

**A novel interaction between the 5' untranslated region of the
virus genome and Musashi homolog 2 is essential for
Chikungunya virus genome replication**

Kaiwen Sun

**Submitted in accordance with the requirements for the degree of
Doctor of Philosophy**

School of Molecular and Cellular Biology

Faculty of Biological Sciences

The University of Leeds

December, 2022

The candidate confirms that the work submitted is his own and that appropriate credit has been given where reference has been made to the work of others.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

© 2022, The University of Leeds, Kaiwen Sun

The right of Kaiwen Sun to be identified as the author of this work has been asserted by him in accordance with the Copyright, Designs and Patents Act 1988.

ACKNOWLEDGMENT

First, I would like to express my most sincere gratitude to my primary supervisor Dr Andrew Tuplin, whose unparalleled help, support and guidance made my project possible. It has been a fantastic four-year working with you, and I am incredibly grateful for giving me the opportunity to try various biological techniques and equipment. I genuinely benefited a lot and could not have asked for a better supervisor. And huge thanks to my second supervisor Prof. Mark Harris. I really appreciate your invaluable advice and suggestions regarding my experimental designs and analysis.

I would also like to thank Dr Brian Jackson from the Protein Production Facility for his kind assistance to an amateur like me during the AKTA training.

Special thanks to the fabulous members of the Tuplin group: Dr Marietta Müller, Dr Catherine Kendall, Dr Oliver Prosser, Kate Loveday, Veronica DeJesus and Francesca Appadoo. We have always been a harmonious group that never hesitates to assist each other, and it is my absolute pleasure to be part of it. I would like to particularly thank Dr Marietta Müller for her brilliant RNA pulldown work and her generous help during my first few months in the lab.

To my darling partner and soulmate Cayla, I could never get to where I am now without your company and emotional support. It has been difficult for us both as the last time we went back home was three years ago, and yet you are always here for me when I am down. Meeting you is my greatest privilege and achievement.

Lastly, I would like to thank my parents for their endless and patient support throughout my PhD. Your love and faith in me mean more than anything.

ABSTRACT

Chikungunya virus (CHIKV) is a single-stranded, positive-sense alphavirus of the *Togaviridae* family and is transmitted among humans via *Aedes* spp. mosquitos. Typical symptoms of CHIKV infection include debilitating arthralgia which can persist for months or years. The recent re-emergence of CHIKV raises serious global health concerns due to high rates of morbidity and the lack of licensed antiviral drugs or clinically approved vaccines. Current knowledge about the molecular mechanisms controlling CHIKV replication and virus-host interactions is limited. Previous studies from our group have mapped six stem-loops within the 5' untranslated region (5' UTR) and the first ~200nt of ORF-1. Phenotypic analysis demonstrated that they are RNA replication elements (RREs) required for virus genome replication through structure-dependent mechanisms, which involve vertebrate and invertebrate-specific factors. However, the aspect of molecular virology of how the RREs function or through what interactions are yet to be investigated.

In this study, reverse genetics and biochemical approaches were used to identify and confirm a specific interaction between cellular RNA binding protein Musashi homolog 2 and this structured region of the CHIKV genome. Using electromobility shift assay, I confirmed the direct interaction between MSI2 and the 5' UTR of the CHIKV genome, with the binding site being the single-stranded region upstream of the AUG start codon. Using infectious virus and sub-genomic replicon systems, combined with RNA silencing and drug inhibition assays, it was demonstrated for the first time that MSI2 is required for CHIKV genome replication. A CHIKV *trans*-complementation system and strand-specific qRT-PCR were used to show that MSI2 is required for the initiation of negative-strand synthesis, possibly by functioning as a molecular switch for translation and replication as MSI2 also interacts directly or indirectly with viral non-structural proteins nsP1 and nsP3 – both are essential components of the viral replication complex. These findings provide novel insights into how CHIKV exploits cellular components for its replication and identify potential targets for antiviral therapy.

LIST of ABBREVIATIONS

Abbreviation	Full Description
°C	Degree Celsius
3'	Three prime
5'	Five prime
A or ATP	Adenosine
A226V	Alanine to valine at position 226 in Envelope protein E1
aa	Amino acid
ADP	Adenosine diphosphate
<i>Ae.</i>	<i>Aedes</i>
AGO2	Argonaute
AUD	Alphavirus unique domain
BCA	Bicinchoninic acid
BHK-21	Baby hamster kidney fibroblasts
bp	Base pair
BSA	Bovine serum albumin
BSM	Binding site mutant
C or CTP	Cytosine
C2C12	Mouse myoblast cell
cDNA	Complementary deoxyribonucleic acid
CHIKV	Chikungunya virus
Co-IP	Co-immunoprecipitation
CombA	Combination mutant A
CP	Capsid protein
CRM1	Chromosomal maintenance protein 1
CSE	Conserved sequence elements
C-terminus	Carboxy-terminus
DAZAP1	DAZ associated protein 1
DC-SIGN	Dendritic cell-specific intercellular adhesion molecule-3-grabbing non-integrin

DENV	Dengue virus
DEPC	Diethylpyrocarbonate
DHX9	DExH-box RNA helicase 9
DMEM	Dulbecco's modified eagle media
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTPs	Deoxyribonucleotide triphosphates
dsRNA	Double-stranded RNA
<i>E. coli</i>	<i>Escherichia coli</i>
ECSA	East Central South African
EDTA	Ethylenediaminetetraacetic acid
EEEV	Eastern equine encephalitis virus
eIF	Eukaryotic translation initiation factor
EMSA	Electromobility Shift Assay
ER	Endoplasmic reticulum
FBS	Foetal bovine serum
g	Gram
G or GTP	Guanosine
G3BP	GTPase activating protein (SH3 domain) binding protein
GMP	Guanosine monophosphate
GTase	Guanylyltransferase
HCV	Hepatitis C virus
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His	Histidine
hnRNP	Heterogeneous nuclear ribonucleoprotein
Huh7	Human hepatocellular carcinoma
HVD	Hypervariable domain
IC	Infectious clone
IFN	Interferon
IgG	Immunoglobulin G

IL-6	Interleukin 6
IOL	Indian Ocean Lineage
IP	Immunoprecipitation
ISG	IFN-stimulated genes
kb	Kilobase
kDa	Kilodalton
L	Litre
LB	Lysogeny broth
log	Logarithm
L-SIGN	Lymph node-specific intercellular adhesion molecule-3-grabbing non-integrin
m ⁷ G	N ⁷ -methylguanosine
MAR	Monomeric ADP-ribose
MCP	Monocyte chemotactic protein
mg	Milligram
miRNA	MicroRNA
mL	Millilitre
mM	Millimolar
MOI	Multiplicity of infection
MOPS	3-N-morpholinopropanesulfonic acid
mRNA	Messenger RNA
MS	Mass spectrometry
MSI	Musashi homolog
MTase	Guanine-7-methyltransferase
MTT	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
ng	Nanogram
NRAMP	Natural resistance-associated macrophage protein
nsP	Non-structural protein
nt	Nucleotide
N-terminus	Amino-terminus
NTF-2	Nuclear-transport factor 2

NTPase	Nucleoside-triphosphatase
OAS	2', 5'-oligoadenylate synthetase
ONNV	O'nyong nyong virus
ORF	Open reading frame
p	Probability
P/S	Penicillin/streptomycin
p62	Precursor of envelope protein E3 and E2
PAMP	Pathogen-associated molecular patterns
PAR	Poly (ADP-ribose)
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PFU	Plague forming unit
PK	Pseudoknot
PKR	Protein kinase R
PLB	Passive lysis buffer
pmol	Picomole
ppp	Triphosphate
PPR	Pattern recognition receptors
pSTAT	Phosphorylated signal transducer and activator of transcription
qRT-PCR	Quantitative reverse transcription polymerase chain reaction
RD	Rhabdomyosarcoma
RdRp	RNA-dependent RNA polymerase
RIG-I	Retinoic acid-inducible protein I
RISC	RNA-induced silencing complex
RLR	RIG-I-like receptors
RNA	Ribonucleic acid
RNAi	RNA interference
Ro	Ro 08-2750
RRE	RNA replication element
RRM	RNA recognition motif

RRV	Ross River virus
RSE	Repeated sequence element
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
SDS-PAGE	Sodium dodecyl sulphate–polyacrylamide gel electrophoresis
SFV	Semliki Forest virus
SGP	Sub-genomic promoter
SH3	Src homology 3
SHAPE	Selective 2'-hydroxyl acylation analysed by primer extension
shRNA	Small hairpin RNA
SINV	Sindbis virus
siRNA	Small interfering RNA
spp.	Species
ssRNA	Single-stranded RNA
SVG-2	Human foetal astrocyte cell
TAE	Tris-acetic acid-EDTA buffer
TARDBP	TAR DNA binding protein
TATase	Terminal adenosyl transferase
TF	TransFrame
TLR	Toll-like receptors
TMS-MS	Tandem mass tagging and MALDI-TOF quantitative mass spectrometry
TRBP	TAR-RNA-binding protein
tRNA	Transfer RNA
U	Enzyme unit
U or UTP	Uracil
UTR	Untranslated region
VEEV	Venezuelan equine encephalitis virus
vRC	Virus replication complex
µg	Microgram
µL	Microlitre

Table of Contents

ACKNOWLEDGMENT	III
ABSTRACT	IV
LIST of ABBREVIATIONS.....	V
TABLE of FIGURES and TABLES	XV
CHAPTER 1: INTRODUCTION.....	1
1.1 General Introduction to Chikungunya virus	1
1.1.1 Introduction to Alphaviruses	1
1.1.2 CHIKV Origin and Re-emergence	3
1.1.3 Clinical presentation	6
1.1.4 CHIKV pathogenesis	6
1.2 Molecular Biology of CHIKV.....	9
1.2.1 Genome Organisation	9
1.2.2 Virion structure	11
1.2.3 Non-structural proteins.....	12
1.2.4 Structural proteins	20
1.3 CHIKV lifecycle	24
1.4 Preliminary work deciphering the structure of the 5' end of the CHIKV genome	29
1.4.1 The dynamic stem-loops within the 5' end of the CHIKV genome.....	29
1.4.2 Identification of MSI2 as a binding partner to the 5' end of the CHIKV genome	32
1.5 Aim of This Study.....	34
CHAPTER 2: MATERIALS & METHODS.....	35
2.1 Materials.....	35

2.1.1 Human cell culture	35
2.1.2 Infectious virus construct	36
2.1.3 Sub-genomic replicon constructs	37
2.1.4 Human Musashi homolog 2 RNA recognition motifs Escherichia coli expression vector	40
2.1.5 Critical reagents, kits, antibodies and primers	41
2.2 Methods	43
2.2.1 Agarose gel electrophoresis	43
2.2.2 Denaturing MOPS agarose gel electrophoresis	43
2.2.3 Plasmid RNA linearisation	43
2.2.4 <i>In vitro</i> RNA transcription	44
2.2.5 Virus production	44
2.2.6 Virus infection	45
2.2.7 Plaque assay virus titration	45
2.2.8 RNA extraction from infected cells	45
2.2.9 Determine the cell viability of Ro 08-2750 with MTT assay	46
2.2.10 Treatment with Ro 08-2750 and infection with CHIKV	46
2.2.11 Small interfering RNA knockdown of MSI1 and MSI2	47
2.2.12 Maintenance of stable MSI2 knockdown RD cell lines using short hairpin RNA interference (shRNA)	47
2.2.13 Quantification of CHIKV genomic copies	48
2.2.14 Strand-specific quantification of CHIKV RNA	49
2.2.15 Sub-genomic replicon transfection and analysis	49
2.2.16 Trans-complementation Assay	50
2.2.17 3' end biotinylation of CHIKV RNA and pulldown with infected RD lysates	51

2.2.18 Co-immunoprecipitation (Co-IP) of MSI2 with CHIKV proteins	53
2.2.19 Expression, purification and ion exchange of His-tagged MSI2 protein	53
2.2.20 Electromobility shift assay (EMSA)	55
2.2.21 Western blot and Coomassie blue staining	55
2.2.22 Statistical analysis	56
CHAPTER 3: PHENOTYPIC ANALYSIS OF THE ROLE OF MSI2 IN CHIKV REPLICATION USING A SMALL MOLECULE INHIBITOR	57
3.1 Introduction	57
3.2 Results	59
3.2.1 Inhibition of the MSI2 RNA-recognition motif 1 impaired CHIKV genome replication	59
3.2.2 MSI2 actively participates in the CHIKV replicase complex	68
3.3 Discussion	72
3.3.1 Limitations of RNA pulldown and co-immunoprecipitation	73
3.3.2 Issues with the incompatibility of Huh7 with the <i>trans</i> -complementation systems	74
CHAPTER 4: PHENOTYPIC AND REVERSE GENETIC ANALYSIS OF THE ROLE OF MSI2 in CHIKV GENOME REPLICATION USING RNA SILENCING	75
4.1 Introduction	75
4.1.1 MSI2-specific RNA interference	75
4.1.2 The predicted MSI2 binding sites within the CHIKV 5' UTR	77
4.2 Results	80
4.2.1 The effect of MSI 1+2 siRNA knockdown on CHIKV genome replication in Huh7 cells	80
4.2.2 MSI2 positively facilitates CHIKV replication in RD cells	86
4.2.3 Downregulation of MSI2 by shRNA impaired CHIKV genome replication in RD	

cells	88
4.2.4 Mutagenesis of predicted MSI2 binding sequence on CHIKV genome significantly impaired virus replication in RD cells.....	93
4.3 Discussion	95
4.3.1 Possible reasons for incompatibility of lentivirus transduction with Huh7 cells	96
4.3.2 The proposed role of MSI2 during CHIKV genome replication	97
CHAPTER 5: BIOCHEMICAL ANALYSIS OF THE NOVEL INTERACTION BETWEEN MSI2 AND CHIKV 5' RNA	99
5.1 Introduction	99
5.1.1 Methodological rationale for investigating CHIKV RNA-MSI2 interaction...	99
5.1.2 Hypothesis underpinning the biochemical analysis	101
5.2 Results	102
5.2.1 Purification of MSI2 protein containing the two RNA binding domains....	102
5.2.2 MSI2 binds specifically to the 5' region of the CHIKV genome.....	106
5.2.3 MSI2 specifically binds to the single-stranded region upstream of the AUG start codon	111
5.3 Discussion	115
5.3.1 The proposed mechanism of MSI2 during CHIKV genome replication.....	115
5.3.2 Evaluation of EMSA	117
Chapter 6: DISCUSSION AND CONCLUSION.....	118
6.1 Summary on the role of MSI2 in CHIKV genome replication	118
6.2 Model of the proposed mechanism of MSI2 regulating the initiation of CHIKV negative-sense strand synthesis.....	120
6.3 Future perspectives	122
6.4 Conclusion	124

REFERENCES..... 125

TABLE of FIGURES and TABLES

<i>Figure 1.1.1. Phylogenic tree of alphaviruses and representative variants/subtypes (5).....</i>	<i>2</i>
<i>Figure 1.1.2. Phylogenic tree of CHIKV from each outbreak based on the complete genome sequences from GenBank (29).....</i>	<i>4</i>
<i>Figure 1.1.3. Geographical distribution of CHIKV infection cases reported in 2020 (40).</i>	<i>5</i>
<i>Figure 1.2.1. Schematic representation of the CHIKV genome.</i>	<i>11</i>
<i>Figure 1.2.2 CHIKV virion architecture (105).....</i>	<i>12</i>
<i>Figure 1.2.3. The domain organisation of alphavirus non-structural proteins (108).</i>	<i>13</i>
<i>Figure 1.2.4. The genomic organisation and processing of alphavirus ORF-2 containing the structural proteins (188).....</i>	<i>21</i>
<i>Figure 1.3.1. Schematic illustration of the CHIKV replication cycle (247).</i>	<i>28</i>
<i>Figure 1.4.1. Schematic representation of selective 2'-hydroxyl acylation analysed by primer extension (SHAPE)-constrained thermodynamically derived model of CHIKV genomic RNA folding (nucleotides 1-303) (83).</i>	<i>31</i>
<i>Figure 1.4.2. The schematic representation of the MSI2 protein (258).</i>	<i>34</i>
<i>Table 1. The name, type and species of cells used throughout this study.....</i>	<i>35</i>
<i>Figure 2.1. The representative gene map of infectious virus construct WT Full-length ICRES.</i>	<i>36</i>
<i>Figure 2.2. Replicon constructs.</i>	<i>39</i>
<i>Figure 2.3. The representative gene map of pET-22HT-MSI2(8-193) expressing the two RNA recognition motifs of human MSI2 (266).</i>	<i>40</i>
<i>Table 2. Detailed list of critical reagents, constructs and commercial assays used in this study.</i>	<i>42</i>
<i>Figure 3.1. Schematic illustration of the procedures of co-immunoprecipitation and RNA affinity purification.</i>	<i>59</i>

