain has sequence features of natively unfolded proteins. It has been demonstrated that the deletion

of this domain in SFV nsP3 resulted in low viral pathogenicity, suggesting its importance in viral RNA transcription regulation [61].

Table 1. Non-structural protein (nsP) domain functions and available crystal structures.

Non-Structural Protein	Domain Function	Virus	PDB ID	Reference
nsP1	mRNA capping	SFV	1FW5	[45]
nsP2	NTPase/HelicaseProtease	-	-	-
		VEEV	2HWK, 5EZQ	[29,30]
		CHIKV	3TRK	[31]
nsP3	Macrodomain	SINV	4GUA	[32]
		VEEV	3GQE	[60]
		CHIKV	3GPG	[60]
nsP4	RNA-dependent RNA Polymerase	SINV	4GUA	[32]
		-	-	[32]

PDB: Protein data bank; SFV: Semliki Forest virus; VEEV: Venezuelan equine encephalitis virus; CHIKV: chikungunya virus; SINV: sindbis virus.

2.1.4. Non-Structural Protein 4 (nsP4)

The nsP4 polymerase is the most highly conserved protein in alphaviruses, with the most divergent being >50% identical in amino acid sequence when compared with other alphaviral nsP4s [62,63]. The nsP4 contains the core RNA-dependent RNA polymerase (RdRp) domain at the C-terminal end, determined to be solely responsible for the RNA synthetic properties of the viral RC [64]. The RdRp participates in replicating the genomic RNA via a negative strand RNA and transcribing the 26S subgenomic RNA. The N-terminal domain is alphavirus-specific and may be partially disordered structurally. It appears to be important for the interaction with polyprotein P123 to form RCs that are capable of synthesizing minus strands from plus-strand templates [64–66]. Deletion and mutation studies of the RdRp domain of SINV nsP4 demonstrated terminal adenylyltransferase (TATase) activity, suggesting it has a potential role in maintenance of the 3' poly-A tail at the 3'-end of positive-sense RNAs [66]. Comparison of the secondary structure of CHIKV RdRp with the structures of picornaviral polymerases showed a classical, basic RdRp architecture with well-defined fingers, palm containing the GDD active site and thumb domains [67]. A comprehensive review of the nsP4 was recently published that illustrated the fundamental functions of nsP4. It detailed the importance of P123 interaction with nsP4, the importance of RCs' association with the cellular membrane and the possible interactions of RCs with host factors during viral replication [68].

2.2. Viral Target Proteins for Drug Development

Alphavirus replication is a delicate process, requiring specific protein-protein interactions among the nsPs and host factors to effectively form the highly organized RCs at early and late stages of infection. During infection, the formation of cytoplasmic vacuoles is induced in host cells. These vacuoles contain small cell membrane invaginations called spherules where the RC proteins nsP1 to nsP4, host factors as well as newly synthesized viral RNA localize [69–73]. These structures serve as compartments to facilitate virus propagation, by allowing spatial separation and regulation of RNA translation, replication and packaging of the viral genome. They protect viral RCs and genomic RNA from degradation by cellular proteases and prevent recognition by antiviral double-stranded RNA sensors, such as RIG-I and MDA-5 [74–77]. Although no 3-D reconstruction of alphavirus replication compartments has been published yet, there are multiple reports revealing the morphologies through electron tomography of other positive sense RNA viruses known to form similar replication sites [78–81]. Alphavirus RCs are widely accepted to reside on the invaginated cell membrane, with RNA replication taking place in the spherule lumen [70,82,83]. In recent years, attempts have

been made to shed light into the late RC's organization, with a number of studies adopting systems such as yeast two-hybrid [84,85], immunoprecipitation [86–89] and ELISA [85] to map the interactions among the nsPs. Six novel interactions are identified in CHIKV (nsP1-nsP1, nsP1-nsP2, nsP1-nsP3, nsP1-nsP4, nsP2-nsP4 and nsP4-nsP4) [84], some of which are similarly shown in SINV [88–90] and SFV [86,87]. It has also been demonstrated that nsP1 is involved in the recruitment of other nsPs into the spherules [87] and its membrane association is crucial for SFV replication [46]. These data suggest that nsP1's interaction with all other nsPs is absolutely essential to keep the RC intact and functional. Therefore, nsP1 is an attractive target for drug development. Firstly, perturbing its affinity for the cell membrane could potentially inhibit it from anchoring in the spherules. This will eventually stop the recruitment of the other nsPs, thus preventing the initiation of RC formation. Secondly, disrupting the intrinsic interactions among the nsPs could prevent proper conformational arrangements of the RCs and this could ultimately impede viral replication. However, there is an information deficit in the understanding of nsP interactions and without this knowledge, targeted inhibition on RCs remains difficult. Future work will still need to address the relative importance of the arrangements between the nsPs in the RCs.

nsP2 protein is another excellent viral target for inhibition due to its role in viral replication and host evasion strategies [91]. nsP2 is multifunctional: it has RNA helicase, RNA triphosphatase, nucleoside triphosphatase and auto-protease activities. It is often described as an important co-factor for the maturation of viral RC [16,37,92]. The nsP2 protease function is especially of interest, as proteases of other viruses (such as Human Immunodeficiency Virus (HIV) and Hepatitis C Virus (HCV)) have been successfully targeted, leading to the development of FDA-approved inhibitors [93–95]. Although the sequence identity of the different alphavirus nsP2 proteases is very low, their active site residues are conserved [29]. Cys478 and His548 of CHIKV nsP2 protease (referred as Cys1013 and His1083 in [31]) are two residues that form the catalytic dyad (Figure 2A), where substrates with defined recognition sequences are cleaved (Figure 2B) [1,31,96–98]. It has been demonstrated that a Cys478 to Ala mutation produces inactive protease and completely abolishes CHIKV replication [98]. Recently, the first structure of a peptide-like E64d inhibitor-bound VEEV nsP2 protease (PDB entry 5EZX) was reported (Table 2) [30]. Although E64d is not a viable therapeutic candidate as it could only inhibit protease function and was ineffective in inhibiting viral replication, the structure provided invaluable insights into the roles of the catalytic residues and possible orientation of the substrates during catalysis. The inhibitor was shown to bind beneath a β -hairpin in the interface of the protease and SAM MTase domains. This further demonstrated that the SAM MTase domain is required for proteolysis, with at least three of its residues (Arg662, Lys705 and Lys706) being used by the cysteine protease for substrate binding and recognition. His546 (referred as His548 in CHIKV nsP2) in the protease domain is the only residue adopting a different conformation, with minimal overall structural changes observed when compared to the free enzyme. It was also suggested that the interaction of the carbonyl oxygen of the ester on E64d with the NH of Cys477 (referred as His478 in CHIKV nsP2) stabilized the transition state of the proteolytic reaction. These observations are crucial information as they emphasize the importance of residues within and around the active site cleft. Therefore, targeting these residues would be an applicable strategy to inhibit the enzyme function, which could consequentially inhibit viral replication. Moreover, nsP2 is also identified as a virulence factor. nsP2 plays a role in shutting off host cell mRNA transcription and/or translation [56,59]. A portion of the nsP2 is localized to the nucleus [99,100] and it inhibits host antiviral response by suppressing type I/II interferon-stimulated JAK/STAT signaling [101,102]. A recent study showed that mutations in the nuclear localization signal (NLS) of SFV nsP2 (⁶⁴⁹DDR⁶⁵¹ and ⁶⁴⁹RDD⁶⁵¹) completely blocked nsP2 from entering the nucleus and reduced SFV-induced cell death [103]. This was likely due to retention of nsP2 in the cytoplasm that prevented its association with host factors in the nucleus to shut off host antiviral functions. The NLS of nsP2 could be a good inhibitory target for preventing nsP2 translocation into the nucleus. However, this will only be effective against specific alphaviruses

that use nsP2 to inhibit cellular mRNA transcription as it has been demonstrated otherwise for VEEV nsP2 [59].

3. Chemotherapeutics Targeting Viral Non-Structural Proteins

3.1. Antivirals against nsP2 Protease

The reemergence of CHIKV and its subsequent epidemics exemplify the threat alphaviruses currently have on human health. A comprehensive review of the antiviral efforts towards alphaviruses was recently published, where the various types of inhibitors discovered and developed as potential therapeutics were described [104]. It is of prime importance that resolved structures are employed in facilitating the discovery of novel drug compounds. The nsP2 protease's function as a nonstructural polyprotein processor is absolutely essential for virus replication. Thus, inhibiting this function could potentially mitigate CHIKV infection. This strategy has successfully been used against several viruses like HCV and HIV and has led to the development of remarkably potent drugs [95,105–107]. With the recent availability of various alphavirus nsP2 protease crystal structures, they have become the most targeted viral domain for in silico drug design to date. Furthermore, crystal structures of the VEEV, CHIKV and SINV nsP2 protease domain Protein Data Bank (PDB) entries 2HWK [29], 5EZQ [30], 3TRK [31] and 4GUA [32] respectively displayed highly conserved folds despite having low sequence identity, rendering it a potential target both for specific and pan-alphavirus protease inhibition. Bassetto et al. have effectively employed homology modeling and computer-aided drug design strategies to discover the first few potential compounds against CHIKV nsP2 protease, with the support of a combination of cell-based virus inhibition and cell-free protease assays [108]. Through the virtual screening of about five million commercially available compounds on the VEEV nsP2 protease structure, several were shown to selectively inhibit CHIKV-induced cell death. Compound 25, in particular, displayed the best antiviral activity ($EC_{50} = 3.2 \mu M$) and is predicted to target the central portion of the nsP2 protease active site (Table 2) [108]. Continuing from one of the other lead compounds by Bassetto et al., Das et al. described a set of related compounds, with a few capable of inhibiting CHIKV nsP2 protease as well as virus replication [109]. These results further prove that these compounds derived from in silico drug design are indeed nsP2 protease inhibitors. Many more potential inhibitors originating from the structure-activity relationship and molecular dynamics simulation strategies have been reported and illustrated to dock in the protease's active site with good affinity [110–112]. However, the biological validation of their antiviral activity and/or target specificity is lacking. Aside from screening commercially available compound libraries, repurposing of FDA-approved protease inhibitors is another strategy that could accelerate the identification and development of specific and potent CHIKV protease inhibitors. Lopinavir and Nelfinavir, both potent HIV protease inhibitors, were validated to have modest antiviral activity on CHIKV (Table 2) ($EC_{50} = 32 \mu M$ and $14 \mu M$ respectively). However, they are clearly toxic on the host cell as they display poor selectivity index values ($CC_{50} = 44 \mu M$ and $22 \mu M$ respectively) [113]. Nonetheless, these inhibitors could serve as a stepping stone for the development of novel alphavirus protease inhibitors.

Table 2. List of compounds shown to inhibit nsP2 protease activity.

Molecule	Structure	Virus	EC ₅₀	Reference
Compound 25		CHIKV	3.2 μ M	[108]
Compound 1c		CHIKV	27 μ M	[109]
Nelfinavir		CHIKV	14 μ M	[113]
E64d—cysteine protease inhibitor		VEEV	ND	[30]

ND: not determined.

3.2. Inhibitors of Other nsPs

Inhibitors targeting other nsPs have also been reported, where their mechanisms of action on alphavirus replication were evaluated through reverse genetic mutation of drug-resistant mutants. A number of highly selective CHIKV and VEEV nsP1 capping enzyme inhibitors postulated to disrupt the GTase activity of nsP1 were reported recently [114–116]. MADTP-372 was demonstrated to be potent and selective in inhibiting the induction of cytopathic effect by CHIKV and VEEV [116]. Due to the close proximity of Pro34 residue (mutated in MADTP-372-resistant variant) to His37 residue (a putative acceptor for the guanylation), it was postulated that MADTP-372 inhibited the GTase reaction by either blocking the binding of m⁷Gppp to nsP1 or the guanylation process itself. However, due to the absence of structural information of alphavirus nsP1, other possible scenarios that involve indirect effects cannot be ruled out. Nucleoside analogs have also been proven to be effective against several alphaviruses. Notably, Ribavirin (a guanosine analog widely used against a number of other RNA viruses) demonstrated effective inhibition of CHIKV and SFV genome replication by depleting GTP pools [117–119]. The resulting GTP deficiency in host cells could prevent proper capping of newly synthesized viral mRNA by nsP1, and consequently allow cellular 5' exonucleases to degrade the uncapped mRNA. The absence of the cap structure could also impede viral mRNA translation. In addition, Ribavirin has been proposed to directly inhibit nsP4 RdRp through its interaction with Cys483 residue, resulting in an increase in replication fidelity [120,121]. Another modified nucleoside analog, β -D-N⁴-hydroxycytidine (NHC), was effective in reducing the rates of VEEV release and its infectivity [122]. NHC is by far the most potent nsP4 inhibitor ever reported (EC₅₀ = 0.426 μ M, CC₅₀ > 200 μ M) and could potentially be a substitute for Ribavirin as it is inefficient in developing NHC resistant mutants. Another well-studied nsP4 inhibitor, Favipiravir (T-705; a purine analog approved in Japan for the treatment of influenza infections), was shown to

exert broad-spectrum anti-alphavirus activity in vitro and provide protection in a mouse model of lethal CHIKV infection [123]. Favipiravir selectively inhibits CHIKV nsP4 RdRp function through its interaction with Lys291 residue. Interestingly, this particular lysine residue is conserved in the polymerases of other positive-sense RNA viruses. Thus this may provide an explanation for Favipiravir's broad-spectrum antiviral activity. Through the screening of chemical compound libraries, Compound-A was also recently found to inhibit CHIKV infection. Compound-A was demonstrated as a specific nsP4 RdRp inhibitor that could potentially inhibit RdRp's ribonucleotide selection function by targeting Met2295 residue [124]. It demonstrated potent CHIKV antiviral activity but exhibited very poor selectivity ($EC_{50} = 1.29 \mu M$, $CC_{50} = 5.84 \mu M$). Nonetheless, Compound-A could serve as a starting point for chemical modifications so as to reduce its toxicity.