Figure 3.2.1. Detection of MSI2 expression in RD and Huh7 cells.	61
Figure 3.2.2. MTT assay of Ro 08-2750 in Huh7 and RD cells.	62
Figure 3.2.3. Inhibition of the RNA-binding activity of MSI2 reduced CHIKV replication.	66
Figure 3.2.4. Positive- (left) and negative-sense (right) RNA was quantified by reverse transcription followed by qPCR under Ro 08-2750 treatment in Huh7.	67
Figure 3.2.5. Biotinylation of CHIKV 5' RNA for pulldown using CHIKV-infected cell lysates.	69
Figure 3.2.6. Forward and reverse co-immunoprecipitation of MSI2 with nsP1 and nsP3 using CHIKV infected RD cell lysates.	71
Figure 4.1.1. Schematic representation of the stem-loop structures and site of mutagenesis within the CHIKV 5' end (83).	78
Figure 4.1.2. RNA folding model generated by M Fold predictions.....	79
Figure 4.2.1. Evaluation of the efficiency of MSI knockdown by western blot in Huh7 and RD cells 24 h after siRNA transfection.	83
Figure 4.2.2. Comparison of the effect between MSI1, MSI2 and MSI1+2 siRNA knockdowns on CHIKV replication in Huh7 cells.	86
Figure 4.2.3. Replication phenotype of MSI2 siRNA knockdown compared to mock and scrambled samples in RD cells.	87
Figure 4.2.4. Replication phenotype of MSI2 shRNA knockdown compared to mock and scrambled samples in RD cells.	90
Figure 4.2.5. MSI2 is required for the efficient CHIKV genome replication in RD cells.	92
Figure 4.2.6. Mutagenesis of the MSI2 binding sequence on the CHIKV genome significantly impairs replication.....	94
Figure 5.1.1. The schematic illustration of EMSA (310).	100
Figure 5.2.1. Expression and purification of MSI2 (8-193).	103
Figure 5.2.2. Further purification of MSI2 (8-193) using ion exchange.	105
Figure 5.2.3. MSI2 (8-193) titration of labelled WT CHIKV 5' RNA.....	107
Figure 5.2.4. WT and ComBA RNA competition assay.	110

Figure 5.2.5. The ability of un-labelled WT and CombA RNA to compete with labelled WT RNA for MSI2 interaction.111

Figure 5.2.6. WT and MSI2 BSM RNA competition titration assay.....114

Figure 6.1. Proposed mechanism by which MSI2 and DHX9 influence CHIKV translation/replication switch.121

CHAPTER 1: INTRODUCTION

1.1 General Introduction to Chikungunya virus

1.1.1 Introduction to Alphaviruses

The *Togaviridae* is categorised as a family of enveloped positive-sense, single-stranded RNA viruses with a broad host range. The genome of togaviruses is approximately 11-12 kb, flanked by a 5' cap 0 structure and a 3' poly (A) tail. Following recent changes to virus taxonomy, the *Rubivirus* genus was re-classified as a member of the *Metonaviridae* family and no longer belongs to the *Togaviridae* family, hence leaving alphavirus as its only genus (1). So far, there are 31 identified alphaviruses, including Chikungunya virus (CHIKV), O' nyong-nyong virus (ONNV), Semliki Forest virus (SFV) and Sindbis virus (SINV) etc. (2, 3) (Figure 1.1.1). These are classified as 'Old World alphaviruses', which are phylogenetically distinct from the 'New World alphaviruses', such as Venezuelan equine encephalitis virus (VEEV) and Eastern equine encephalitis virus (EEEV) (4). The characteristic symptom of the New World alphavirus infection is encephalitis, while for the Old World alphavirus it is arthralgia and fever (4).

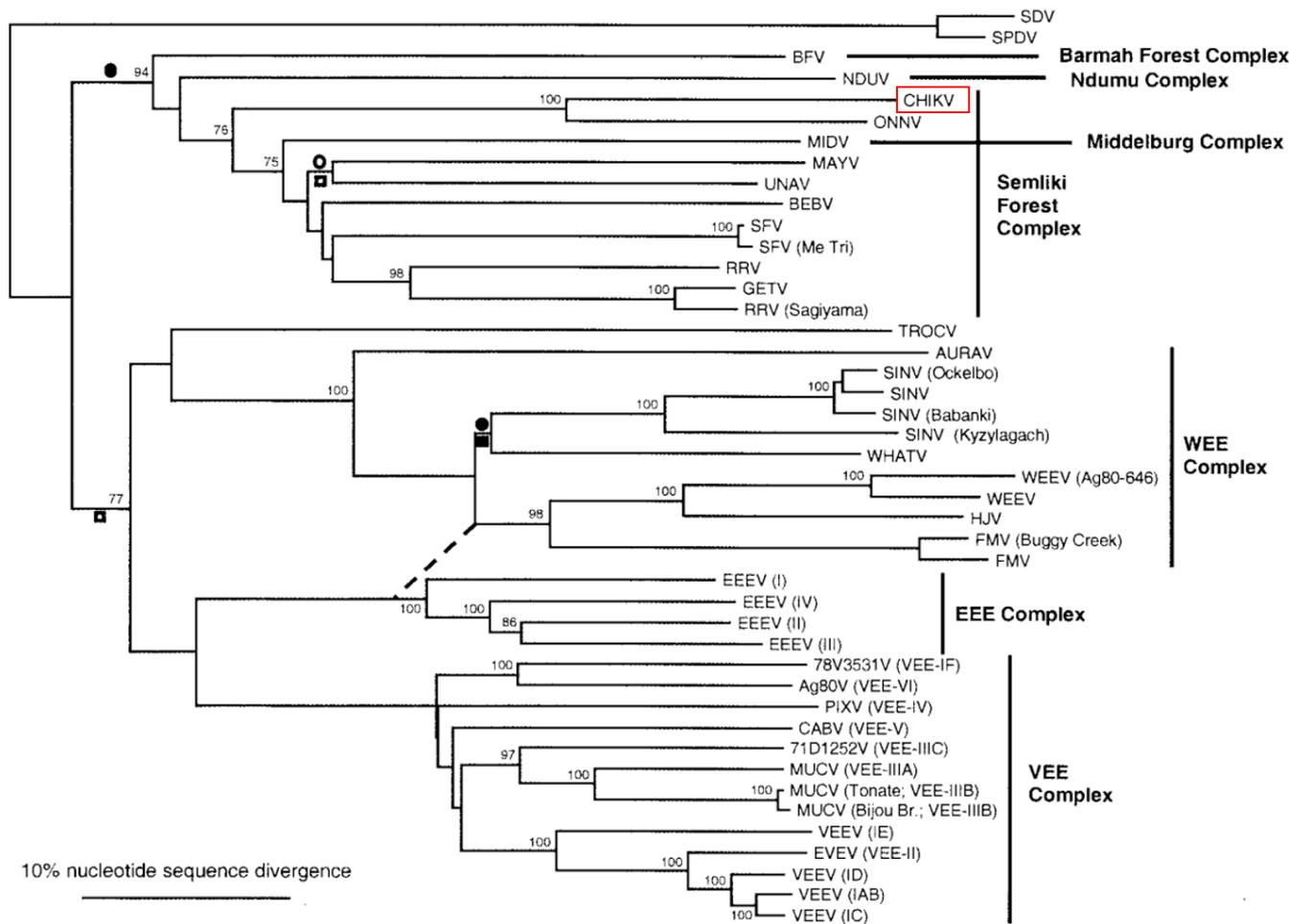


Figure 1.1.1. Phylogenetic tree of alphaviruses and representative variants/subtypes (5).

The phylogeny was generated from nucleotide sequence alignments around the E1 structural gene. CHIKV is highlighted in red. The scale bars represent the divergence of 10% genome sequence. The symbols next to the branch indicate: open circle: hypothetical Old to New World introduction (New World origin); closed circle: New to Old World introduction (New World origin); open square: Old to New World introduction (Old World origin); closed square: New to Old World introduction (Old World origin).

1.1.2 CHIKV Origin and Re-emergence

CHIKV was first reported and identified during an outbreak in Tanzania in 1952, with the most severely impacted region being the Makonde Plateau, where households were commonly infested with mosquitos due to water storage (6). Chikungunya means 'that which bends up' in the local language to describe the agonising posture of infected individuals (6). Symptoms of CHIKV infection typically include acute febrile symptoms, myalgia, rash and severe arthralgic joint pain, which may persist for years in some patients (7). The typical time elapse between the initial virus infection and the display of symptoms is between 4 to 7 days (8).

CHIKV is transmitted to humans mainly by *Aedes albopictus* and *Aedes aegypti* mosquitos (9). Before the 21st century, scattered reports of CHIKV outbreaks were recorded around the world, namely Angola (10), Bernin (11), Burundi (12), Cameroon (13), the Central African Republic (14), Democratic Republic of the Congo (15), Gabon (16), Guinea (17), Kenya (18), Liberia (19), Madagascar (20), Malawi (21), Nigeria (22), Uganda (23), Senegal (24), Sierra Leone (25), southern Africa (25), Sudan (26) and Tanzania (6). Among these territories, Africa was believed to be the origin of the virus, which spread across Asia presumably by human migratory events (27). Three lineages of CHIKV have been identified according to their nucleic acid sequence and geographical distribution: the West African lineage, the Asian lineage and the East-Central-Southern African (ECSA) lineage (Figure 1.1.2) (28).

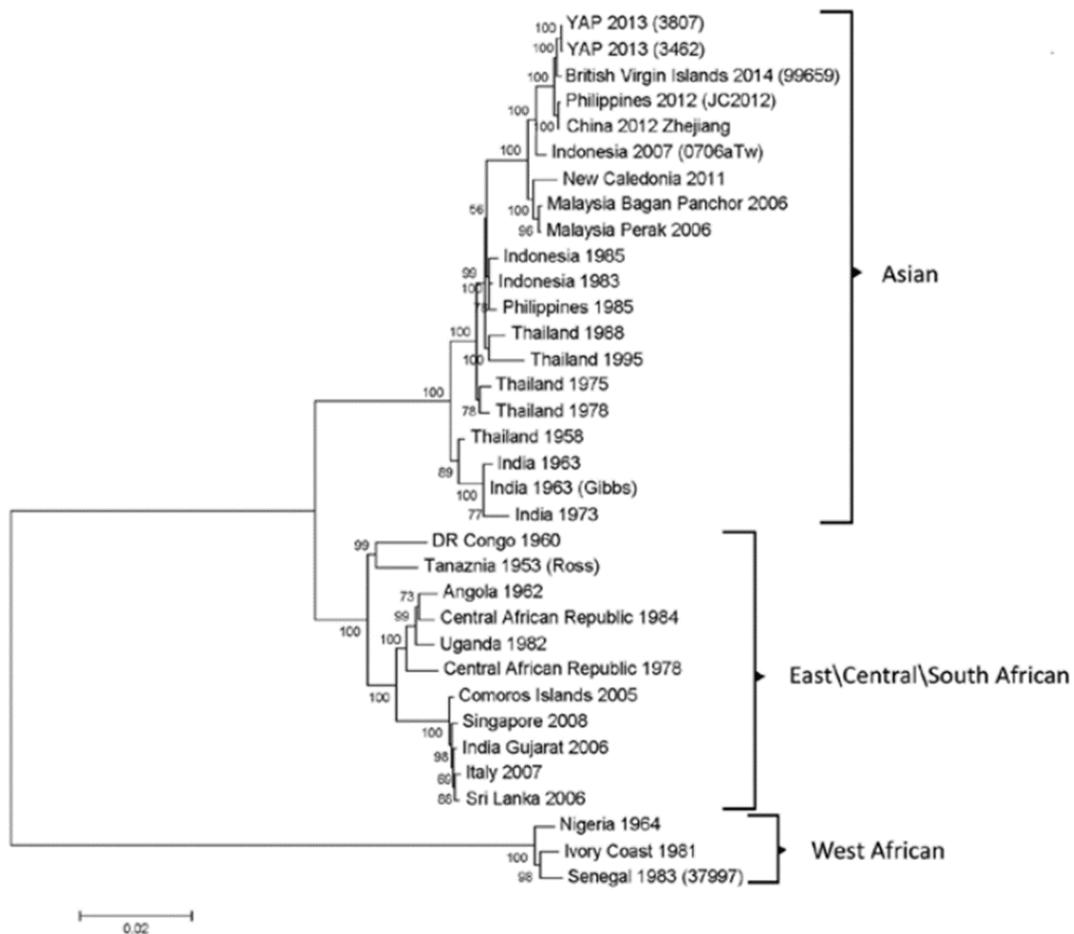


Figure 1.1.2. Phylogenetic tree of CHIKV from each outbreak based on the complete genome sequences from GenBank (29).

The scale bar indicates the number of nucleotide substitutions per site. The lineages of the strains are annotated at the right.

After approximately 50 years of low incidence, CHIKV recently re-emerged due to global urbanisation and expansion of geographical distribution of *Aedes* mosquitos, causing epidemic outbreaks across regions in Asia, Africa, America, the Middle East and Europe (Figure 1.1.3) (7, 30, 31). Although the disease is self-limiting in most cases, increasing evidence shows that the re-emergence is associated with high rates of morbidity, neurological complications and increased mortality (32, 33). In 2004, there was an explosive CHIKV outbreak in Kenya that spread as an epidemic across multiple

Indian Ocean islands (34). The epidemic strain, belonging to the ECSA lineage, had a single mutation in the envelope protein gene (E1-A226V), which significantly enhanced CHIKV infectivity in *Aedes albopictus* mosquitos, therefore accelerating and enlarging the dissemination of disease (35). It has been proposed that as valine possess more methyl groups than alanine, the decrease in polarity may alter the cholesterol dependency for CHIKV, leading to more efficient membrane fusion and virus entry (36). Notably, their nucleotide identity shared 99.9% similarity to the African strains isolated from the following outbreak in India between 2005-2006, whereas the latter does not possess an A226V mutation, suggesting that the two outbreaks were likely to be caused by two different virus strains (37). Furthermore, the rapid spread of CHIKV strains of Asian lineage was responsible for the autochthonous outbreak in the Philippines, the Caribbean and most territories of the Americas from 2013 to 2016 (29, 38, 39).

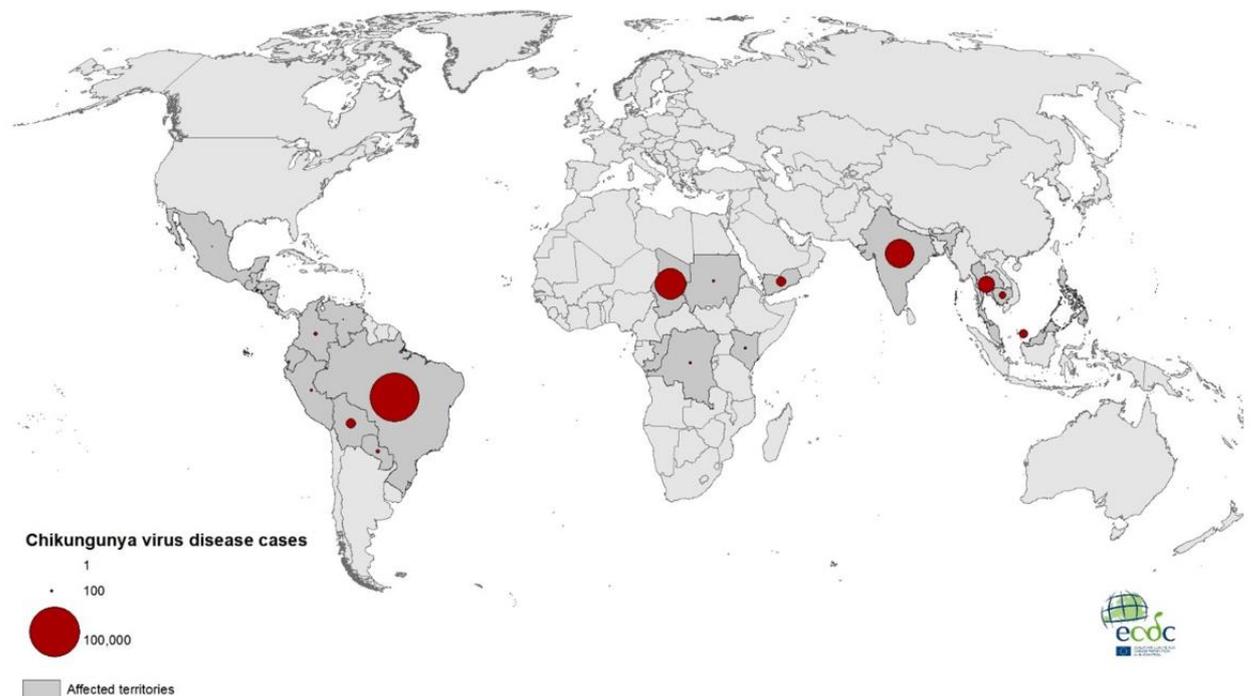


Figure 1.1.3. Geographical distribution of CHIKV infection cases reported in 2020 (40). Cases of CHIKV infection have been reported in several countries in Asia, Africa, and South America between January to December 2020. The affected territories were marked in grey, with the number of cases annotated in red circle.

1.1.3 Clinical presentation

The viral stage (day 1-4 after disease onset), coupled with the convalescence stage (day 5-14 after disease onset), describes the acute infectious phase of CHIKV infection (41). Around 90% cases have been reported to suffer from polyarthrititis, the hallmark of CHIKV infection (42). Other common symptoms of CHIKV infection include acute fever and erythematous, maculopapular rash, which can be spread over the entire body during the acute infectious phase (43). A minor of infection cases was related to retinitis, conjunctivitis and episcleritis (44). Remarkably, CHIKV infection also targets neonatal neurodevelopment, which was discovered for the first time during the Réunion Island outbreak (45). Mother-to-child transmission of CHIKV has been documented to cause microcephaly and cerebral palsy (46). The characteristic joint pain of CHIKV infection frequently takes place in proximal to wrists and ankles, suggesting the manifestation of inflammatory polyarthrititis (24, 47). About 50% of patients from the Réunion Island outbreak also experienced gastrointestinal discomfort (48). The acuteness of these symptoms usually alleviates after approximately 3 days, with the mitigation of skin rash followed by desquamation (24). While rare cases of asymptomatic infections were recorded, persistent swelling joint pain can also last up to 36 months during the so-called chronic infectious phase (49). The year long suffering of joint pain severely affected the physiological and mental health of the patient and subsequently increased the socioeconomical burden (50).

1.1.4 CHIKV pathogenesis

As the first line of defence against pathogens, the type I interferon (IFN) pathway of the innate immune system has been well-reported to play an important anti-CHIKV role (51-53). Deficiency in type I IFN has been shown to lead to lethal CHIKV infections in mice (52). Activation of type I IFN requires the recognition of pathogen-associated molecular patterns (PAMPs) by pattern recognition receptors (PPRs) (54). Multiple cellular PPRs have been shown to participate in CHIKV detection, including Toll-like receptors (TLRs) and cytosolic retinoic acid-inducible gene I (RIG-I)-like receptors (RLRs)

(55). The double-stranded RNAs (dsRNAs) synthesised as the replication intermediate of CHIKV can be recognised by TLRs/RLRs, triggering a series of downstream signalling cascade which results in the production of type I IFNs (56). It has been reported that CHIKV infection not only led to high genomic copies of several antiviral factors such as IFN α , IFN β and TLR3, but the latter was also a major determinant for the severity of disease outcome (57, 58). Another IFN-induced antiviral protein called 2', 5'-oligoadenylate synthetase (OAS) 3 was shown to effectively antagonise CHIKV infection during the early stages of virus replication (59).

The onset of the acute phase of CHIKV infection is associated with high viraemic titres of CHIKV in blood. This is accompanied by the overexpression of several pro- and anti-inflammatory cytokines and chemokines (60). In particular, the production of interleukin 6 (IL-6) in infected osteoblasts could be maintained at high levels for more than a month (61). The overexpression of IL-6 was detected to be positively correlated with high virus titre, as well as arthralgic persistence in CHIKV-infected patients (51). With the expression of IL-6 receptors on osteoblasts, positive feedback production of IL-6 may be induced, leading to the persistence of arthralgia (62, 63). This has significant implications for the role of monocytes/macrophages in CHIKV pathogenesis. The IL-6 pathway triggers the expression of monocyte chemotactic protein (MCP)-1, which is an important chemokine regulating the migration and infiltration of monocytes/macrophages to the site of infection (64, 65). A high amount of MCP-1, 2 and 3 were detected during the acute CHIKV infection, while inhibition of MCP strongly abrogates pathology (60, 66). CHIKV infection in animal models deficient in MCP-1 signalling showed an increase and prolongation of arthritis, with impediments of both pro- and anti-inflammatory responses (67). It has been speculated that monocytes/macrophages were one of the cellular factors to drive CHIKV dissemination (62, 68).

As the progression of the disease continues, cellular and humoral adaptive immune responses to virus infections initiate the specific, adaptive antiviral response. In

humans, immunoglobulin G3 (IgG3) has been found to be the dominant isotype for CHIKV-specific response (69). High and persistent viral load leading to severe disease outcome was seen in B cell knockout mice, while prophylactic administration of CHIKV-specific monoclonal antibodies to T and B cell-deficient mice progressively inhibited persistent viraemia (70, 71). Using chimeric viruses, it has been demonstrated that the primary target of antibody-mediated immune response is the CHIKV envelope glycoprotein E2 (69, 72). In particular, the key antigenic determinant has been reported to be the domain B of E2, which after fusion with the surface-exposed area of domain A, sufficiently protected mice from virus accumulation (73). Other CHIKV-specific neutralising antibodies also recognise epitopes on capsid and E1 protein, with essential clinical implications that they have been shown to cross-reactive with most CHIKV genotypes as well as other alphaviruses such as SINV (74, 75). On the other hand, given the potential pro-viral role of B cells and antibodies implicated in the entry of many viruses, the antibody-mediated response to CHIKV infection may contribute to disease exacerbation known as antibody-dependent enhancement (76). A recent study suggested that antibodies presented at the sub-neutralising level enhanced CHIKV infection and aggravated the severity of disease as a result of increased CHIKV attachment (77). Therefore, careful considerations need to be taken when developing humoral-based anti-CHIKV therapies.

T cell-mediated cytotoxicity is a critical antiviral resource regulating the elimination of infected cells. Flow cytometry analysis of circulating T lymphocytes in CHIKV-infected patients demonstrated that CD8⁺ T cells peaked during the early stages of the disease, whereas CD4⁺ T cells dominated the immune response in the later stages (78). High percentages of both T cells in activated and effector forms were found in CHIKV-infected patients as well as patients with rheumatoid arthritis, suggesting that the diagnosis of the latter should take CHIKV infection into account (43). Interestingly, a pathogenic role of CD4⁺ T cells has been elucidated that they were specifically responsible for developing joint inflammation without antagonising virus replication and dissemination (79). Despite their ability to suppress viraemia, CD4⁺ T cells

activated via whole inactivated virus vaccination to B cell-deficient mice showed earlier and more severe manifestations of joint disease compared to unvaccinated controls (80). A follow-up study further demonstrated the association between CD4+ T cells and disease phenotype that restoration of severe joint swelling in T cell receptor knockout mice was achieved by the transfer of splenic CD4+ T cells (81).

1.2 Molecular Biology of CHIKV

1.2.1 Genome Organisation

CHIKV is a ~70 nm enveloped, positive-sense single-stranded (ss) RNA virus with a ~11.8 kb genome. It contains two open reading frames (ORFs) flanked by 5' and 3' untranslated regions (UTR) (Figure 1.2.1). The 5' UTR contains 76 nucleotides and is capped by 5' type-0 N-7-methylguanosine for commencing cap-dependent translation. The 3' UTR contains a poly (A) tail, which mimics cellular mRNAs poly (A) tail during initiation of virus genome translation (82). Additionally, there is a UTR at the junction region between ORF-1 and ORF-2 and is designated the sub-genomic promoter (SGP) as its complementary negative-sense strand contains the promoter sequence for the synthesis of sub-genomic RNA (83).

ORF-1, which occupies the 5'-proximal two third of the genome, encodes the viral non-structural proteins 1-4 (nsP1-4). They are translated directly from the genomic RNA, predominantly synthesising the polyprotein P123, due to the presence of an opal stop codon UGA between nsP3 and nsP4 coding regions (84). Through the viral protease nsP2, the polyprotein precursor is post-translationally processed into the early replicase P123 + nsP4, where nsP4 functions as the viral RNA-dependent RNA polymerase (RdRp). Replication of the full-length negative-sense strand RNA is then initiated by genomic RNA foldback, to enable interaction between host and viral proteins associated with the viral 5' and 3' UTRs (85). This leads to the formation of a dsRNA replication intermediate (86). Concomitantly, the early replicase P123 + nsP4 is

further processed into individual nsP1, nsP2, nsP3 and nsP4. This 'late replicase' can recognise both the genomic and subgenomic RNA replication promoters, thereby generating the positive-sense genomic and sub-genomic RNA. The latter encodes the five structural polypeptides of ORF-2, namely the capsid protein (CP), E3, E2, 6K/TF and E1, which are produced following proteolytic cleavage (87).

The 5' UTR of the CHIKV genome accommodates two conserved sequence elements (CSEs): the first 44 nt of the 5' genome constitutes CSE1, and a 51 nt element designated CSE2 located within the nsP1 coding region (88). Mutagenesis analysis has shown that both sequence and secondary structures involved in the CSEs are crucial for RNA synthesis (89-91). A recent study from our group specifically demonstrated that CSE2, which consists of two stem-loops in CHIKV, acts in a host-specific manner that is required for genome replication in vertebrates but not mosquito cells (83). Given that mutations of CSE2 in SINV resulted in replication diminishment exclusively in invertebrate cells, these studies highlighted the divergence of the replication mechanism employed by different alphaviruses (92). Importantly, the 5' UTR RNA structures of alphaviruses has also been shown to function as *cis*-acting elements, which are indispensable for the initiation of negative-sense strand synthesis (86, 93). This suggests that although virus genome replication initiates at the 3' UTR, the sequence and structural elements within the 5' UTR is also crucial for this process (85).

The 3' end of most alphaviruses share a common sequence arrangement, with short repeated sequence elements (RSEs) followed by a 19 nt CSE4 directly upstream of the poly (A) tail (94). The sequence and location of CSE4 are extremely conserved in all alphaviruses (94). It has been shown that mutations in CSE4 lead to reduced SINV replication and plaque size (95). Specifically, the last 13 nt of CSE4, together with a minimum of 12 poly (A) residues, are the essential requirements for the initiation of negative-sense strand synthesis (96). A follow-up study identified the single cytosine of CSE immediately upstream of the poly (A) tail being the initiation site for negative-sense strand synthesis (97). Moreover, the 3' UTR has been reported to directly

interact with the cellular protein HuR, which normally regulates the expression of cellular mRNA (98). A U-rich sequence upstream of the CSE4 has a high affinity for HuR, leading to its relocalisation and sequestration in the cytoplasm to progress the infection cycle (99). Given its potency in interacting with host factors and being highly conserved, the 3' UTR of alphavirus has been suggested to play a specific role in mosquito adaptation (66).

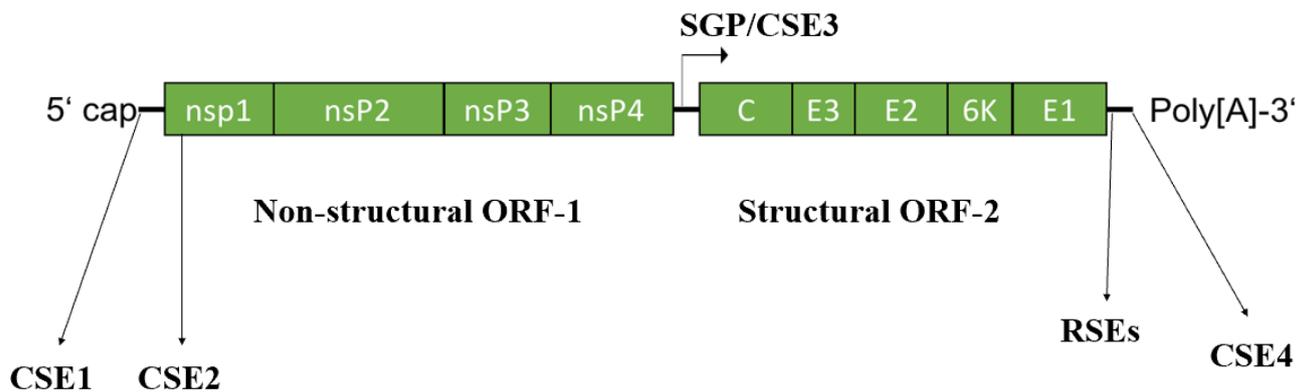


Figure 1.2.1. Schematic representation of the CHIKV genome.

CHIKV contains two ORFs flanked by a 5' m7G cap and a 3' poly(A) tail. The first one encodes the four non-structural proteins: nsP1, nsP2, nsP3 and nsP4; ORF-2 encodes the five structural proteins expressed from the sub-genomic promoter (SGP): capsid, E3, E2, 6K and E1. The position of the CSEs and RSEs are annotated along the genome accordingly.

1.2.2 Virion structure

CHIKV is a spherical, icosahedral, enveloped virus with a diameter of approximately 70nm and a triangulation number of 4 (Figure 1.2.2) (100). Average alphavirus particles have a molecular mass of 5.2×10^6 Da and a density of 1.22 g/cm^3 (101, 102). The virion is surrounded by a host-derived lipid bilayer containing 240 units of E1 and E2 (82). These two transmembrane proteins are arranged in an icosahedral lattice, forming 80 protruding spikes via heterogenous trimerisation (103, 104). Inside the

virion, it harbours the virus positive-sense, single-stranded RNA genome, which is packaged into the nucleocapsid consisting of capsid proteins.

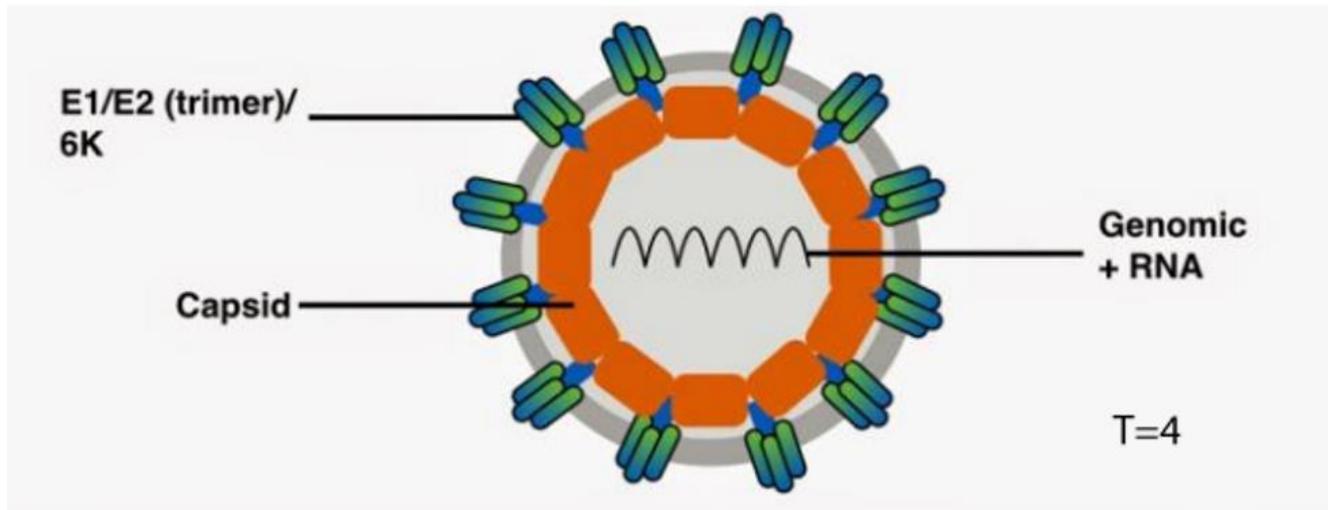


Figure 1.2.2 CHIKV virion architecture (105).

CHIKV is a ~70nm enveloped virus. The positive-sense, single-stranded RNA genome is packaged within the nucleocapsid with an icosahedral symmetry of T=4. The nucleocapsid is further surrounded by a host-derived lipid membrane with trimers of glycoprotein E1 and E2 forming the protruding spikes.

1.2.3 Non-structural proteins

The four non-structural proteins of CHIKV are expressed from ORF-1 (Figure 1.2.3). Briefly, nsP1 is responsible for the association between the replicase and intracellular membranes and functions as methyltransferase and guanylylmethylase during CHIKV replication; nsP2 plays various roles including acting as protease, helicase, RNA triphosphatase and modulating pathogenicity and host protein shutoff; nsP3 is a three-domain phosphoprotein which has been shown to interact with multiple cellular factors such as Ras GTPase-activating protein-binding protein 1/2 (G3BP1/2) and DExH-box helicase 9 (106, 107) ; nsP4 is the viral RdRp.

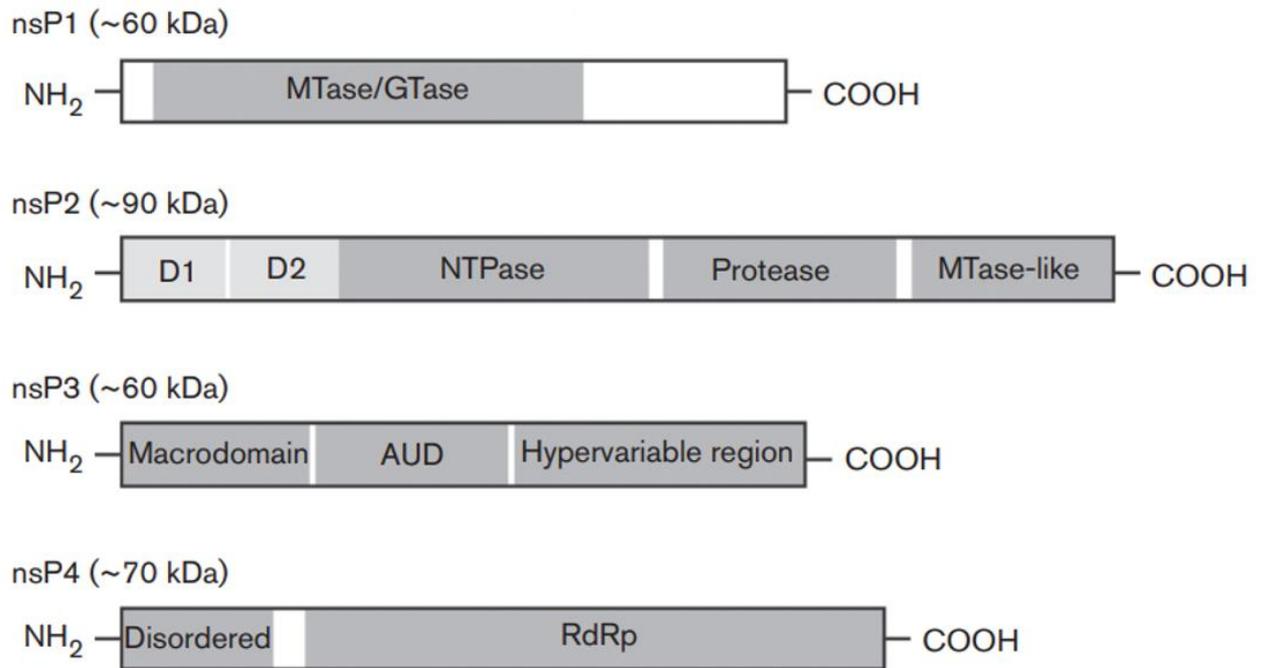


Figure 1.2.3. The domain organisation of alphavirus non-structural proteins (108).