4. Concluding Remarks and Perspectives

The current understanding of interactions through structural evaluation of the nsP2 protease domain has opened avenues for the development of specific inhibitors. It is by far the most well-defined out of the other nsP domains and its structural information has been inspirational for computational biologists and chemists alike in designing many series of compounds, mainly targeting its catalytic binding pocket. The representation of E64d inhibitor bound to the nsP2 catalytic pocket is a major advancement towards identifying the key residues that may be important for substrate binding and recognition [30]. However, the lack of structural information on the other viral nsPs makes understanding of the mode of action of inhibitors and rational designing of specific and efficient inhibitors challenging. It is worthwhile to note that for all nsP1 and nsP4 inhibitors discussed earlier, the direct interactions with their targeted residues were never demonstrated. In particular, obtaining the crystal structure of alphavirus nsP4 RdRp would allow more robust comparison of this polymerase to other viral RdRps and new relationships between RdRp-encoding viruses could be formed. Resolving the structure of the RC will also allow for a better understanding of the function of the polymerase and the entire RC.

Acknowledgments: This work is supported by core research grants provided to the Singapore Immunology Network by the Biomedical Research Council (BMRC), Agency for Science, Technology and Research (A*STAR), Singapore. We would also like to thank Kai-Er Eng from SgN, A*STAR for critical discussion and proofreading of the manuscript.

Author Contributions: Farhana Abu Bakar and Lisa F. P. Ng wrote the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Strauss, J.H.; Strauss, E.G. The alphaviruses: Gene expression, replication, and evolution. *Microbiol. Rev.* **1994**, *58*, 491–562. [\[PubMed\]](#)
2. Laine, M.; Luukkainen, R.; Toivanen, A. Sindbis viruses and other alphaviruses as cause of human arthritic disease. *J. Intern. Med.* **2004**, *256*, 457–471. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Steele, K.E.; Twenhafel, N.A. Review paper: Pathology of animal models of alphavirus encephalitis. *Vet. Pathol.* **2010**, *47*, 790–805. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Borgherini, G.; Poubeau, P.; Staikowsky, E.; Lory, M.; Le Moullec, N.; Becquart, J.P.; Wengling, C.; Michault, A.; Paganin, F. Outbreak of chikungunya on Reunion Island: Early clinical and laboratory features in 157 adult patients. *Clin. Infect. Dis.* **2007**, *44*, 1401–1407. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Simon, F.; Parola, P.; Grandadam, M.; Fourcade, S.; Oliver, M.; Brouqui, P.; Hance, P.; Kraemer, P.; Ali Mohamed, A.; de Lamballerie, X.; et al. Chikungunya infection: An emerging rheumatism among travelers returned from Indian Ocean islands. Report of 47 cases. *Medicine* **2007**, *86*, 123–137. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Pialoux, G.; Gauzere, B.A.; Jaureguierry, S.; Strobel, M. Chikungunya, an epidemic arbovirolosis. *Lancet Infect. Dis.* **2007**, *7*, 319–327. [\[CrossRef\]](#)

7. Manimunda, S.P.; Vijayachari, P.; Uppoor, R.; Sugunan, A.P.; Singh, S.S.; Rai, S.K.; Sudeep, A.B.; Muruganandam, N.; Chaitanya, I.K.; Guruprasad, D.R. Clinical progression of chikungunya fever during acute and chronic arthritic stages and the changes in joint morphology as revealed by imaging. *Trans. R. Soc. Trop. Med. Hyg.* **2010**, *104*, 392–399. [[CrossRef](#)] [[PubMed](#)]
8. Ali Ou Alla, S.; Combe, B. Arthritis after infection with Chikungunya virus. *Best Pract. Res. Clin. Rheumatol.* **2011**, *25*, 337–346. [[CrossRef](#)] [[PubMed](#)]
9. Holbrook, M.R.; Gowen, B.B. Animal models of highly pathogenic RNA viral infections: Encephalitis viruses. *Antivir. Res.* **2008**, *78*, 69–78. [[CrossRef](#)] [[PubMed](#)]
10. Zacks, M.A.; Paessler, S. Encephalitic alphaviruses. *Vet. Microbiol.* **2010**, *140*, 281–286. [[CrossRef](#)] [[PubMed](#)]
11. Renault, P.; Solet, J.L.; Sissoko, D.; Balleydier, E.; Larrieu, S.; Filleul, L.; Lassalle, C.; Thiria, J.; Rachou, E.; de Valk, H.; et al. A major epidemic of chikungunya virus infection on Reunion Island, France, 2005–2006. *Am. J. Trop. Med. Hyg.* **2007**, *77*, 727–731. [[PubMed](#)]
12. World Health Organization. *Chikungunya—Fact Sheet*; World Health Organization: Geneva, Switzerland, 2017; pp. 1–5.
13. Strauss, E.G.; Rice, C.M.; Strauss, J.H. Complete nucleotide sequence of the genomic RNA of Sindbis virus. *Virology* **1984**, *133*, 92–110. [[CrossRef](#)]
14. Strauss, E.G.; Rice, C.M.; Strauss, J.H. Sequence coding for the alphavirus nonstructural proteins is interrupted by an opal termination codon. *Proc. Natl. Acad. Sci. USA* **1983**, *80*, 5271–5275. [[CrossRef](#)] [[PubMed](#)]
15. De Groot, R.J.; Hardy, W.R.; Shirako, Y.; Strauss, J.H. Cleavage-site preferences of Sindbis virus polyproteins containing the non-structural proteinase. Evidence for temporal regulation of polyprotein processing in vivo. *EMBO J.* **1990**, *9*, 2631–2638. [[PubMed](#)]
16. Vasiljeva, L.; Merits, A.; Golubtsov, A.; Sizemskaja, V.; Kaariainen, L.; Ahola, T. Regulation of the sequential processing of Semliki Forest virus replicase polyprotein. *J. Biol. Chem.* **2003**, *278*, 41636–41645. [[CrossRef](#)] [[PubMed](#)]
17. Barton, D.J.; Sawicki, S.G.; Sawicki, D.L. Solubilization and immunoprecipitation of alphavirus replication complexes. *J. Virol.* **1991**, *65*, 1496–1506. [[PubMed](#)]
18. Shirako, Y.; Strauss, J.H. Regulation of Sindbis virus RNA replication: Uncleaved P123 and nsP4 function in minus-strand RNA synthesis, whereas cleaved products from P123 are required for efficient plus-strand RNA synthesis. *J. Virol.* **1994**, *68*, 1874–1885. [[PubMed](#)]
19. Tang, B.L. The cell biology of Chikungunya virus infection. *Cell. Microbiol.* **2012**, *14*, 1354–1363. [[CrossRef](#)] [[PubMed](#)]
20. Perera, R.; Owen, K.E.; Tellinghuisen, T.L.; Gorbalenya, A.E.; Kuhn, R.J. Alphavirus nucleocapsid protein contains a putative coiled coil alpha-helix important for core assembly. *J. Virol.* **2001**, *75*, 1–10. [[CrossRef](#)] [[PubMed](#)]
21. Ekstrom, M.; Liljestrom, P.; Garoff, H. Membrane protein lateral interactions control Semliki Forest virus budding. *EMBO J.* **1994**, *13*, 1058–1064. [[PubMed](#)]
22. Paredes, A.M.; Brown, D.T.; Rothnagel, R.; Chiu, W.; Schoepp, R.J.; Johnston, R.E.; Prasad, B.V. Three-dimensional structure of a membrane-containing virus. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 9095–9099. [[CrossRef](#)] [[PubMed](#)]
23. Cheng, R.H.; Kuhn, R.J.; Olson, N.H.; Rossmann, M.G.; Choi, H.K.; Smith, T.J.; Baker, T.S. Nucleocapsid and glycoprotein organization in an enveloped virus. *Cell* **1995**, *80*, 621–630. [[CrossRef](#)]
24. Mancini, E.J.; Clarke, M.; Gowen, B.E.; Rutten, T.; Fuller, S.D. Cryo-electron microscopy reveals the functional organization of an enveloped virus, Semliki Forest virus. *Mol. Cell* **2000**, *5*, 255–266. [[CrossRef](#)]
25. Lescar, J.; Roussel, A.; Wien, M.W.; Navaza, J.; Fuller, S.D.; Wengler, G.; Wengler, G.; Rey, F.A. The Fusion glycoprotein shell of Semliki Forest virus: An icosahedral assembly primed for fusogenic activation at endosomal pH. *Cell* **2001**, *105*, 137–148. [[CrossRef](#)]
26. Sourisseau, M.; Schilte, C.; Casartelli, N.; Trouillet, C.; Guivel-Benhassine, F.; Rudnicka, D.; Sol-Foulon, N.; Le Roux, K.; Prevost, M.C.; Fsihi, H.; et al. Characterization of reemerging chikungunya virus. *PLoS Pathog.* **2007**, *3*, e89. [[CrossRef](#)] [[PubMed](#)]
27. Kielian, M.; Rey, F.A. Virus membrane-fusion proteins: More than one way to make a hairpin. *Nat. Rev. Microbiol.* **2006**, *4*, 67–76. [[CrossRef](#)] [[PubMed](#)]
28. Marsh, M.; Helenius, A. Virus entry: Open sesame. *Cell* **2006**, *124*, 729–740. [[CrossRef](#)] [[PubMed](#)]

29. Russo, A.T.; White, M.A.; Watowich, S.J. The crystal structure of the Venezuelan equine encephalitis alphavirus nsP2 protease. *Structure* **2006**, *14*, 1449–1458. [[CrossRef](#)] [[PubMed](#)]
30. Hu, X.; Compton, J.R.; Leary, D.H.; Olson, M.A.; Lee, M.S.; Cheung, J.; Ye, W.; Ferrer, M.; Southall, N.; Jadhav, A.; et al. Kinetic, Mutational, and Structural Studies of the Venezuelan Equine Encephalitis Virus Nonstructural Protein 2 Cysteine Protease. *Biochemistry* **2016**, *55*, 3007–3019. [[CrossRef](#)] [[PubMed](#)]
31. Cheung, J.; Franklin, M.; Mancía, F.; Rudolph, M.; Cassidy, M.; Gary, E.; Burshteyn, F.; Love, J. Structure of the Chikungunya virus nsP2 protease. Unpublished work. 2011.
32. Shin, G.; Yost, S.A.; Miller, M.T.; Elrod, E.J.; Grakoui, A.; Marcotrigiano, J. Structural and functional insights into alphavirus polyprotein processing and pathogenesis. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 16534–16539. [[CrossRef](#)] [[PubMed](#)]
33. Atkins, G.J.; Sheahan, B.J.; Liljestrom, P. The molecular pathogenesis of Semliki Forest virus: A model virus made useful? *J. Gen. Virol.* **1999**, *80 Pt 9*, 2287–2297. [[CrossRef](#)] [[PubMed](#)]
34. Rupp, J.C.; Sokoloski, K.J.; Gebhart, N.N.; Hardy, R.W. Alphavirus RNA synthesis and non-structural protein functions. *J. Gen. Virol.* **2015**, *96*, 2483–2500. [[CrossRef](#)] [[PubMed](#)]
35. Gorchakov, R.; Frolova, E.; Sawicki, S.; Atasheva, S.; Sawicki, D.; Frolov, I. A new role for ns polyprotein cleavage in Sindbis virus replication. *J. Virol.* **2008**, *82*, 6218–6231. [[CrossRef](#)] [[PubMed](#)]
36. Lemm, J.A.; Rice, C.M. Assembly of functional Sindbis virus RNA replication complexes: Requirement for coexpression of P123 and P34. *J. Virol.* **1993**, *67*, 1905–1915. [[PubMed](#)]
37. Lemm, J.A.; Rumenapf, T.; Strauss, E.G.; Strauss, J.H.; Rice, C.M. Polypeptide requirements for assembly of functional Sindbis virus replication complexes: A model for the temporal regulation of minus- and plus-strand RNA synthesis. *EMBO J.* **1994**, *13*, 2925–2934. [[PubMed](#)]
38. Cross, R.K. Identification of a unique guanine-7-methyltransferase in Semliki Forest virus (SFV) infected cell extracts. *Virology* **1983**, *130*, 452–463. [[CrossRef](#)]
39. Laakkonen, P.; Hyvönen, M.; Peränen, J.; Kääriäinen, L. Expression of Semliki Forest virus nsP1-specific methyltransferase in insect cells and in *Escherichia coli*. *J. Virol.* **1994**, *68*, 7418–7425. [[PubMed](#)]
40. Ahola, T.; Kääriäinen, L. Reaction in alphavirus mRNA capping: Formation of a covalent complex of nonstructural protein nsP1 with 7-methyl-GMP. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 507–511. [[CrossRef](#)] [[PubMed](#)]
41. Mi, S.; Stollar, V. Expression of Sindbis virus nsP1 and methyltransferase activity in *Escherichia coli*. *Virology* **1991**, *184*, 423–427. [[CrossRef](#)]
42. Decroly, E.; Ferron, F.; Lescar, J.; Canard, B. Conventional and unconventional mechanisms for capping viral mRNA. *Nat. Rev. Microbiol.* **2011**, *10*, 51–65. [[CrossRef](#)] [[PubMed](#)]
43. Laakkonen, P.; Ahola, T.; Kaariainen, L. The effects of palmitoylation on membrane association of Semliki forest virus RNA capping enzyme. *J. Biol. Chem.* **1996**, *271*, 28567–28571. [[CrossRef](#)] [[PubMed](#)]
44. Ahola, T.; Lampio, A.; Auvinen, P.; Kaariainen, L. Semliki Forest virus mRNA capping enzyme requires association with anionic membrane phospholipids for activity. *EMBO J.* **1999**, *18*, 3164–3172. [[CrossRef](#)] [[PubMed](#)]
45. Lampio, A.; Kilpeläinen, I.; Pesonen, S.; Karhi, K.; Auvinen, P.; Somerharju, P.; Kaariainen, L. Membrane binding mechanism of an RNA virus-capping enzyme. *J. Biol. Chem.* **2000**, *275*, 37853–37859. [[CrossRef](#)] [[PubMed](#)]
46. Spuul, P.; Salonen, A.; Merits, A.; Jokitalo, E.; Kaariainen, L.; Ahola, T. Role of the amphipathic peptide of Semliki forest virus replicase protein nsP1 in membrane association and virus replication. *J. Virol.* **2007**, *81*, 872–883. [[CrossRef](#)] [[PubMed](#)]
47. Ahola, T.; Kujala, P.; Tuittila, M.; Blom, T.; Laakkonen, P.; Hinkkanen, A.; Auvinen, P. Effects of palmitoylation of replicase protein nsP1 on alphavirus infection. *J. Virol.* **2000**, *74*, 6725–6733. [[CrossRef](#)] [[PubMed](#)]
48. Gorbalenya, A.E.; Koonin, E.V.; Donchenko, A.P.; Blinov, V.M. Two related superfamilies of putative helicases involved in replication, recombination, repair and expression of DNA and RNA genomes. *Nucleic Acids Res.* **1989**, *17*, 4713–4730. [[CrossRef](#)] [[PubMed](#)]
49. Vasiljeva, L.; Merits, A.; Auvinen, P.; Kaariainen, L. Identification of a novel function of the alphavirus capping apparatus. RNA 5'-triphosphatase activity of nsP2. *J. Biol. Chem.* **2000**, *275*, 17281–17287. [[CrossRef](#)] [[PubMed](#)]
50. Karpe, Y.A.; Aher, P.P.; Lole, K.S. NTPase and 5'-RNA triphosphatase activities of Chikungunya virus nsP2 protein. *PLoS ONE* **2011**, *6*, e22336. [[CrossRef](#)] [[PubMed](#)]