Domain organisation of the alphavirus non-structural proteins nsP1-4. Domains with known enzymatic functions are highlighted in grey boxes at the relative positions within the representation for each protein. The approximate molecular weight of each protein is indicated above the schematic representation, which varies in length to reflect the size difference. MTase, methyltransferase; GTase, guanylyltransferase. NTPase, nucleoside triphosphatase. AUD, alphavirus unique domain.

1.2.3.1 Non-structural protein 1

One of the major roles of nsP1 is the addition of the 5' cap to the nascent genomic and sub-genomic RNA (109). The N-terminal domain of nsP1 contains motifs with methyl- and guanylyl-transferase activities (MTase/GTase) (Figure 1.2.3) (110). In contrast to eukaryotic mRNA capping, nsP1 catalyses the transfer of methyl group from S-adenosylmethionine to the 7th position of GTP, followed by the association of the methylated residue with the alphavirus 5' UTR to achieve a 7-methylguanosine (m7G) cap structure (cap 0) (111). The MTase activity was found to be a prerequisite to GTase, which hydrolyses GMP to ⁷MeGMP (112). Pro³⁴ and His³⁷ of nsP1 are the binding sites

for ⁷MeGMP, while Asp⁸⁹, Arg⁹² and Tyr²⁴⁸ are involved in the MTase activity of nsP1 (113). By possessing the cap-synthesising machinery, this nsP1-mediated capping pathway not only mimics the cellular mRNAs to facilitate cap-dependent translation of the virus genome, but also prevents the viral RNA from being degraded in the processing body (114, 115). Capped viral RNAs are also advantageous in disguising their 'non-self' identifies so that the innate immune response would not be triggered (116).

Interestingly, not all CHIKV genomic RNA are capped by nsP1 during the early stage of infection (117). Although their precise role remains obscure, it was speculated that the non-capped RNA molecules were essential components for productive virus infection as they retained the ability to be packaged into nucleocapsids, displayed higher decay rate and were potent inducers of innate immune response (117).

Also within the N-terminal domain of nsP1 are a membrane-binding amphipathic helix and palmitoylation site, both of which anchor nsP1 and other virus proteins to the host lipid membrane (Figure 1.2.3) (118, 119). nsP1 is the only viral protein which possesses membrane binding ability. The two hydrophobic loops were essential for guiding nsP1 to the inner layer of the plasma membrane via S-palmitoylation (120). Palmitoylated cysteine residues of CHIKV nsP1 have been shown to associate with cholesterol-rich domains of the phospholipid with high affinity (121, 122). Indeed, a recent study using cryo-electron microscopy revealed that the amphipathic helix specifically anchors to the negatively charged cellular phospholipid, while the cytoplasmic domain of nsP1 induces structural changes to the outer layer of the membrane (123). Consistently, nsP1-mediated membrane reorganisation has been visualised to be dodecameric rings, which suggests it may function as a portal, regulating the access of host/virus factors to the replicase complex and the exit of nascent viral RNA (124).

1.2.3.2 Non-structural protein 2

The ~90 kDa nsP2 possesses three catalytic functions, namely helicase, nucleoside triphosphatase (NTPs) and protease (Figure 1.2.3). It was originally believed that nsP2 only consisted of an N-terminal helicase/NTPs domain and a C-terminal protease domain (108). However, crystallography of the nsP2 C-terminal revealed a third domain which was structurally similar to S-adenosyl-L-methionine-dependent RNA MTases (125). Further exploration of the structure of nsP2 demonstrated two more potential domains proximal to the N-terminal, with the first one acting as a putative co-factor to the protease domain and the second involved in promoter recognition (126).

Numerous studies have shown that the helicase/NTPase domain facilitates the 5' to 3' unwinding of RNA secondary structures during alphavirus replication (Figure 1.2.3) (127-129). The NTPases of nsP2 catalyses the Mg^{2+} -dependent formation of the 5' diphosphate group by removing the γ -phosphate, thereby preparing the nascent RNA for nsP1-mediated capping (109, 130). Functional NTPase activity is required for the helicase activity, as mutations of the active-site lysine within the classical nucleotide-binding motif A abolished the hydrolytic ability of NTPase as well as the helicase activity (131). Interestingly, this unwinding process presumably happens concomitantly with the synthesis of nascent RNA controlled by nsP4, whose addition of nucleotides to the growing genomic or sub-genomic RNA immediately precedes the separation and reannealing the dsRNA intermediate (132).

The processing of alphavirus polyproteins exclusively requires the C-terminal protease activity of nsP2 (Figure 1.2.3) (133). The C-terminal recombinant protein alone retains the protease activity identical to full-length nsP2, indicating the partial functional independency of the helicase/NTPase activity (134). It has been demonstrated that the protease domain functions via deprotonating a thiol group of cysteine residue from the active-site dyads Cys¹⁰¹³ and His¹⁰⁸³ (135, 136). The non-structural proteins are initially synthesised as a single polyprotein P1234 and subsequently cleaved into

P123+nsP4 and then nsP1-4 as the virus replication progresses. The protease therefore needs to efficiently catalyse the cleavage of 3 junction sites. During the early stages of virus replication, *cis*-cleavage of the 3/4 junction is favoured as this site can only be cleaved by proteases associated with nsP3 (137). As P1234 accumulates, cleavage could also presumably be processed in *trans* when the polyproteins are in proximity to each other (82). Notably, cleavage of the 2/3 junction requires the removal of nsP1 to expose the active site within the N-terminal of nsP2, whose distant location to the 2/3 junction implies *trans*-cleavage (138, 139). Indeed, 3D modelling of nsP2 indicates the correlations between the stabilisation of N-terminal conformation and its maximal enzymatic activities (132). As a result, during later stages of the replication the protease profile favours the cleavage of 1/2 and 2/3, yielding a pool of stable P34 and nsP4 hypothetically responsible for the temporal regulation RNA synthesis (137). This regulatory mechanism has recently been reported to be crucial for CHIKV adaptations in mosquito cells (140).

Besides the enzymatic roles, nsP2 has also been extensively investigated in host transcriptional shut-off and virus-induced cytotoxicity (141-143). Mutagenesis of nsP2 suggested that surface-exposed residues 674-688 of the protease domain of SINV prominently induce cellular transcriptional shut-off by degrading the catalytic subunit of cellular DNA-dependent RNA polymerase II as well as the characteristic cytopathic effect (144). Such global inhibition of transcription, which can also be induced via nsP2 protease-independent ubiquitination of the catalytic subunit of RNA polymerase II, provides an intervention against cellular antiviral response (145). The C-terminal MTase-like domain of nsP2 also contributes to the virus counter defence strategies by impairing the IFN signalling pathway (Figure 1.2.3) (146). This domain alone has been demonstrated to sufficiently enhance the nuclear export of phosphorylated signal transducer and activator of transcription 1 (pSTAT1), one of the essential transcription factors regulating the activation of IFN-stimulated genes (ISGs) (146, 147). Potential synergetic coordination of the MTase-like domain with the protease domain has also been shown to antagonise the immune response, as mutations disabling the nsP2-

mediated host-cell shut-off rendered nsP2 defective in preventing the cellular stress response (148).

1.2.3.3 Non-structural protein 3

The alphavirus nsP3 (~60 kDa) consists of three well-studied domains: the macrodomain, the alphavirus unique domain (AUD), and the hypervariable domain (HVD) (Figure 1.2.3). The N-terminal macrodomain showed high conservation among alphaviruses, while homologous domains have also been detected in proteins of mammalian cells, bacteria and other positive-sense RNA viruses (149, 150). The macrodomain of alphaviruses possesses nucleic acid binding ability and phosphatase activity, as biochemical analysis of CHIKV macrodomain indicated its binding to DNA, RNA, monomeric ADP-ribose (MAR), poly (ADP-ribose) (PAR), as well as its dephosphorylation of ADP-ribose-1'-phosphate to ADP-ribose (151, 152). However, the macrodomain of SFV has an affinity for PAR but not MAR and exhibits low phosphatase activity to ADP-ribose-1'-phosphate, suggesting similar but not identical functional properties of the macrodomain between different alphaviruses (151, 153). For CHIKV, it has been demonstrated that the nsP3 macrodomain specifically hydrolysed MAR from mono ADP-ribosylated aspartate and glutamate but not lysine residues, while inhibition of the hydrolytic activity led to lethal phenotypes in both mammalian and mosquito cells (3). The ADP-ribose binding region of the macrodomain has been suggested to be critical for virus replication and virulence (154). In addition, the correct assembly of the structural elements within the macrodomain of SFV nsP3 has been shown to be essential for the precise positioning of nsP2 protease active sites to the 2/3 cleavage junction (155).

The central portion of nsP3 harbours the AUD, which is strongly homologous in sequence among alphaviruses (Figure 1.2.3). Compared to the other two domains of nsP3, the function of AUD is not well characterised. Crystallographic analysis of SINV AUD revealed its zinc-binding affinity (138). Four conserved cysteine residues at position 263, 265, 288 and 306, with the former two embedded in the α -helical loops

of the latter two, have been found crucial for virus replication (138). Mutagenesis within the AUD has been characterised to impair negative-sense strand synthesis, polyprotein processing and neurovirulence (156-158). A recent study further demonstrated the role of AUD in alphavirus replication and assembly that one AUD mutant (P247A/V248A) remarkably impeded virus production by blocking the transcription of sub-genomic RNA, resulting in reduced expression of viral structural proteins in vertebrate and invertebrate cells (159). The *in vitro* binding affinity of this mutant to the sub-genomic RNA promoter was dramatically reduced, while the cellular and viral interaction partners have yet to be determined.

Unlike the first two domains, the C-terminal domain of nsP3 is characteristically variable in terms of sequence and length among alphaviruses (Figure 1.2.3). This so-called HVD lacks any major predicted secondary structures and is intrinsically disordered in all alphaviruses (160, 161). Genetic manipulations within the HVD have exhibited high functional tolerance of changes, which provides a plausible insertion site for marker proteins (110, 157, 162). Despite its lack of conservation, most alphaviruses share a common amino acid arrangement within the HVD, including a hyperphosphorylated region and a proline-rich region and repeat motifs of Phe-Gly-Asp-Phe (FGDF) near the C-terminal (163-166). NsP3 is the only non-structural protein that can be phosphorylated, with phosphoserine and phosphothreonine residues detected in SFV-infected mouse cells (167). Deletion of HVD phosphorylated sites is associated with reduced neurovirulence in mice and inhibition of negative-sense strand synthesis (156, 168). The proline rich region of HVD has been shown to bind to Src homology 3 (SH3)-domain containing proteins, which are associated with signalling pathways of cytoskeleton regulation (169). Host cell proteins containing this domain, such as amphiphysin-1 and Bin1/amphiphysin-2, have been reported to interact with the proline-rich region of nsP3 HVD to enhance virus replication (163, 170). Furthermore, the repeat FGDF motifs of HVD have been well-characterised in their interaction with host cell protein G3BP, which is involved in stress granule formation to sequester translational factors (170-174). Structural and biochemical analysis of

nsP3 has found that the FGDF motifs directly bind to the nuclear-transport factor 2 (NTF2)-like domain of G3BP, which is disabled from forming stress granules (175). Further investigation revealed the sequential binding pattern that the N-terminal FGDF motif of HVD binds to G3BP first, followed by the accumulation of nsP3 and G3BP to allow the dimeric interaction with the C-terminal FDGF motif (176). This interaction has been shown to be absolutely required for CHIKV replication in vertebrate cells, shedding light on how viruses hijack and evade the host antiviral system to facilitate their lifecycle (84).

1.2.3.4 Non-structural protein 4

The alphavirus RdRp is encoded by nsP4 and consists of the classic polymerase fingers, thumb and palm domains, the latter accommodates the GDD catalytic triad (177). Biochemical analysis has demonstrated that nsP4 binds directly to the AU-rich region of the 3' CSE4 of virus genomic RNA, and that this interaction is independent of other nsPs or the 5' UTR (178). Both genomic and sub-genomic RNA of alphaviruses are synthesised by nsP4, which uses distinct and independent binding sites for their synthesis, respectively: residues 531-538 for genomic RNA and 329-334 for sub-genomic RNA (37, 179, 180). Interestingly, nsP4 has recently been suggested to predominantly determine the alphavirus replicase activity and selection of template RNA, as the replication activity solely depends on the expression level of nsP4, while nsP4 of several alphaviruses are functionally compatible with heterologous P123 (181). In addition to *de novo* RNA synthesis, the catalytic triad of nsP4 has also shown terminal adenylyltransferase (TATase) activity, which specifically regulates the Mg²⁺-dependent addition of adenine to the 3' end of nascent RNA (182). Loss of function mutation of nsP4 GDD>GAA abolishes both RdRp and TATase activity (182, 183). The ~100 N-terminal of nsP4 is a partially unstructured domain essential for the function of nsP4 (110). The definitely conserved tyrosine residue within the N-terminal of nsP4 has been shown to interact with nsP1 for negative-sense strand synthesis (184). Such direct and indispensable interactions between nsP4 with nsP1 and nsP2, but not nsP3, lead to the proposed synergetic mechanism of alphavirus replication that nsP4 binds

to the viral RNA to synthesise progeny single-stranded RNA, which may then be capped by MTase/GTPase of nsP1 and NTPase of nsP2 described in 1.2.3.1 and 1.2.3.2 (185).

1.2.4 Structural proteins

The five structural proteins of CHIKV are expressed from ORF-2 of the sub-genomic RNA (Figure 1.2.4). The precursor polyprotein containing capsid, p62, 6K/TransFrame (TF) and E1 is further processed to capsid, E3, E2, 6K/TF and E1. The structure and organisation of these structural proteins have been resolved by X-ray crystallography (186, 187).

ORF-2

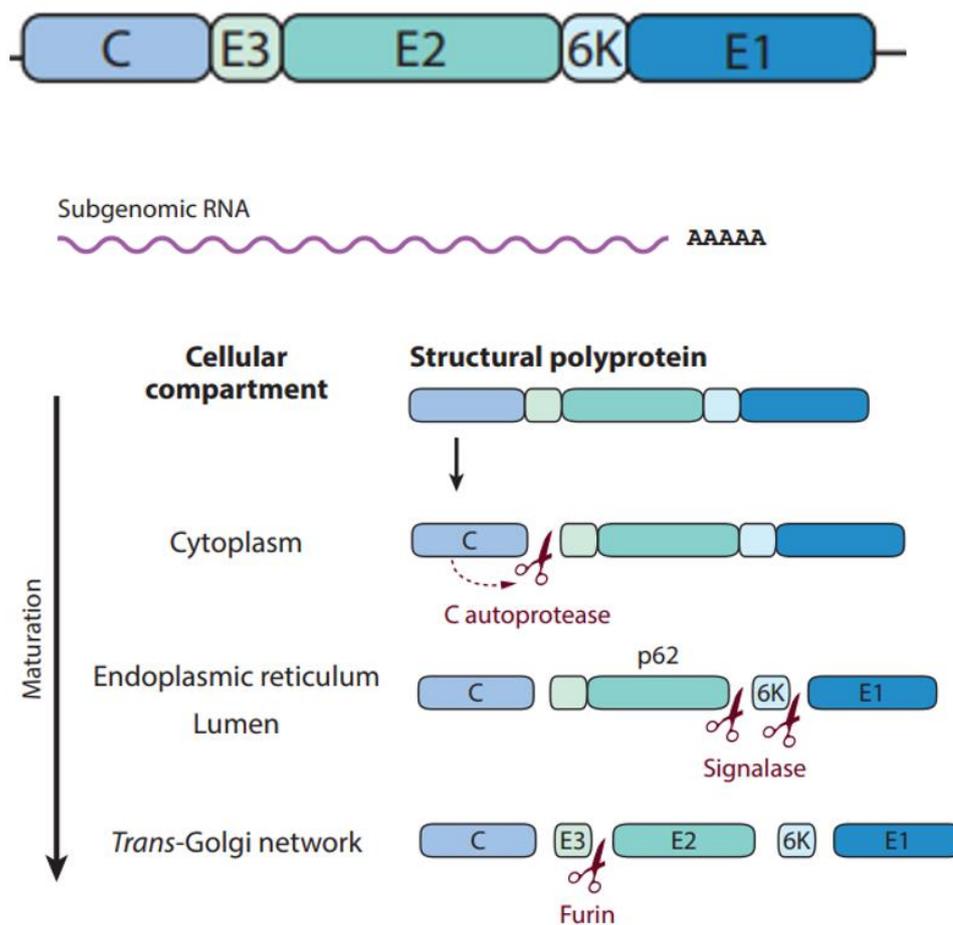


Figure 1.2.4. The genomic organisation and processing of alphavirus ORF-2 containing the structural proteins (188).

The 1244 residues structural precursor protein is translated from the sub-genomic RNA. The *cis*-acting protease of the capsid protein cleaves itself from the polyprotein complex. Upon maturation inside the endoplasmic reticulum, the polyprotein is further cleaved into p62, which is further processed into E3 and E2, and 6K and E1 proteins by cellular furin and signalase.

1.2.4.1 Capsid protein

The CHIKV capsid protein (CP) constitutes the structural subunits of the nucleocapsid by forming pentameric and hexameric capsomeres. The intrinsically disordered N-terminal domain is highly enriched with positively charged residues like arginine and lysine (189, 190). The N-terminal of this ~30 kDa protein contains two nuclear localisation signals between residues 60-77 and 84-99, respectively (191). It has been suggested that despite its cytoplasmic presence, CP is actively shuttling between the host nucleus and cytoplasm via a nuclear export signal between residues 143-155, which interacts with the export receptor chromosomal region maintenance 1 (CRM1) (191). Through residues 81-105, which are structured to coiled-coil α -helix, the N-terminal domain of CP also binds to genomic RNA to form the RNA-capsid complex required for nucleocapsid assembly in the cytoplasm (191, 192). This helix region has been shown to function as a checkpoint control during virus core assembly, as heterologous CP from different alphaviruses depends on the compatibility of their helix regions to form phenotypically mixed capsid-like particles (193). The C-terminal of CP possesses a serine protease activity which *cis*-cleaves itself from the nascent structural precursor polyprotein (Figure 1.2.4) (194). Site-directed mutagenesis of the C-terminal of SINV CP revealed several conserved residues, including the autoproteolytic triad His¹³⁹, Asp¹⁶¹ and Ser²¹³, which regulate the cleavage at Trp²⁶¹ (195).

1.2.4.2 Envelope proteins 1-3

The envelope proteins E3 and E2 are processed from the precursor protein termed p62 (Figure 1.2.4). Glycosylated p62 is processed and matured at the endoplasmic reticulum (ER) of infected cells, where the furins and signalases from host cells cleave p62 to render E3 and E2 (196, 197). The translocation signal attributed to the polar residues within the N-terminal of E3 targets p62 to ER for subsequent processing (198). The furin loop at the E3-E2 junction prevents the premature cleavage of the p62-E1 heterodimer, whose subsequent release, in turn, leads to the increased disorder of the furin loop (187). E3 is an α/β glycoprotein with its β -hairpin enclosed with three α -helices, forming a horseshoe-shaped structure (187). E3 participates in stabilising the structure of E2 by conformational modifications so that the fusion loop can be inserted (187). It is also essential for the stabilisation and protection of the p62-E1 complex from a low pH cellular environment, as well as facilitating the endocytosis of E2-E1 spikes (199).

Both E1 and E2 proteins are glycosylated and palmitoylated (200). The E2 glycoprotein is involved in the recognition of cellular receptors and is a transmembrane protein with three domains (A, B and C) (198). The first two domains regulate the conformational creation of a groove, where a β -hairpin is inserted by the premature fusion loop of E1 to form heterodimers representing the protruding spikes on the envelope of alphavirus virion (187). Domain B alone is responsible for the β ribbon-dependent interaction with E3 (199). Asparagine residue at position 263 and 273 are involved in the N-linked glycosylation of the E2 protein (199). Conserved cysteine residues of E2 have been shown to be related to proper virion assembly, as mutations led to defects in polyprotein processing, fusion efficiency and virion morphology (201). Functional analysis mapped to a short linear signal in the cytoplasmic domain of E2 regulating virus budding by interacting with the nucleocapsid (202). On the other hand, the E1 glycoprotein possesses hydrophobic regions involved in the fusion of endosomal and viral membranes (203). Three domains (I, II and III) can be found in the alphavirus E1 protein, with the second one containing a fusion loop necessary for E1-E2 interaction

(204). In terms of the virion morphology, the heterodimers of E1-E2 make up the spikes, which mediate the low pH-induced endocytic fusion for virus entry (199). Recently, the cryptic epitopes of E1 have been reported to dwell in the N-terminal fusion loop in a pH-dependent manner, and E1-specific monoclonal antibodies effectively inhibit virus egress *in vitro* (205, 206). Similar neutralising antibodies have also been identified targeting the acid-sensitive region of E2, further meriting the development of antibody-based anti-CHIKV strategies (207).

1.2.4.3 6K-TF protein

The 6K alphavirus protein is a small, hydrophobic acylated protein of 61 amino acids in length (Figure 1.2.4). It is involved in membrane permeabilisation to facilitate progeny virus budding by inducing the translocation of lipids between the two sides of biological membranes (208). The 6K protein is also considered to constitute a cation-selective ion channel presumably via membrane oligomerisation (209). Genetic manipulations of the 6K proteins have been shown to result in the accumulation of cytoplasmic nucleocapsids due to altered membrane fusion capacity, as well as the instability of spike-dependent virion structure (210, 211). The dual function of the 6K gene was not recognised until the discovery of the TransFrame protein (TF) as a result of frameshifting (212). This ~8 kDa protein is expressed with an approximate efficiency of ~18% by ribosomal-1 frameshifting at a conserved U-rich motif within the 6K gene (212). Comparative analysis showed prominent diversity in the structure of genomic RNA 3' end which stimulates -1 frameshifting (213).

1.3 CHIKV lifecycle

CHIKV is introduced to the host through the bite of infected mosquitos. Inoculation of the virus at the bite site causes the infection of nearby cells, such as skin fibroblasts and dermal macrophages (52, 214, 215). The lifecycle of CHIKV is initiated by the recognition of receptors at the target cell surface. CHIKV cellular receptors include the C-type lectins dendritic cell-specific intercellular adhesion molecule-3-grabbing non-integrin (DC-SIGN) and lymph node-SIGN (L-SIGN) in dendritic cells, laminin receptor in mammalian cells and the metal ion transporter natural resistance-associated macrophage protein (NRAMP) in both mammalian and insect cells (216-218). CHIKV then enters the host cell by clathrin-mediated endocytosis (Figure 1.3.1). Inside the endosome, the E1-E2 heterodimers undergo extensive conformational changes induced by the low pH microenvironment, exposing the E1 fusion loop responsible for membrane fusion (4, 219). The trimers of E1-E2 are then reorganised to form E1 trimers, unmasking its hydrophobic domain to become fusion active (220). Distortion of the cellular membrane begins with the synergetic interactions between E1 trimers, which concomitantly distort the virus envelope by pulling the transmembrane regions against each other (221). Such membrane deformations in the host cell and virus envelope lead to the fusion of the outer layer of the phospholipid, followed by the mixing of the inner layer to complete membrane fusion (4). Direct delivery of the nucleocapsid to the cytoplasm has also been detected, where the virion penetrates the plasma membrane directly via a membranous pore formed by undefined virus and host proteins (222).

Once the nucleocapsid is released into the cytosol, the viral genomic RNA is rapidly liberated by the large cellular ribosomal subunits, which disrupt the structure of the nucleocapsid (223, 224). The viral genomic RNA has been suggested to contain direct binding sites to the capsid protein, which is proposed to increase RNA stability and promote translation (225). The first event after genome release is the canonical cap-dependent translation of ORF-1 of the genomic RNA from a single AUG initiation codon,

producing the non-structural polyproteins P123 and P1234 (Figure 1.3.1). The critical feature for producing one or the other of the two polyproteins is the UGA stop codon near the end of the nsP3 coding region. Mutational analysis within and next to this stop codon has reported that the single cysteine residue immediately following the UGA is critical for its translational readthrough, which in ~10% cases produces P1234 (226). The *cis*-acting protease activity of nsP2 rapidly catalyses the cleavage at the 3/4 junction, releasing the active nsP4 with RdRp activity. Individual nsP4 has been indicated to be highly unstable and must be protected by the virus replication complex (vRC) (137). vRC is the nsP3-activated, host cell protein-containing, functional P123 + nsP4 replicase required for the initiation and elongation of negative-sense strand synthesis (227). The palmitoylated nsP1 associates with the cellular membrane by an amphipathic helix within its MTase/GTase domain, which regulates the attachment of vRC to the membranes (227-229). Overlapping functions of non-structural proteins have been suggested, with P123 + nsP4 regulating negative-sense strand synthesis as well as the formation of membrane invagination termed spherules (230). The location of the spherule is host-specific, with plasma membrane-associated in mammalian cells and cytopathic vacuole-associated in insect cells (231). The formation of spherules can be induced by P123 + nsP4 in the absence of genomic template RNA, while nsP1 coupled with uncleaved P23 and active nsP4 is also capable of forming replication-competent but morphologically irregular spherules (232). Colocalisation of the dsRNA intermediate with spherules have been demonstrated as a protective mechanism against the cellular immune response (233, 234). During the synthesis of the negative-sense strand, P123 + nsP4 interacts with the 3' end promoter sequence of genomic RNA, as well as *cis*-acting elements within the 5' UTR to presumably induce genome circularisation with host cell factors (85). This early phase of alphavirus replication typically takes 3 to 4 hours, producing negative-sense RNA to act as the template for synthesising genomic RNA and sub-genomic RNA, which is transcribed from the promoter at the intragenic region and harbours the genes for structural proteins (Figure 1.3.1) (235).

As the replication cycle progresses, ORF-1 translation leads to the accumulation of the non-structural polyprotein P123, which is rapidly *trans*-cleaved into nsP1, nsP2 and nsP3 (Figure 1.3.1). This cleavage process is crucial for switching the preference from negative-sense to positive-sense strand synthesis of genomic and sub-genomic RNA (236). The protease activity of nsP2 must sequentially *cis*-cleave the 1/2 and 3/4 junction at specifically conserved residues, while the *trans*-cleavage of 2/3 junction requires the release of nsP1 to expose the N-terminal active site of nsP2 (134, 139, 237). The production of individual nsPs irreversibly terminates the short-lived synthesis of negative-sense strand RNA, suggesting the subtle temporal regulation of the viral replicase (238). Synthesis of positive-sense genomic and sub-genomic RNA requires the recognition of the promoter sequence by two independent RNA binding sites nsP4 (179, 180). Cleaved nsP1 is solely responsible for the membrane association of vRC, while nsP1 and nsP3 are both required to guide the translocation of spherules from the plasma membrane to endosomal compartments within the cytosol (229). Notably, although nsP4 possess the RdRp activity for RNA synthesis, mutations in nsP2 and nsP3 impair the synthesis of both genomic and sub-genomic RNA, indicating their multifunctional role as co-transcription factors to nsP4 (157, 159, 239, 240). In addition, nsP3 has been shown to block the formation of stress granules by sequestering its structural subunit G3BP in mammalian cells and Rasputin in insect cells (241).

During the late stages of the CHIKV infection cycle, abundant sub-genomic RNA replicated from the negative-sense strand leads to ORF-2 translation. CP is the first to be produced from sub-genomic RNA translation. After its autocatalytic cleavage of the structural precursor protein, CP associates with the nascent genomic RNA to form the nucleocapsid in the cytoplasm (82). The exposure of the signal peptide within the N-terminal of p62 then interacts with the ER membrane to direct the polyprotein to its lumen, where cellular furin and signalase process the polyprotein into E3, E2, 6K and E1 (196). In ~10% cases, the UUUUUUA sequence within the 6K gene causes ribosomal frameshift to produce a single TF protein instead of 6K and E1 (212, 213). The 6K protein has been implicated in the cytoplasmic trafficking of virus glycoproteins from

the vesicular system to the plasma membrane, while the TF protein is dispensable for virus replication but is involved in virion assembly (242-244). Cleaved E2 and E1 heterodimerise and incorporate into the plasma membrane with their C-terminal cytoplasmic tail, which facilitates virus budding by interacting with the *de novo* nucleocapsids (245). A portion of the host lipid bilayer is acquired by virion during budding. In mosquito cells, virus budding has been detected in the lumen of cytopathic vacuoles, which promotes dissemination efficiency and immune evasion via cell-to-cell spread using the host secretory pathway (246). Following viraemia, CHIKV can further spread to the spleen, lymph nodes, muscles and liver (53, 71, 79).

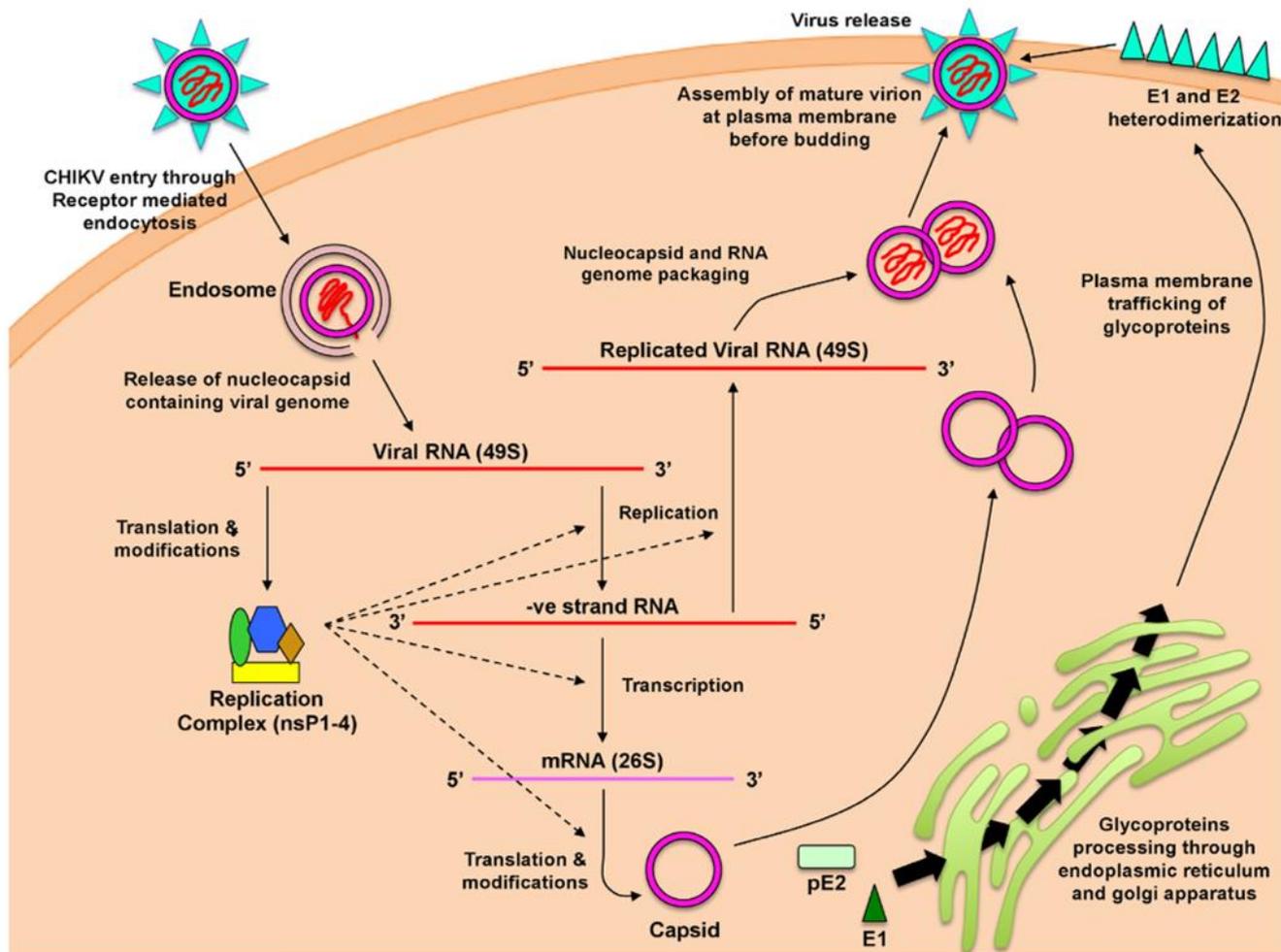


Figure 1.3.1. Schematic illustration of the CHIKV replication cycle (247).

The virus enters susceptible host cells by associating with cellular receptors, triggering clathrin-mediated endocytosis. The acidic environment within the endosome leads to conformational modification of the CHIKV envelope proteins, which regulate membrane fusion with the host membrane, causing the nucleocapsid to be released into the cytoplasm. The viral genomic RNA is unmasked by the large ribosomal subunit, followed by the translation of the four non-structural proteins. This virus replicase P123 + nsP4 and nsP1-4 is involved in the synthesis of negative-sense genomic RNA, the synthesis of nascent positive-sense genomic RNA (49S), as well as the transcription of the sub-genomic RNA (26S), which encodes the five structural proteins of ORF-2. Translation of the sub-genomic RNA produces the structural precursor protein C-p62-6K-E1. The autolytic protease of the capsid protein cleaves itself from the polyprotein, which is further post-translationally modified in the ER and Golgi apparatus. E1 glycoproteins heterodimerise with E2 and are translocated to the plasma membrane, where they are subsequently incorporated into the nascent virion as trimeric spikes. The cytoplasmic capsid protein packages the nascent genomic RNA to assemble into the icosahedral nucleocapsid. CHIKV virion is then released while obtaining a lipid bilayer from the host cell plasma membrane.

1.4 Preliminary work deciphering the structure of the 5' end of the CHIKV genome

1.4.1 The dynamic stem-loops within the 5' end of the CHIKV genome

Arthropod-borne viruses such as CHIKV have to evolutionarily adapt to both vertebrate and invertebrate cellular environments. Previous work from our group, using selective 2' hydroxyl acylation analysed by primer extension (SHAPE) constrained thermodynamic mapping and reverse genetic analysis, demonstrated that nt 1-330 at the 5' end of the CHIKV genome encode 5 RNA stem-loops (SL47 –SL246) and a pseudoknot (PK) interaction that base pairs with the ORF-1 AUG start codon (Figure 1.4.1) (83). Synonymous substitutions disrupting each stem-loop in turn inhibited virus replication in a host-specific manner. SL47, which is highly conserved among the Old World alphaviruses, facilitated genome replication in both human- and mosquito-derived cells; SL85-246 were required for host-dependent genome replication - SL165 and SL194 (CSE2) were exclusively required in mammalian cells, whereas SL246 exclusively enhanced replication in mosquito cells (83). All of the stem-loops functioned as RNA replication elements (RREs) predominantly in a structure-dependent manner, with no known roles in the regulation of virus translation. Notably, the function of SL194 requires both the structural integrity of the stem-loop and the sequence conservation of the single-stranded terminal loop, as a single C>U mutation within the terminal loop region severely impaired CHIKV replication (83). In addition, inconsistency between SHAPE mapping and structural prediction of SL85 was hypothesised to indicate that this region might be able to dissociate the base-pairing of the stem to form structurally dynamic interactions with other parts within the CHIKV 5' end, such as the adjacent AUG start codon (83). These RREs have been proposed to be involved in the viral lifecycle via interacting with host-specific factors (83). Yet, the mechanism of how they interact or through which host proteins remains obscure.