51. Rikonen, M.; Peränen, J.; Kääriäinen, L. ATPase and GTPase activities associated with Semliki Forest virus nonstructural protein nsP2. *J. Virol.* **1994**, *68*, 5804–5810. [[PubMed](#)]
52. Kim, K.H.; Rumenapf, T.; Strauss, E.G.; Strauss, J.H. Regulation of Semliki Forest virus RNA replication: A model for the control of alphavirus pathogenesis in invertebrate hosts. *Virology* **2004**, *323*, 153–163. [[CrossRef](#)] [[PubMed](#)]
53. Vasiljeva, L.; Valmu, L.; Kääriäinen, L.; Merits, A. Site-specific protease activity of the carboxyl-terminal domain of Semliki Forest virus replicase protein nsP2. *J. Biol. Chem.* **2001**, *276*, 30786–30793. [[CrossRef](#)] [[PubMed](#)]
54. Ramakrishnan, C.; Kutumbarao, N.H.V.; Suhitha, S.; Velmurugan, D. Structure–function relationship of Chikungunya nsP2 protease: A comparative study with papain. *Chem. Biol. Drug Des.* **2017**, *89*, 772–782. [[CrossRef](#)] [[PubMed](#)]
55. Lulla, A.; Lulla, V.; Tints, K.; Ahola, T.; Merits, A. Molecular Determinants of Substrate Specificity for Semliki Forest Virus Nonstructural Protease. *J. Virol.* **2006**, *80*, 5413–5422. [[CrossRef](#)] [[PubMed](#)]
56. Gorchakov, R.; Frolova, E.; Frolov, I. Inhibition of transcription and translation in Sindbis virus-infected cells. *J. Virol.* **2005**, *79*, 9397–9409. [[CrossRef](#)] [[PubMed](#)]
57. Breakwell, L.; Dosenovic, P.; Karlsson Hedestam, G.B.; D’Amato, M.; Liljestrom, P.; Fazakerley, J.; McInerney, G.M. Semliki Forest virus nonstructural protein 2 is involved in suppression of the type I interferon response. *J. Virol.* **2007**, *81*, 8677–8684. [[CrossRef](#)] [[PubMed](#)]
58. Frolov, I.; Garmashova, N.; Atasheva, S.; Frolova, E.I. Random insertion mutagenesis of sindbis virus nonstructural protein 2 and selection of variants incapable of downregulating cellular transcription. *J. Virol.* **2009**, *83*, 9031–9044. [[CrossRef](#)] [[PubMed](#)]
59. Bhalla, N.; Sun, C.; Matthew Lam, L.K.; Gardner, C.L.; Ryman, K.D.; Klimstra, W.B. Host translation shutoff mediated by non-structural protein 2 is a critical factor in the antiviral state resistance of Venezuelan equine encephalitis virus. *Virology* **2016**, *496*, 147–165. [[CrossRef](#)] [[PubMed](#)]
60. Malet, H.; Coutard, B.; Jamal, S.; Dutartre, H.; Papageorgiou, N.; Neuvonen, M.; Ahola, T.; Forrester, N.; Gould, E.A.; Lafitte, D.; et al. The crystal structures of Chikungunya and Venezuelan equine encephalitis virus nsP3 macro domains define a conserved adenosine binding pocket. *J. Virol.* **2009**, *83*, 6534–6545. [[CrossRef](#)] [[PubMed](#)]
61. Tuittila, M.T.; Santagati, M.G.; Røyttä, M.; Määttä, J.A.; Hinkkanen, A.E. Replicase complex genes of Semliki Forest virus confer lethal neurovirulence. *J. Virol.* **2000**, *74*, 4579–4589. [[CrossRef](#)] [[PubMed](#)]
62. Forrester, N.L.; Guerbois, M.; Seymour, R.L.; Spratt, H.; Weaver, S.C. Vector-borne transmission imposes a severe bottleneck on an RNA virus population. *PLoS Pathog.* **2012**, *8*, e1002897. [[CrossRef](#)] [[PubMed](#)]
63. Weston, J.; Villoing, S.; Bremont, M.; Castric, J.; Pfeffer, M.; Jewhurst, V.; McLoughlin, M.; Rodseth, O.; Christie, K.E.; Koumans, J.; et al. Comparison of two aquatic alphaviruses, salmon pancreas disease virus and sleeping disease virus, by using genome sequence analysis, monoclonal reactivity, and cross-infection. *J. Virol.* **2002**, *76*, 6155–6163. [[CrossRef](#)] [[PubMed](#)]
64. Rubach, J.K.; Wasik, B.R.; Rupp, J.C.; Kuhn, R.J.; Hardy, R.W.; Smith, J.L. Characterization of purified Sindbis virus nsP4 RNA-dependent RNA polymerase activity in vitro. *Virology* **2009**, *384*, 201–208. [[CrossRef](#)] [[PubMed](#)]
65. Rupp, J.C.; Jundt, N.; Hardy, R.W. Requirement for the Amino-Terminal Domain of Sindbis Virus nsP4 during Virus Infection. *J. Virol.* **2011**, *85*, 3449–3460. [[CrossRef](#)] [[PubMed](#)]
66. Tomar, S.; Hardy, R.W.; Smith, J.L.; Kuhn, R.J. Catalytic core of alphavirus nonstructural protein nsP4 possesses terminal adenylyltransferase activity. *J. Virol.* **2006**, *80*, 9962–9969. [[CrossRef](#)] [[PubMed](#)]
67. Chen, M.W.; Tan, Y.B.; Zheng, J.; Zhao, Y.; Lim, B.T.; Cornvik, T.; Lescar, J.; Ng, L.F.P.; Luo, D. Chikungunya virus nsP4 RNA-dependent RNA polymerase core domain displays detergent-sensitive primer extension and terminal adenylyltransferase activities. *Antivir. Res.* **2017**, *143*, 38–47. [[CrossRef](#)] [[PubMed](#)]
68. Pietila, M.K.; Hellstrom, K.; Ahola, T. Alphavirus polymerase and RNA replication. *Virus Res.* **2017**, *234*, 44–57. [[CrossRef](#)] [[PubMed](#)]
69. Grimley, P.M.; Berezesky, I.K.; Friedman, R.M. Cytoplasmic structures associated with an arbovirus infection: Loci of viral ribonucleic acid synthesis. *J. Virol.* **1968**, *2*, 1326–1338. [[PubMed](#)]
70. Kujala, P.; Ikaheimonen, A.; Ehsani, N.; Vihinen, H.; Auvinen, P.; Kaariainen, L. Biogenesis of the Semliki Forest virus RNA replication complex. *J. Virol.* **2001**, *75*, 3873–3884. [[CrossRef](#)] [[PubMed](#)]

71. Frolova, E.I.; Gorchakov, R.; Pereboeva, L.; Atasheva, S.; Frolov, I. Functional Sindbis virus replicative complexes are formed at the plasma membrane. *J. Virol.* **2010**, *84*, 11679–11695. [[CrossRef](#)] [[PubMed](#)]
72. Atasheva, S.; Fish, A.; Fornerod, M.; Frolova, E.I. Venezuelan equine Encephalitis virus capsid protein forms a tetrameric complex with CRM1 and importin α/β that obstructs nuclear pore complex function. *J. Virol.* **2010**, *84*, 4158–4171. [[CrossRef](#)] [[PubMed](#)]
73. Neuvonen, M.; Kazlauskas, A.; Martikainen, M.; Hinkkanen, A.; Ahola, T.; Saksela, K. SH3 domain-mediated recruitment of host cell amphiphysins by alphavirus nsP3 promotes viral RNA replication. *PLoS Pathog.* **2011**, *7*, e1002383. [[CrossRef](#)] [[PubMed](#)]
74. Ahmad, S.; Hur, S. Helicases in Antiviral Immunity: Dual Properties as Sensors and Effectors. *Trends Biochem. Sci.* **2015**, *40*, 576–585. [[CrossRef](#)] [[PubMed](#)]
75. Andrejeva, J.; Childs, K.S.; Young, D.E.; Carlos, T.S.; Stock, N.; Goodbourn, S.; Randall, R.E. The V proteins of paramyxoviruses bind the IFN-inducible RNA helicase, mda-5, and inhibit its activation of the IFN- β promoter. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 17264–17269. [[CrossRef](#)] [[PubMed](#)]
76. Diao, F.; Li, S.; Tian, Y.; Zhang, M.; Xu, L.G.; Zhang, Y.; Wang, R.P.; Chen, D.; Zhai, Z.; Zhong, B.; et al. Negative regulation of MDA5- but not RIG-I-mediated innate antiviral signaling by the dihydroxyacetone kinase. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 11706–11711. [[CrossRef](#)] [[PubMed](#)]
77. Yoneyama, M.; Kikuchi, M.; Natsukawa, T.; Shinobu, N.; Imaizumi, T.; Miyagishi, M.; Taira, K.; Akira, S.; Fujita, T. The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. *Nat. Immunol.* **2004**, *5*, 730–737. [[CrossRef](#)] [[PubMed](#)]
78. Ertel, K.J.; Benefield, D.; Castano-Diez, D.; Pennington, J.G.; Horswill, M.; den Boon, J.A.; Otegui, M.S.; Ahlquist, P. Cryo-electron tomography reveals novel features of a viral RNA replication compartment. *eLife* **2017**, *6*. [[CrossRef](#)] [[PubMed](#)]
79. Koppek, B.G.; Perkins, G.; Miller, D.J.; Ellisman, M.H.; Ahlquist, P. Three-dimensional analysis of a viral RNA replication complex reveals a virus-induced mini-organelle. *PLoS Biol.* **2007**, *5*, e220. [[CrossRef](#)] [[PubMed](#)]
80. Fontana, J.; Lopez-Iglesias, C.; Tzeng, W.P.; Frey, T.K.; Fernandez, J.J.; Risco, C. Three-dimensional structure of Rubella virus factories. *Virology* **2010**, *405*, 579–591. [[CrossRef](#)] [[PubMed](#)]
81. Welsch, S.; Miller, S.; Romero-Brey, I.; Merz, A.; Bleck, C.K.; Walther, P.; Fuller, S.D.; Antony, C.; Krijnse-Locker, J.; Bartenschlager, R. Composition and three-dimensional architecture of the dengue virus replication and assembly sites. *Cell Host Microbe* **2009**, *5*, 365–375. [[CrossRef](#)] [[PubMed](#)]
82. Froshauer, S.; Kartenbeck, J.; Helenius, A. Alphavirus RNA replicase is located on the cytoplasmic surface of endosomes and lysosomes. *J. Cell Biol.* **1988**, *107*, 2075–2086. [[CrossRef](#)] [[PubMed](#)]
83. Hellstrom, K.; Kallio, K.; Merilainen, H.M.; Jokitalo, E.; Ahola, T. Ability of minus strands and modified plus strands to act as templates in Semliki Forest virus RNA replication. *J. Gen. Virol.* **2016**, *97*, 1395–1407. [[CrossRef](#)] [[PubMed](#)]
84. Rana, J.; Rajasekharan, S.; Gulati, S.; Dudha, N.; Gupta, A.; Chaudhary, V.K.; Gupta, S. Network mapping among the functional domains of Chikungunya virus nonstructural proteins. *Proteins* **2014**, *82*, 2403–2411. [[CrossRef](#)] [[PubMed](#)]
85. Sreejith, R.; Rana, J.; Dudha, N.; Kumar, K.; Gabrani, R.; Sharma, S.K.; Gupta, A.; Vratil, S.; Chaudhary, V.K.; Gupta, S. Mapping interactions of Chikungunya virus nonstructural proteins. *Virus Res.* **2012**, *169*, 231–236. [[CrossRef](#)] [[PubMed](#)]
86. Lulla, V.; Sawicki, D.L.; Sawicki, S.G.; Lulla, A.; Merits, A.; Ahola, T. Molecular defects caused by temperature-sensitive mutations in Semliki Forest virus nsP1. *J. Virol.* **2008**, *82*, 9236–9244. [[CrossRef](#)] [[PubMed](#)]
87. Salonen, A.; Vasiljeva, L.; Merits, A.; Magden, J.; Jokitalo, E.; Kaariainen, L. Properly folded nonstructural polyprotein directs the semliki forest virus replication complex to the endosomal compartment. *J. Virol.* **2003**, *77*, 1691–1702. [[CrossRef](#)] [[PubMed](#)]
88. Shirako, Y.; Strauss, J.H. Requirement for an aromatic amino acid or histidine at the N terminus of Sindbis virus RNA polymerase. *J. Virol.* **1998**, *72*, 2310–2315. [[PubMed](#)]
89. Sawicki, D.L.; Sawicki, S.G. A second nonstructural protein functions in the regulation of alphavirus negative-strand RNA synthesis. *J. Virol.* **1993**, *67*, 3605–3610. [[PubMed](#)]
90. Fata, C.L.; Sawicki, S.G.; Sawicki, D.L. Modification of Asn374 of nsP1 suppresses a Sindbis virus nsP4 minus-strand polymerase mutant. *J. Virol.* **2002**, *76*, 8641–8649. [[CrossRef](#)] [[PubMed](#)]