Despite the fact that the initiation of genome replication for positive-sense RNA

viruses takes place at the 3' end of the genome, the RREs within the 5' end also play critical roles in this process (248, 249). For instance, mutagenesis of the 51nt CSE, which is constituted by SL165 and SL194 in CHIKV, severely inhibited virus replication in both mammalian and insect cell lines (90, 91). Genome circularisation resulting from stabilised long-range RNA-RNA interactions in dengue virus and poliovirus provided insights into the replication process that RdRp binds to the 5' region initially and then translocates to the 3' promoter 'guided' by the circularisation (250, 251). It has also been speculated that such interactions may regulate template specificity as well as serve as a translation-transcription switch, which induces the dissociation of the ribosome from the template to initiate replication (85). Furthermore, in mosquito-borne flaviviruses which are also single-stranded, positive-sense RNA viruses, the complementary interactions between the 5' and 3' sequence motifs, termed downstream of AUG region, have been demonstrated to be crucial for virus replication (252, 253). Given the indispensable functional importance of the 5' region, this study sought to identify mammalian host cell proteins essential for CHIKV replication and analyse their interaction with the 5' CHIKV genome to unveil the mechanism.

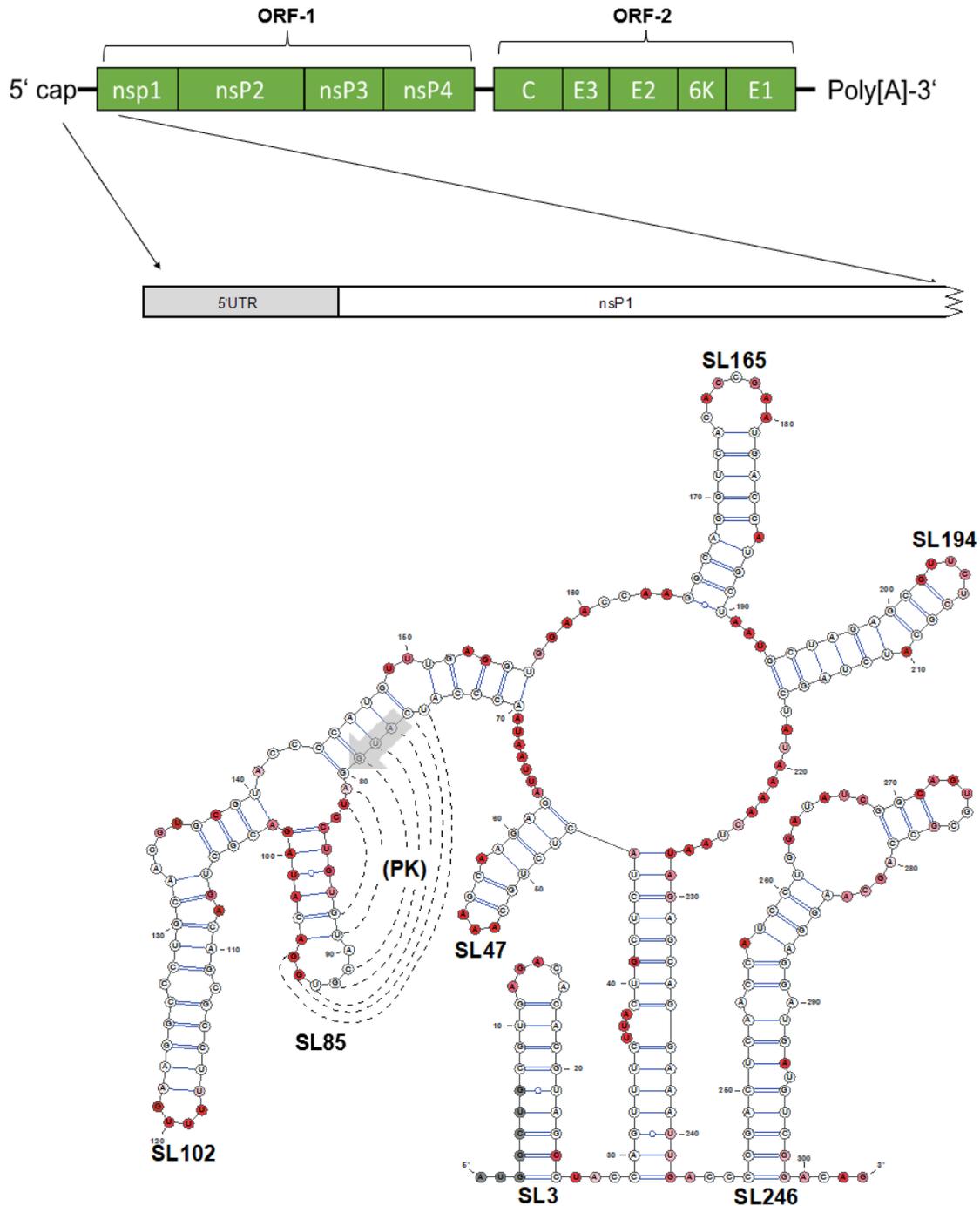


Figure 1.4.1. Schematic representation of selective 2'-hydroxyl acylation analysed by primer extension (SHAPE)-constrained thermodynamically derived model of CHIKV genomic RNA folding (nucleotides 1-303) (83).

The AUG start codon of ORF-1 was marked by the grey arrow. The colour of individual nucleotides indicates their corresponding SHAPE reactivity. High reactivity (red)

indicates unpaired nucleotides while low reactivity (white) indicates base-paired nucleotides. The stem-loops were designated SL3, SL47, SL85, SL102, SL165, SL194 and SL246. Pseudoknot = PK.

1.4.2 Identification of MSI2 as a binding partner to the 5' end of the CHIKV genome

Using RNA affinity purification, coupled with quantitative tandem mass tagging and MALDI-TOF quantitative mass spectrometry (TMT-MS) from CHIKV-infected human rhabdomyosarcoma (RD) cell extract, our group previously identified a list of viral and cellular proteins interacting with the SL47-SL246 bait RNA (Figure 1.4.1). Two CHIKV non-structural proteins, nsP1 and nsP3, and two cellular proteins, DExH-box RNA helicase 9 (DHX9) and Musashi homolog 2 (MSI2) were significantly enriched ($p < 0.05$) relative to bead and non-specific RNA controls. The involvement of both nsP1 and nsP3 in CHIKV replication has been detailed in 1.2.3 above. This study focused on the role of MSI2 in CHIKV replication as it has not been previously identified, while DHX9 has been shown to promote CHIKV nsPs translation whilst simultaneously inhibiting genome replication (106).

The Musashi protein family play an essential role in regulating gene expression of neurones during central nervous system maturation (254). Defective MSI expression has been shown to be related to aberrant proliferation and reduced stability of the stem cells in the central nervous system (255). The Musashi family include several RNA binding proteins including MSI1, MSI2, TAR DNA binding protein (TARDBP), DAZ associated protein 1 (DAZAP1) and heterogeneous nuclear ribonucleoprotein (hnRNP) A/B/D/D-like (256). Their overexpression is frequently associated with the development of several tumour types. MSI2 and its paralogs, which are genes resulting from duplication events, contain two consensus RNA-recognition motifs (RRM): RRM1 preferentially binds to (G/A) U₁₋₃AGU, which can be found at a single-stranded region (nt 63-67) within the CHIKV 5' UTR (Figure 1.4.1); RRM2 promotes binding affinity by

binding to the same motif (257) (Figure 1.4.2). MSI1 and MSI2 are evolutionarily conserved and share 75% homology, which occasionally results in functional redundancy (258, 259). MSI2 is ubiquitously cytoplasmic expressed and is involved in the regulation of IL-6 signalling, whose downstream cascades include PI3K, MAPK and JAK/STAT, all of which contribute to important cellular processes (258). Progressive tumorigenesis of oncogenic cells induced by MSI2-mediated pathways has also been well-documented (260, 261). The genes of the Musashi family have been found to be frequently associated with post-transcriptional modifications such as alternative splicing, which results in the formation of isoforms differing in molecular weight and RNA binding affinity (256). For MSI2, several RRM-containing isoforms have been recorded on UniprotKB, suggesting that the overlapping functions of MSI2 also might exist among its isoforms (262).

So far, limited studies have been performed on the effect of MSI on positive-sense RNA viruses. One Zika virus study has shown that MSI1 binds to the three consensus binding sites in the single-stranded region of the 3' UTR of the viral genome to facilitate replication (263). Virus-induced interference of the interaction between MSI1 and its original targets may lead to direct pathogenic consequences in the neurodevelopmental process. Importantly, different N-terminal isoforms of MSI1 arising from alternative splicing could potentially impact the permissiveness of Zika virus replication, due to the altered affinity and kinetics of MSI1-Zika interaction (256). The significance of Musashi proteins in virus replication has been emphasised that UAG sequence motif, where the RRMs of Musashi proteins predominantly binds, was highly accessible and conserved in mosquito-borne flaviviruses such as Zika and dengue virus (264). Furthermore, the difference in binding affinity and kinetics resulting from post-transcriptional modifications of MSI isoforms could potentially lead to a facilitative or detrimental effect on virus replication (256).



Figure 1.4.2. The schematic representation of the MSI2 protein (258).

MSI2 is a ~35 kDa RNA-binding protein ubiquitously expressed in the cytoplasm. It contains two RRM, the first one preferentially binds to (G/A) U1-3AGU, while the second one facilitates the binding affinity.

1.5 Aim of This Study

There is currently no CHIKV vaccine or specific antiviral therapy, due in part to a lack of detailed mechanism of CHIKV replication and interactions with host cell factors. A deeper understanding of how host cell proteins is essential for developing antiviral treatments. Preliminary studies in our group using RNA affinity purification and TMS-MS identified MSI2 was significantly enriched as a binding partner to the CHIKV 5' nt 1-303. This implied a potential pro-viral role of MSI2 in CHIKV genome replication. Therefore, the aim of this study is to investigate the importance and mechanism of MSI2 during the CHKV lifecycle in human cells. The specific objectives of this study are:

1. Validate the role of MSI2 during CHIKV replication.
2. Confirm which stage of the CHIKV replication cycle is MSI2-dependent.
3. Identify potential MSI2 RNA binding site within the CHIKV genome.

CHAPTER 2: MATERIALS & METHODS

2.1 Materials

2.1.1 Human cell culture

Human Rhabdomyosarcoma (RD), human hepatoma liver (Huh7) and Baby Hamster Kidney (BHK) cells (Table 1) were grown in Dulbecco's modified eagle's medium (DMEM; Sigma) containing 10% fetal bovine serum (FBS; PAA), 1x penicillin-streptomycin (Sigma), 25mM HEPES in 0.85% NaCl (Lonza) and 1% non-essential amino acids mixture (NEAA; Lonza). Both cell lines were plated as monolayers in 75 cm² cell culture flasks at 37°C/5% CO₂. Trypsin-EDTA solution (Sigma) was used to split and passage the cell lines.

Cell line	Cell type	Origin species
Huh7	Liver (Hepatocellular carcinoma)	Human
RD	Muscle (Rhabdomyosarcoma)	Human
BHK-21	Kidney (fibroblast)	Hamster

Table 1. The name, type and species of cells used throughout this study.

2.1.2 Infectious virus construct

The wild-type CHIKV infectious clone designated ICRES used throughout this study were derived from the LR2006_OPY1 La Réunion island isolate of the ECSA lineage (265). CHIKV MSI2 BSM was customised and synthesised by Azenta Life Sciences. Figure 2.1 shows a plasmid map of the CHIKV full-length ICRES vector.

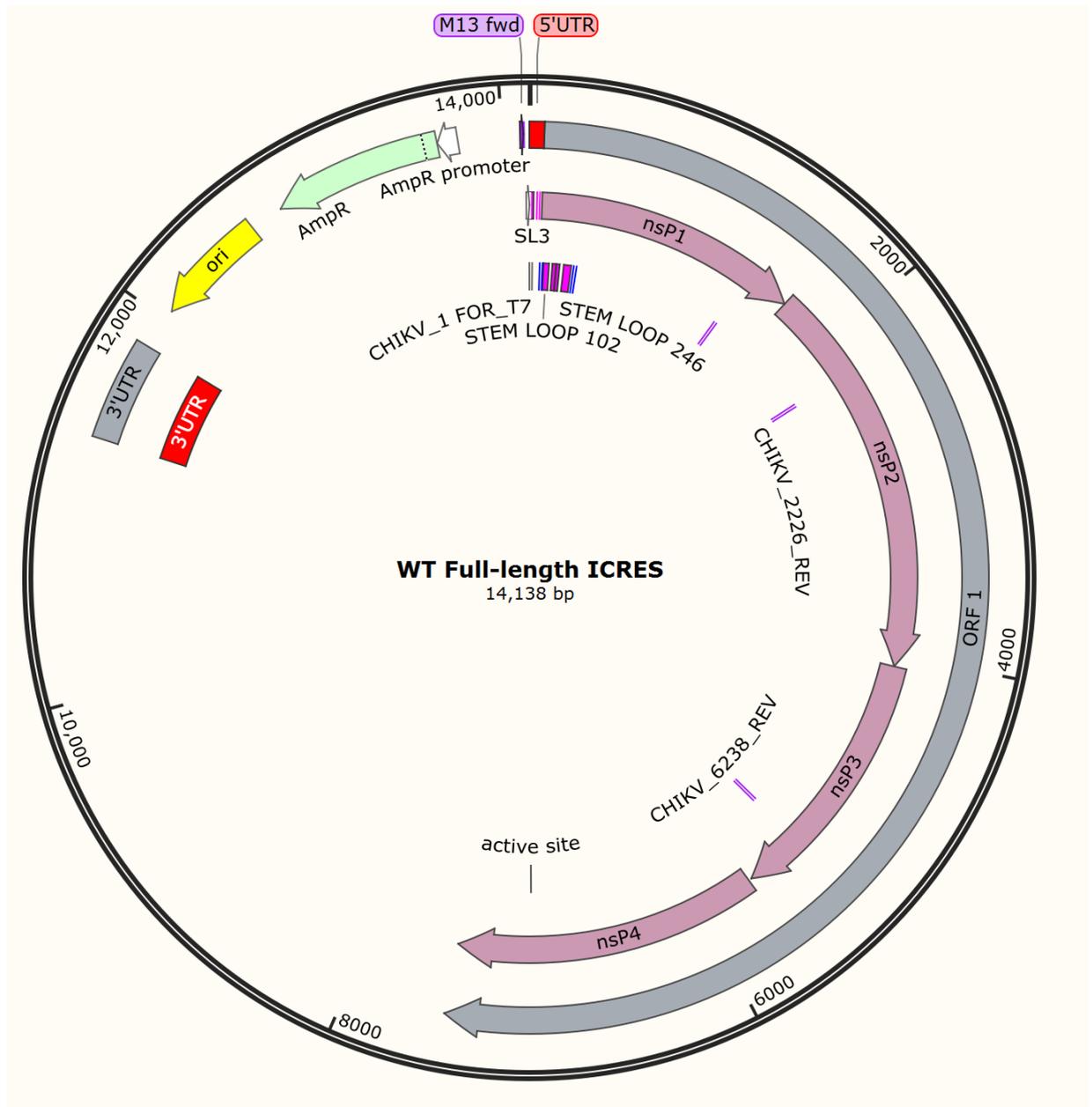


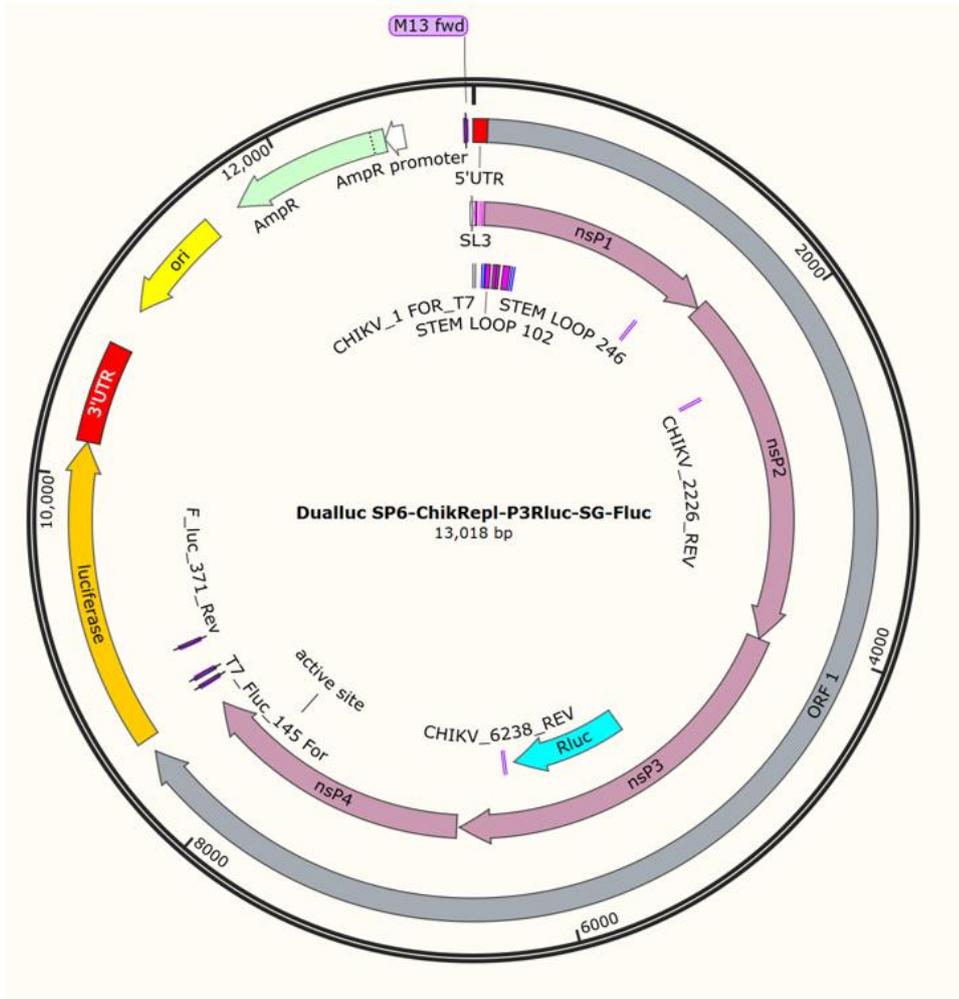
Figure 2.1. The representative gene map of infectious virus construct WT Full-length ICRES.

2.1.3 Sub-genomic replicon constructs

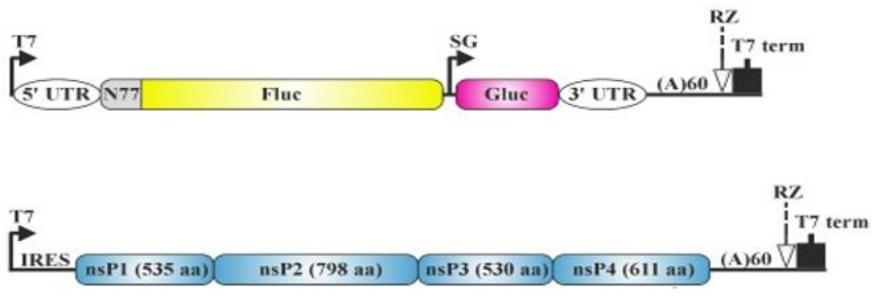
Sub-genomic replicon studies were carried out utilising a sub-genomic replicon construct derived from the CHIKV ECSA genotype. The replicon 'SP6-ChikRepl-P3Rluc-SG-Fluc' was simplified as CHIKV dual-luciferase replicon, which contains an SP6 RNA polymerase binding site, a *Renilla* luciferase gene inserted between nsP3 and nsP4, and a firefly luciferase gene replacing the entire ORF-2. CMV-MSI2 BSM-Fluc-Gluc was customised and synthesised by Azenta Life Sciences. Figure 2.2A shows the plasmid map of the CHIKV dual-luciferase replicon.

The sub-genomic replicon constructs for trans-complementation assay were kindly provided by Prof Andres Merits, University of Tartu, Estonia. CMV-Fluc-Gluc contains a firefly and gaussian luciferase gene replacing ORF-1 and ORF-2, respectively, it is therefore simplified as WT reporter. CMV-P1234-CHIKV only contains CHIKV nsPs, it is therefore simplified as CHIKV polymerase. Figure 2.2B shows the genomic structure of the constructs.

A)



B)



**CMV-Fluc-Gluc
WT Reporter**

**CMV-P1234-CHIKV
CHIKV Polymerase**

Figure 2.2. Replicon constructs.

A). The representative gene map of SP6-ChikRepl-P3Rluc-SG-Fluc, simplified as the dual-luciferase replicon (265). **B).** Schematic representation of the two components of the trans-complementation assay. The upper construct CMV-Fluc-Gluc, has a firefly luciferase gene replacing the ORF-1 and a Gaussia luciferase gene replacing ORF-2, it is therefore termed WT reporter; the lower construct CMV-P1234-CHIKV, only expresses CHIKV nsPs, and is therefore termed CHIKV polymerase (140).

2.1.4 Human Musashi homolog 2 RNA recognition motifs Escherichia coli expression vector

pET-22HT-MSI2(8-193) expressing the two RNA recognition motifs of human MSI2 in bacterial cells for protein purification was a gift from Sean Ryder (Addgene plasmid # 60356; <http://n2t.net/addgene:60356>; RRID: Addgene_60356). Figure 2.3 shows the plasmid map of this vector.

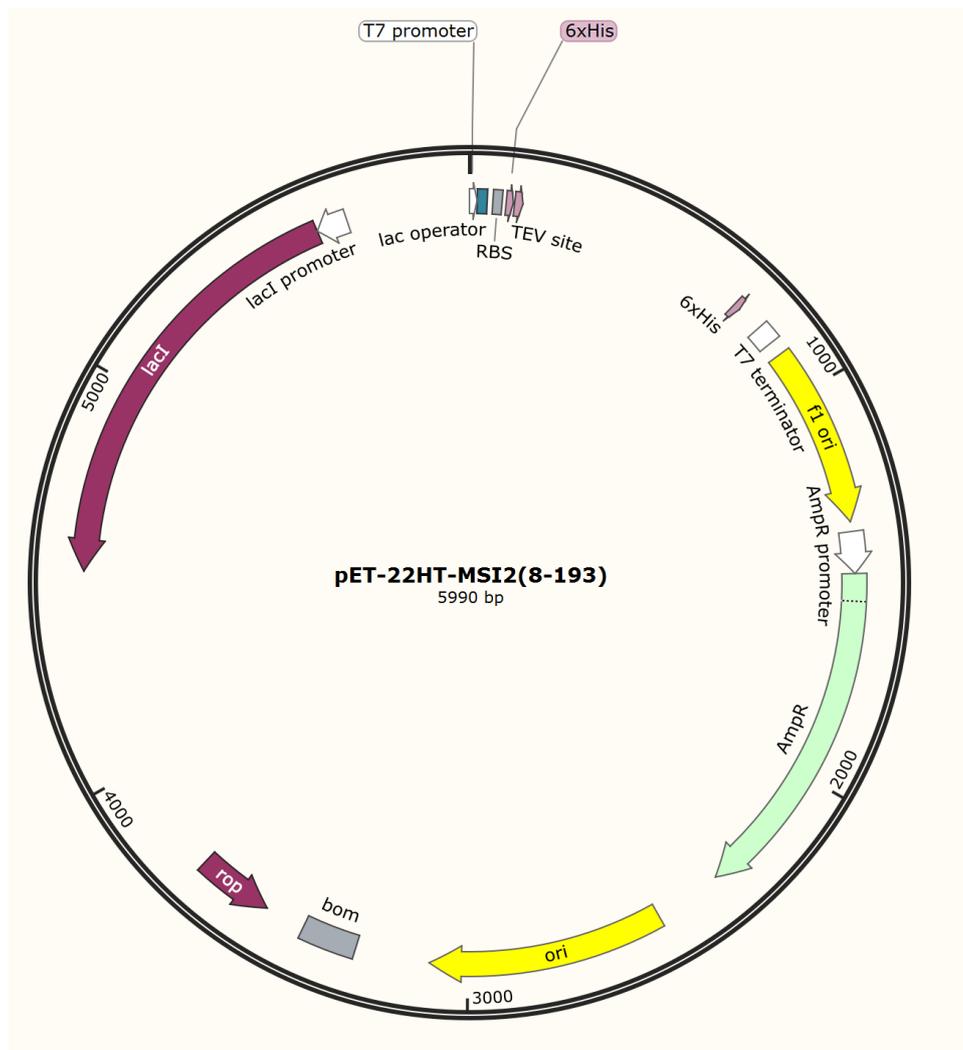


Figure 2.3. The representative gene map of pET-22HT-MSI2(8-193) expressing the two RNA recognition motifs of human MSI2 (266).

2.1.5 Critical reagents, kits, antibodies and primers

Reagents, kits and antibodies which are critical to this study are summarised in Table

2. Detail about primers can be found in the method section.

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-mouse monoclonal β -Actin	Sigma-Aldrich	Cat# A1978
Anti-rabbit monoclonal MSI2	Abcam	Cat# ab76148
Anti-mouse monoclonal MSI1	Abcam	Cat# ab129819
Anti-rabbit polyclonal nsP1	This paper	N/A
Anti-rabbit polyclonal nsP3	This paper	N/A
Anti-rabbit polyclonal capsid	This paper	N/A
Donkey anti-rabbit IRDye [®] 680 LT	LICOR	Cat# 926-68023
Donkey anti-mouse IRDye [®] 800 CW	LICOR	Cat# 926-32212
Bacterial and virus strains		
CHIKV ICRES cDNA clone	Isolate LR2006 OPY1. East Central South African genotype	(265)
CHIKV SP6-ChikRepl-P3Rluc-SG-Fluc	Andres Merits	(265)
BL21(DE3) Competent E. coli	New England Biolabs	Cat# C2527H
<i>Mix & Go!</i> DH5 α Competent Cells	Zymo Research	Cat# T3007
CMV-P1234-CHIKV	Andres Merits	(140)
CMV-P1234 ^{GAA} -CHIKV	Andres Merits	(140)
CMV-Fluc-Gluc	Andres Merits	(140)
CMV-MSI2 BSM-Fluc-Gluc	This paper	N/A
CHIKV MSI2 BSM	This paper	N/A
Chemicals, peptides, and recombinant proteins		
Human MSI1 siRNA	Santa Cruz Biotechnology	Cat# sc-106836
Human MSI1 shRNA	Santa Cruz Biotechnology	Cat# sc-106836-SH
Human MSI2 siRNA	Santa Cruz Biotechnology	Cat# sc-75834

Human MSI2 shRNA	Santa Cruz Biotechnology	Cat# sc-75834-SH
Ro 08-2750	TOCRIS	Cat# 2272
ATP, [γ -32P]	Perkin Elmer	Cat# NEG502A250UC
pET-22HT-MSI2 (8-193)	Sean Ryder	Cat# 60356; http://n2t.net/addgene:60356 ; RRID: Addgene_60356 (266)
Dynabeads™ Protein G for Immunoprecipitation	Thermo Fisher Scientific	Cat# 10003D
Critical commercial assays		
RNA Clean & Concentrator	Zymo Research	Cat# R1013
Dual-Luciferase® Reporter Assay System	Promega	Cat# E1910
HisTrap™ FF	Cytiva	Cat# 17531901
PD-10 Desalting Column, with Sephadex G-25 resin	Cytiva	Cat# 17085101
SCRIPT cDNA Synthesis Kit	Jena Bioscience	Cat# PCR-511S
2x qPCR BIO SyGreen Blue Mix Lo-ROX	PCR Biosystems	Cat# PB20.15-05

Table 2. Detailed list of critical reagents, constructs and commercial assays used in this study.

2.2 Methods

2.2.1 Agarose gel electrophoresis

The size and purity of plasmid DNA were determined by agarose gel electrophoresis, carried out using 1% agarose gels (1% analytical grade agarose (Sigma Aldrich), 1x TAE buffer (40mM Tris-Acetate, 1mM EDTA), 1:10,000 dilution SYBR SAFE DNA stain (Life Technologies)). Samples were mixed with 6x Purple Gel Loading Dye (ThermoScientific) and loaded alongside 1KB Plus DNA ladder (ThermoScientific). Gels were run in 1x TAE buffer at 80V for 60 minutes before visualisation under UV transillumination.

2.2.2 Denaturing MOPS agarose gel electrophoresis

The purity, integrity and size of RNA products of *in vitro* transcription were determined by denaturing formaldehyde agarose gel electrophoresis in MOPS buffer. 1% agarose were prepared using 1% analytical grade agarose (Sigma Aldrich), 10x MOPS buffer (0.4M MOPS pH 7.0, 0.1M NaOAc, 0.01M EDTA pH 8.0), ddH₂O and 1:10,000 dilution SYBR SAFE DNA stain (Life Technologies). After melting the agarose, formaldehyde was added to a final concentration of 1.7% v/v. Samples were mixed with 2x denaturing RNA loading dye (ThermoScientific), heated to 65°C for 10 minutes and cooled on ice. Samples were loaded alongside Millennium RNA ladder (Ambion). Gels were run in 1x MOPS buffer at 100V for 60 minutes before visualisation under UV transillumination.

2.2.3 Plasmid RNA linearisation

10 µg of plasmid DNA was incubated with 10x CutSmart buffer (New England Biolabs), 5U NotI-HF (New England Biolabs) and nuclease-free H₂O at 37°C for 3 h. The restriction enzyme was inactivated at 65°C for 20 min. The linearised DNA was purified by ethanol precipitation. Briefly, 0.1 volume of sodium acetate (3 M, pH 5.2) and 2.5 volumes of ice-cold 95% ethanol were added to the DNA mixture. Mix thoroughly by vortexing and incubate at -80°C for 1 h. Centrifuge at top speed at 4°C for 30 min. Remove the supernatant and wash the DNA pellet with 1 ml 70% ethanol; centrifuge

at top speed at 4°C for 15 min. Resuspend the RNA in 30 µL nuclease-free H₂O and determine concentration by NanoDrop spectrometry and integrity by 1% agarose gel electrophoresis.

2.2.4 *In vitro* RNA transcription

2 µg of *NotI*-*HF* linearised cDNA plasmid was used as the template for the *in vitro* synthesis of 5' capped CHIKV subgenomic replicon and uncapped CHIKV pulldown baits using SP6 and T7 polymerase, respectively (ThermoScientific). Briefly, the DNA template was mixed with 5x transcription buffer (ThermoScientific), 4 mM ATP (New England Biolabs), 4 mM GTP (New England Biolabs), 4 mM CTP (New England Biolabs), 4 mM (UTP), 0.02 M MgCl₂, pyrophosphatase (New England Biolabs), RNase inhibitor and 40 U SP6 (with 0.06 µmol 3'-O-Me-m⁷G(5')ppp(5')G (New England Biolabs) for capped RNA) or T7 polymerase in an 80 µl reaction system. The mixture was incubated at 40°C for 3 h for SP6 or 37°C for 2.5 h for T7. Template DNA was removed by TURBO DNase (Invitrogen) and the RNA was purified as follows. The sample was made up to 200 µl using DEPC-H₂O and then mixed with one volume of phenol: chloroform: isoamyl alcohol pH 6.7 (Ambion). After separating the immiscible layers by centrifugation, 0.1 volume of 3 M sodium acetate (pH 5.2) and 2.5 volumes of 95% ethanol were added to the upper aqueous phase. RNA was precipitated at -80°C for at least 1 h and was then washed once with 1 mL 70% ethanol. Pelleted RNA was resuspended in 150-200 µl DEPC-H₂O. The RNA integrity was confirmed by MOPS denaturing agarose gel electrophoresis and quantified by NanoDrop spectrometry.

2.2.5 Virus production

1µl of plasmid DNA was transformed into *Mix & Go* competent cells (Zymo Research Corporation) following the manufacturer's instructions. Plasmid cDNA was purified using GeneJET Plasmid Maxiprep kits (Thermo Fisher Scientific). DNA linearisation and RNA *in vitro* transcription were performed as mentioned in 2.2.3 and 2.2.4.

BHK cells were washed, trypsinised and resuspended in 10 ml DMEM. Centrifuge at 1500 g for 3 min and resuspend the pellet in 10 ml ice-cold DEPC-PBS. Repeat twice before the final resuspension in 10 ml ice-cold DEPC-PBS and count the number of cells. Centrifuge the cells at 1500 g for 3 min and resuspend to $\geq 3 \times 10^6$ cells/ml in ice-cold DEPC-PBS. Add 400 μ l cells (1.2×10^6 cells) in a 4mm pre-cooled electrocuvette, then add 2 μ g 5' capped RNA into the bottom of the cuvette and swirl gently. Electroporate the RNA into the cells at 260 V for 25 ms with a single square wave pulse using a Bio-Rad electroporator. Quickly resuspend the mixture in 10 ml DMEM before seeding into a T75 flask. After 24h, the supernatant was aspirated and titred by plaque assay.

2.2.6 Virus infection

Infection of RD and Huh7 cells was performed on 12-well plates at a multiplicity of infection (MOI) of 0.1, 24 h post MSI2 siRNA transfection. Wild-type ICRES (7×10^7 pfu/mL) was diluted in complete DMEM, and 0.5 mL was added per well. After 10 min of adsorption on a rocking platform and 50 min incubation at 37°C/5% CO₂, the inoculum was replaced with 1 mL complete DMEM. At 24 hpi, the supernatant media was harvested for subsequent plaque assay and the monolayer was lysed with IP lysis buffer (25mM Tris-HCl pH 7.4, 150mM NaCl, 1mM EDTA, 1% NP-40, 5% glycerol).

2.2.7 Plaque assay virus titration

BHK cells were plated in 12-well plates to allow confluent monolayer formation. Virus stocks were serially diluted in complete DMEM and added 0.5mL per well. After 10 min of adsorption on a rocking platform and 50 min incubation at 37°C/5% CO₂, the inoculum was aspirated and replaced with 1 mL complete DMEM. Monolayers were fixed with 4% formaldehyde for 30 min 48 hpi. Plaque formation was visualised after staining with 0.1% crystal violet for 20 min.

2.2.8 RNA extraction from infected cells

Infected cells in 12-well plates were washed with 1x PBS, aspirated and treated with

500 µl TRIzol reagent. The monolayer was removed and lysed by washing the TRIzol over the wells. 100 µl chloroform was thoroughly mixed with samples and incubated at RT for 3 min. The upper aqueous phase was separated by centrifuging at 12000 x g for 15 min at 4°C. Samples were then mixed with 250 µl isopropanol and pelleted by centrifuging at 12000 x g for 10 min at 4°C. The pellet was washed with 500 µl ice-cold 75% ethanol and resuspended in 40 µl diethylpyrocarbonate (DEPC) treated H₂O.