91. Fros, J.J.; Pijlman, G.P. Alphavirus Infection: Host Cell Shut-Off and Inhibition of Antiviral Responses. *Viruses* **2016**, *8*. [[CrossRef](#)] [[PubMed](#)]
92. Lemm, J.A.; Bergqvist, A.; Read, C.M.; Rice, C.M. Template-dependent initiation of Sindbis virus RNA replication in vitro. *J. Virol.* **1998**, *72*, 6546–6553. [[PubMed](#)]
93. Manns, M.P.; von Hahn, T. Novel therapies for hepatitis C—One pill fits all? *Nat. Rev. Drug Discov.* **2013**, *12*, 595–610. [[CrossRef](#)] [[PubMed](#)]
94. Pokorna, J.; Machala, L.; Rezacova, P.; Konvalinka, J. Current and Novel Inhibitors of HIV Protease. *Viruses* **2009**, *1*, 1209–1239. [[CrossRef](#)] [[PubMed](#)]
95. De Clercq, E. The design of drugs for HIV and HCV. *Nat. Rev. Drug Discov.* **2007**, *6*, 1001–1018. [[CrossRef](#)] [[PubMed](#)]
96. Strauss, E.G.; de Groot, R.J.; Levinson, R.; Strauss, J.H. Identification of the active site residues in the nsP2 proteinase of Sindbis virus. *Virology* **1992**, *191*, 932–940. [[CrossRef](#)]
97. Golubtsov, A.; Kääriäinen, L.; Caldentey, J. Characterization of the cysteine protease domain of Semliki Forest virus replicase protein nsP2 by in vitro mutagenesis. *FEBS Lett.* **2006**, *580*, 1502–1508. [[CrossRef](#)] [[PubMed](#)]
98. Rausalu, K.; Utt, A.; Quirin, T.; Varghese, F.S.; Zusinaite, E.; Das, P.K.; Ahola, T.; Merits, A. Chikungunya virus infectivity, RNA replication and non-structural polyprotein processing depend on the nsP2 protease's active site cysteine residue. *Sci. Rep.* **2016**, *6*, 37124. [[CrossRef](#)] [[PubMed](#)]
99. Rikkinen, M.; Peranen, J.; Kaariainen, L. Nuclear and nucleolar targeting signals of Semliki Forest virus nonstructural protein nsP2. *Virology* **1992**, *189*, 462–473. [[CrossRef](#)]
100. Rikkinen, M.; Peranen, J.; Kaariainen, L. Nuclear targeting of Semliki Forest virus nsP2. *Arch. Virol. Suppl.* **1994**, *9*, 369–377. [[PubMed](#)]
101. Fros, J.J.; van der Maten, E.; Vlak, J.M.; Pijlman, G.P. The C-terminal domain of chikungunya virus nsP2 independently governs viral RNA replication, cytopathicity, and inhibition of interferon signaling. *J. Virol.* **2013**, *87*, 10394–10400. [[CrossRef](#)] [[PubMed](#)]
102. Bourai, M.; Lucas-Hourani, M.; Gad, H.H.; Drosten, C.; Jacob, Y.; Tafforeau, L.; Cassonnet, P.; Jones, L.M.; Judith, D.; Couderc, T.; et al. Mapping of Chikungunya virus interactions with host proteins identified nsP2 as a highly connected viral component. *J. Virol.* **2012**, *86*, 3121–3134. [[CrossRef](#)] [[PubMed](#)]
103. Tamm, K.; Merits, A.; Sarand, I. Mutations in the nuclear localization signal of nsP2 influencing RNA synthesis, protein expression and cytotoxicity of Semliki Forest virus. *J. Gen. Virol.* **2008**, *89*, 676–686. [[CrossRef](#)] [[PubMed](#)]
104. Ching, K.-C.; Ng, L.F.P.; Chai, C.L.L. A compendium of small molecule direct-acting and host-targeting inhibitors as therapies against alphaviruses. *J. Antimicrob. Chemother.* **2017**, *72*, 2973–2989. [[CrossRef](#)] [[PubMed](#)]
105. Kwong, A.D.; Kauffman, R.S.; Hurter, P.; Mueller, P. Discovery and development of telaprevir: An NS3-4A protease inhibitor for treating genotype 1 chronic hepatitis C virus. *Nat. Biotechnol.* **2011**, *29*, 993–1003. [[CrossRef](#)] [[PubMed](#)]
106. Howe, A.Y.; Venkatraman, S. The Discovery and Development of Boceprevir: A Novel, First-generation Inhibitor of the Hepatitis C Virus NS3/4A Serine Protease. *J. Clin. Transl. Hepatol.* **2013**, *1*, 22–32. [[CrossRef](#)] [[PubMed](#)]
107. Lv, Z.; Chu, Y.; Wang, Y. HIV protease inhibitors: A review of molecular selectivity and toxicity. *HIV AIDS* **2015**, *7*, 95–104. [[CrossRef](#)]
108. Bassetto, M.; de Burghgraeve, T.; Delang, L.; Massarotti, A.; Coluccia, A.; Zonta, N.; Gatti, V.; Colombano, G.; Sorba, G.; Silvestri, R.; et al. Computer-aided identification, design and synthesis of a novel series of compounds with selective antiviral activity against chikungunya virus. *Antivir. Res.* **2013**, *98*, 12–18. [[CrossRef](#)] [[PubMed](#)]
109. Das, P.K.; Puusepp, L.; Varghese, F.S.; Utt, A.; Ahola, T.; Kananovich, D.G.; Lopp, M.; Merits, A.; Karelson, M. Design and Validation of Novel Chikungunya Virus Protease Inhibitors. *Antimicrob. Agents Chemother.* **2016**, *60*, 7382–7395. [[CrossRef](#)] [[PubMed](#)]
110. Singh, K.D.; Kirubakaran, P.; Nagarajan, S.; Sakthiah, S.; Muthusamy, K.; Velmurgan, D.; Jeyakanthan, J. Homology modeling, molecular dynamics, e-pharmacophore mapping and docking study of Chikungunya virus nsP2 protease. *J. Mol. Model.* **2012**, *18*, 39–51. [[CrossRef](#)] [[PubMed](#)]

111. Nguyen, P.T.; Yu, H.; Keller, P.A. Identification of chikungunya virus nsP2 protease inhibitors using structure-base approaches. *J. Mol. Graph. Model.* **2015**, *57*, 1–8. [[CrossRef](#)] [[PubMed](#)]
112. Byler, K.G.; Collins, J.T.; Ogungbe, I.V.; Setzer, W.N. Alphavirus protease inhibitors from natural sources: A homology modeling and molecular docking investigation. *Comput. Biol. Chem.* **2016**, *64*, 163–184. [[CrossRef](#)] [[PubMed](#)]
113. Bhakat, S.; Delang, L.; Kaptein, S.; Neyts, J.; Leyssen, P.; Jayaprakash, V. Reaching beyond HIV/HCV: Nelfinavir as a potential starting point for broad-spectrum protease inhibitors against dengue and chikungunya virus. *RSC Adv.* **2015**, *5*, 85938–85949. [[CrossRef](#)]
114. Gigante, A.; Canela, M.D.; Delang, L.; Priego, E.M.; Camarasa, M.J.; Querat, G.; Neyts, J.; Leyssen, P.; Perez-Perez, M.J. Identification of 1,2,3 triazolo[4,5-d]pyrimidin-7(6H)-ones as novel inhibitors of Chikungunya virus replication. *J. Med. Chem.* **2014**, *57*, 4000–4008. [[CrossRef](#)] [[PubMed](#)]
115. Gigante, A.; Gomez-SanJuan, A.; Delang, L.; Li, C.; Bueno, O.; Gamo, A.M.; Priego, E.M.; Camarasa, M.J.; Jochmans, D.; Leyssen, P.; et al. Antiviral activity of 1,2,3 triazolo[4,5-d]pyrimidin-7(6H)-ones against chikungunya virus targeting the viral capping nsP1. *Antivir. Res.* **2017**, *144*, 216–222. [[CrossRef](#)] [[PubMed](#)]
116. Delang, L.; Li, C.; Tas, A.; Quérat, G.; Albulescu, I.C.; de Burghgraeve, T.; Guerrero, N.A.S.; Gigante, A.; Piorowski, G.; Decroly, E.; et al. The viral capping enzyme nsP1: A novel target for the inhibition of chikungunya virus infection. *Sci. Rep.* **2016**, *6*, 31819. [[CrossRef](#)] [[PubMed](#)]
117. Briolant, S.; Garin, D.; Scaramozzino, N.; Jouan, A.; Crance, J.M. In vitro inhibition of Chikungunya and Semliki Forest viruses replication by antiviral compounds: Synergistic effect of interferon- α and ribavirin combination. *Antivir. Res.* **2004**, *61*, 111–117. [[CrossRef](#)] [[PubMed](#)]
118. Rothan, H.A.; Bahrani, H.; Mohamed, Z.; Teoh, T.C.; Shankar, E.M.; Rahman, N.A.; Yusof, R. A combination of doxycycline and ribavirin alleviated chikungunya infection. *PLoS ONE* **2015**, *10*, e0126360. [[CrossRef](#)] [[PubMed](#)]
119. Scholte, F.E.; Tas, A.; Martina, B.E.; Cordioli, P.; Narayanan, K.; Makino, S.; Snijder, E.J.; van Hemert, M.J. Characterization of synthetic Chikungunya viruses based on the consensus sequence of recent E1-226V isolates. *PLoS ONE* **2013**, *8*, e71047. [[CrossRef](#)] [[PubMed](#)]
120. Coffey, L.L.; Beehar, Y.; Borderia, A.V.; Blanc, H.; Vignuzzi, M. Arbovirus high fidelity variant loses fitness in mosquitoes and mice. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 16038–16043. [[CrossRef](#)] [[PubMed](#)]
121. Rozen-Gagnon, K.; Stapleford, K.A.; Mongelli, V.; Blanc, H.; Failloux, A.-B.; Saleh, M.-C.; Vignuzzi, M. Alphavirus Mutator Variants Present Host-Specific Defects and Attenuation in Mammalian and Insect Models. *PLoS Pathog.* **2014**, *10*, e1003877. [[CrossRef](#)] [[PubMed](#)]
122. Urakova, N.; Kuznetsova, V.; Crossman, D.K.; Sokratian, A.; Guthrie, D.B.; Kolykhalov, A.A.; Lockwood, M.A.; Natchus, M.G.; Crowley, M.R.; Painter, G.R.; et al. β -D-N(4)-hydroxycytidine is a potent anti-alphavirus compound that induces high level of mutations in viral genome. *J. Virol.* **2017**. [[CrossRef](#)] [[PubMed](#)]
123. Delang, L.; Guerrero, N.S.; Tas, A.; Quérat, G.; Pastorino, B.; Froeyen, M.; Dallmeier, K.; Jochmans, D.; Herdewijn, P.; Bello, E.; et al. Mutations in the chikungunya virus non-structural proteins cause resistance to favipiravir (T-705), a broad-spectrum antiviral. *J. Antimicrob. Chemother.* **2014**, *69*, 2770–2784. [[CrossRef](#)] [[PubMed](#)]
124. Wada, Y.; Orba, Y.; Sasaki, M.; Kobayashi, S.; Carr, M.J.; Nobori, H.; Sato, A.; Hall, W.W.; Sawa, H. Discovery of a novel antiviral agent targeting the nonstructural protein 4 (nsP4) of chikungunya virus. *Virology* **2017**, *505*, 102–112. [[CrossRef](#)] [[PubMed](#)]

© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).