2.2.9 Determine the cell viability of Ro 08-2750 with MTT assay

RD and Huh7 cells were seeded in 96-well plates at 7.5x10⁴ cells/well and incubated for 24 h as described previously. RD cells were then treated with increasing doses (1, 3, 5, 7.5, 10, 20 µM) of Ro 08-2750 (TOCRIS) and DMSO (Fisher Chemical). Huh7 cells were then treated with increasing dose (0.1, 0.5, 1, 3, 5, 10, 20 µM) of Ro 08-2750 (TOCRIS) and DMSO (Fisher Chemical). The inhibitor stock was dissolved in DMSO and aliquoted for the serial dilutions to be added to the cells with complete DMEM. After 24 h incubation, media/inhibitor was removed and 20 µl of 5 mg/mL 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) were added to the cells and incubated at 37°C for 3 h. The media was then removed and the cells were lysed in DMSO and absorbance was recorded at 570nm on an Infinite F50 microplate reader (Tecan). The absorbance values were normalised to untreated control cells.

2.2.10 Treatment with Ro 08-2750 and infection with CHIKV

RD and Huh7 cells were seeded in 12-well plates at 1x10⁵ cells/well the day before. The cells were then infected with ICRES CHIKV at MOI=0.1 and non-lethal doses of serial dilutions of Ro 08-2750 were diluted in 500µl complete DMEM and adsorbed to the cells for 1 h at 37°C. The virus or inhibitor mixture was removed and cells were thoroughly washed with PBS and incubated for 24 h in Ro 08-2750 diluted in DMEM. Supernatants were then collected to determine the titre of CHIKV released by plaque assay as described in 2.2.7.

2.2.11 Small interfering RNA knockdown of MSI1 and MSI2

RD and Huh7 cells were seeded at 1×10^5 cells/well in 12-well plates in antibiotic-free medium. After 20-24 h till the confluency reached about 80%, cells were washed with PBS and incubated with 1x Opti-MEM + GlutaMAX (Gibco) for 20 min at 37°C/5% CO₂. For each well, 50 pmol MSI1 (Santa Cruz) and/or MSI2 siRNA (Santa Cruz) were mixed with 100 µl Opti-MEM and incubated at room temperature for 1 min. No siRNAs were added to mock samples and 50 pmol of scrambled siRNA (QIAGEN) was used to evaluate specificity. Correspondingly, 3 µl Lipofectamine RNAiMAX (Invitrogen) was mixed with 100 µl Opti-MEM and incubated at room temperature for 1 min. The siRNA and Lipofectamine RNAiMAX were then thoroughly mixed and incubated at room temperature for 5 min before adding to the cells. The viability of cells was regularly monitored by microscopy to ensure that cytotoxicity was limited prior to sample collection. After 24 h, cells were either lysed to check for knock efficiency by western blot or used for subsequent CHIKV infection.

2.2.12 Maintenance of stable MSI2 knockdown RD cell lines using short hairpin RNA interference (shRNA)

Human embryonic kidney 293 (HEK 293T) cells were plated in antibiotic-free DMEM in 6-well plates and incubated for 24 h as previously described. When the confluency reached >80%, for each transfection, in one tube 300 µL OptiMEM was mixed with 1 µg p8.9 packaging plasmids, 1 µg envelope plasmids and 1.5 µg MSI2 shRNA/scrambled shRNA (Santa Cruz Biotechnology); in another tube, 300 µL of OptiMEM was mixed with 5 µL lipofectamine 2000 (Invitrogen). Both tubes were gently mixed by flicking and incubated at room temperature for 5 min, followed by mixing the contents from both tubes and incubate for 20 min at room temperature. The antibiotic-free DMEM was aspirated and monolayers were washed once with warm PBS. 800 µL Opti-MEM was added to each well, followed by dropwise addition of the lentiviral plasmids/shRNA mixture onto the cells and incubated at 37°C. After 6 h, the media was changed to antibiotic-free DMEM.

After 48h post-transfection, the supernatant containing the lentivirus was harvested and filtered through a 0.45 µm filter. RD and Huh7 cells were seeded at 1×10^5 cells/well the day prior to transduction. 1 mL of the lentivirus supernatant and polybrene (MERCK) were added to each well. The media was removed and changed to antibiotic-free DMEM after 6 h. 2.5 µg/well puromycin was added to the cells after 72 h, cell viability was checked on a regular basis and the cells were passaged and maintained with 2.5 µg puromycin. The efficiency and specificity of the shRNA were confirmed by western blot.

2.2.13 Quantification of CHIKV genomic copies

MSI2 knockdown RD and Huh7 cells were infected with wild-type CHIKV at MOI=0.1. After 24h, total RNA was extracted from the monolayer using TRI Reagent Solution (Applied Biosystems) as previously described. Purified RNA samples were reverse transcribed to cDNA using High-Capacity RNA-to-cDNA kit (Thermo Fisher Scientific), according to the manufacturer's instructions. Briefly, 2 µg of RNA was added to the 20 µl reaction system. Negative controls were included without adding the reverse transcriptase enzyme mix. Samples were incubated at 37°C for 1h and the reaction was terminated by heating at 95°C for 5 min. The resulting cDNA was used for qPCR following the manufacturer's instructions of 2xqPCRBIO SyGreen Blue Mix Lo-ROX (PCR Biosystems). CHIKV positive-sense strand primers (Forward: CCGACTCAACCATCCTGGAT; Reverse: GGCAGACGCAGTGGTACTTCCT) were selected for amplification and to detect the difference in the viral RNA level. Carboxy-X-rhodamine (ROX) passive reference dye was included in each reaction for signal calibration. The reaction condition was set up that the enzyme was initially activated by heating to 95°C for 2 min, followed by 40 cycles of amplification (95°C for 5 sec; 60°C for 30 sec). A dissociation curve 60°C-95°C as pre-defined by the Mx3005P thermal cycler (Agilent technologies). *In vitro* transcribed CHIKV ICRES RNA was also reverse transcribed and the consequent cDNA was serially diluted to create a standard to quantify copy numbers in the respective samples.

2.2.14 Strand-specific quantification of CHIKV RNA

MSI2 knockdown RD and Huh7 cells were infected with wild-type CHIKV at MOI=0.1. After 8h, 10h 12h and 24h, total RNA was extracted from the monolayer using TRI Reagent Solution (Applied Biosystems) according to the manufacturer's protocol. ssqPCR was performed according to the protocol described by Plaskon and colleagues (267). Briefly, 250 ng of RNA were reverse-transcribed with CHIKV strand-specific primers (positive: GGCAGTATCGTGAATTCGATGCGTCTGCTCTGTCTA CATGA; negative: GGCAGTATCGTGAATTCGATGCGACACGGAGACGCCA ACATT) using the SCRIPT cDNA Synthesis Kit (Jena Bioscience) according to the manufacturer's instructions. 100ng of strand-specific cDNA was used as the template for the quantitative PCR using the qPCRBIO SyGreen Blue Mix Lo-ROX (PCR Biosystems) with CHIKV strand-specific primers [(positive forward: AATAAATCATAAGACACGGAGACGCCAACATT; positive reverse: GGCAGT ATCGTGAATTCGATGC); (negative forward: GGCAGTATCGTGAATTCGATGC; negative reverse: AATAAATCATAAGTCTGCTCTGTCTACATGA)] at 95°C for 2 min, 40 cycles of (95°C for 5 sec, 60°C for 30 sec), and a dissociation curve 60°C-95°C as pre-defined by the Mx3005P thermal cycler (Agilent technologies). In vitro transcribed CHIKV ICRES RNA was also reverse transcribed and the consequent cDNA was serially diluted series to create a standard to quantify copy numbers in the respective samples.

2.2.15 Sub-genomic replicon transfection and analysis

1µl of plasmid DNA was transformed into *Mix & Go* competent cells (Zymo Research Corporation) following the manufacturer's instructions. Plasmid cDNA was purified using GeneJET Plasmid Maxiprep kits (Thermo Fisher Scientific). DNA linearisation and RNA *in vitro* transcription were performed as mentioned in 2.2.3 and 2.2.4. After transfections with siRNAs (24 h for MSI2), RD and Huh7 cells were washed with PBS and incubated with 1 mL Opti-MEM (Gibco) for 30 min at 37°C/5% CO₂ before the transfection medium was added. 5 µl of Lipofectamine 2000 (Invitrogen) was diluted in 100 µl Opti-MEM and incubated under room temperature for 5 min; correspondingly,

1000 ng of capped subgenomic replicon RNA were mixed with 100 μ L Opti-MEM and incubated under room temperature for 5 min. The transfection reagent and RNA were then mixed together and incubated at room temperature for 25 min before evenly adding to the monolayer. The transfection medium was removed and replaced with fresh complete DMEM after 6 h. At 8 h and 24 h post replicon transfection, cells were carefully washed twice with PBS, lysed with 300 μ l 1x passive lysis buffer (PLB, Promega) at room temperature for 15 min and stored at -80°C until used. The analysis was carried out using the luciferase assay reagent and Stop & Glo reagent (Promega) and a FluroStar Optima luminometer to measure the levels of luciferase expression. The data was processed and recorded as relative light units (RLU).

2.2.16 Trans-complementation Assay

1 μ l of plasmid DNA was transformed into *Mix & Go* competent cells (Zymo Research Corporation) following the manufacturer's instructions. Plasmid cDNA was purified using GeneJET Plasmid Maxiprep kits (Thermo Fisher Scientific). The *trans*-complementation assay was performed in RD cells as previously described (140). Briefly, MSI2 knockdown RD cells were seeded in 12-well plates at 1×10^5 cells/well the day before transfection. The $> 80\%$ RD cells were transfected with 1 μ g of the homologues pair of plasmids, namely the plasmid encoding for the wild-type/MSI2 BSM CHIKV template RNA and the plasmid encoding for wild-type/GAA CHIKV replicase proteins, using Lipofectamine 2000 (Invitrogen) according to the manufacturer's protocol. The cells were incubated at 37°C 5% CO_2 for 6h before changing the media from Opti-MEM to antibiotic-free DMEM.

Samples were collected at 8 h and 24 h post-transfection as previously described. The monolayers were washed once with 1xPBS and then lysed with 1xPLB (Promega) under room temperature for 15 min with frequent rocking. The signal of firefly luciferase and Gaussia luciferase was measured using the Dual-Luciferase[®] Reporter Assay System (Promega). The analysis was carried out using the FluroStar Optima luminometer and the data was processed and recorded as RLU as described in 2.2.15.

2.2.17 3' end biotinylation of CHIKV RNA and pulldown with infected RD lysates

The infected cell lysates were prepared by seeding cell lysates into 6 T75 flasks at 1×10^6 cells/flask. On the next day, the cells were uninfected with CHIKV ICRES at an MOI of 10 in 4 ml DMEM/flask. The virus was allowed to absorb to cells for 1 h at 37°C 5% CO₂. 6 ml DMEM/flask was then added and the cells were further incubated for 23 h. The medium was removed and cells were washed twice with PBS. Scrape the attached cells from the bottom of the flasks using a cell scraper in 10 ml fresh PBS. Centrifuge the floating cells at 1000 g for 5 min. PBS was removed and the pellet was resuspended in 750 µl PBS. The cells were centrifuged again at 1000 g for 5 min. Remove PBS and lyse the cells in 1 ml IP lysis buffer (25 mM Tris-HCl pH 7.4, 150 mM NaCl, 1 mM EDTA, 1% NP-40, 5% glycerol). Incubate on ice for 30 min with frequent vortexing. Cell debris was pelleted by centrifuging at top speed at 4°C for 10 min and the supernatant cell lysates were transferred to fresh tubes to determine total protein concentration using the Pierce BCA Protein Assay kit (Thermo Scientific). The concentration should be at least higher than 2 mg/ml.

The RNA bait of the first 337 nt of CHIKV 5' end was prepared by amplifying the 5' end region of the ICRES CHIKV (forward primer: TAATACGAC TCACTATAGGGATGGCTGCGTGAGACACACG; reverse primer: CGCACTG CGCATCGGGCAGA) and *in vitro* transcribed. A negative control RNA of the same length was prepared by amplifying a region of firefly luciferase (forward primer: GTGGACATTACCTACGCCGAGT; reverse primer: GAAGCCCTGGTAGTCG GTCTTG). The RNA was *in vitro* transcribed and purified as mentioned in 2.2.4. The RNA was then biotinylated following the manufacturer's protocol (ThermoFisher Scientific). Briefly, purified RNA was incubated overnight at 16°C mixed with 10x RNA ligase reaction buffer, RNase inhibitor, biotinylated cytidine bisphosphate, T4 RNA ligase, 30% PEG and DEPC-free H₂O. Biotinylated RNA was purified using phenol-chloroform

precipitation and was resuspended in ~13 μL DEPC-free H_2O . 1.5 μL RNA was used to measure concentration with NanoDrop spectrometry and another 1.5 μL was used to check RNA integrity in 1% denaturing MOPS gel. The remaining 10 μL RNA was used in the subsequent folding reaction by first incubating at 95°C for 5 min, followed by 2 min on ice with the addition of 1 μL RNase inhibitor, and then mixed with 16.6 μL 3.3x folding buffer (333 mM HEPES pH 8.0, 13.4 mM MgCl_2 , 333 mM NaCl) and 23.3 μL DEPC- H_2O and incubate at 37°C for 25 min.

Resuspend Dynabeads MyOne Streptavidin T1 (Invitrogen) in the vial by vortexing. ~15 μL beads/sample into fresh tubes and mixed with 1 ml of Binding and Washing Buffer (5 mM Tris pH 7.5, 0.5 mM EDTA, 1 M NaCl) and vortex for 1 min. Separate the beads from the buffer using a magnetic rube rack and remove the buffer. Repeat this washing step for three times. Next, wash beads twice in 100 μL of Solution A (DEPC-treated 0.1 M NaOH, DEPC-treated 0.05 M NaCl) by vortexing for 1 min. Wash beads once in 100 μL Solution B (DEPC-treated 0.1 M NaCl) by vortexing for 1 min. Finally, wash beads twice with 100 μL 20 mM Tris pH 7.5 by vortexing for 1 min.

To immobilise the biotinylated RNA to beads, resuspend the beads in an equal volume to RNA (50 μL) of 1X RNA Capture Buffer (20 mM Tris pH 7.5, 1 M NaCl, 1 mM EDTA) to beads. The biotinylated and folded RNA, plus 1 μL RNase inhibitor, was then added to the beads and was mixed gently by flicking the tube. Incubate the mixture for ≥ 2 h at 4°C on a spinning wheel. The beads were then washed twice with 100 μL 20 mM Tris pH 7.5 by gentle flicking to minimise disturbance to the interaction, followed by one wash in 100 μL 1 x Protein-RNA Binding Buffer (0.02 M Tris pH 7.5, 0.05 M NaCl, 2 mM MgCl_2 , 0.1% Tween20) by gentle flicking. Resuspend the beads in 100 μL protein mix containing 200 μg RD cell lysates, 1 μL RNase inhibitor, 10 μL 10x protein-RNA binding buffer (0.2 M Tris pH 7.5, 0.5 M NaCl, 20 mM MgCl_2 , 1% Tween20) and make it up to 100 μL with 50% glycerol in DEPC- H_2O . Incubate the mixture for ≥ 2 h at 4°C on a spinning wheel. The mixture was then transferred to a pre-cooled small plastic plate and irradiated with ≥ 0.15 J/cm^2 at 254 nm UV light. Wash the crosslinked samples

three times with 100 μ L of Wash buffer 1 (20 mM Tris pH 7.5, 10 mM NaCl, 0.1% Tween20) before eluting the bound proteins by boiling at 95°C in 27 μ L DEPC-H₂O and 9 μ L 4x SDS loading dye (8% SDS, 40% glycerol, 60% 2-mercaptoethanol, 0.008% bromophenol blue and 0.125 M Tris HCl, pH=6.8) for 10 min and analyse on Western blot.

2.2.18 Co-immunoprecipitation (Co-IP) of MSI2 with CHIKV proteins

1x10⁵ RD cells/well were seeded in 12-well plates the day before infection. Uninfected RD cells were used as the negative control. The cells were then infected with ICRES CHIKV at MOI=10. The monolayers were washed with 1xPBS and lysed with IP lysis buffer (25 mM Tris-HCl pH 7.4, 150 mM NaCl, 1 mM EDTA, 1% NP-40, 5% glycerol). Cell lysates were incubated on ice for 1 h with an intermittent vortex to ensure complete lysing, followed by cell debris removal by centrifuging at top speed for 15 min at 4°C. 10 μ L of Dynabeads™ Protein G (Thermo Fisher Scientific) were washed with IP lysis buffer once before mixing with the cell lysates, which was precleared of non-specific binding by incubating at 4°C for 1 h in a tube roller. Input samples were taken at this point and stored at -20°C. The beads were removed with magnetic rack, MSI2 antibody (1:50) was added to the supernatant and incubated at 4°C overnight in a tube roller. The antibody was then caught by adding 20 μ L of Dynabeads™ Protein G and incubated at 4°C for 2 h in a tube roller. The supernatants were removed and the beads were washed four times for 5 min with IP lysis buffer, followed by resuspending in 15 μ L IP lysis buffer and 5 μ L 4x SDS loading dye (8% SDS, 40% glycerol, 60% 2-mercaptoethanol, 0.008% bromophenol blue and 0.125 M Tris HCl, pH=6.8). The mixture was heated at 95°C for 10 min before validating with western blot.

2.2.19 Expression, purification and ion exchange of His-tagged MSI2 protein

The plasmid expressing the RNA binding domains of human MSI2 pET-22HT-MSI2 (8-193) was a kind gift from Sean Ryder (Addgene plasmid # 60356;

<http://n2t.net/addgene:60356>; RRID: Addgene_60356). The plasmid was transformed to DH5 α competent cells by incubating on ice for 30 min, followed by 1 h in 1 mL antibiotic-free LB in a 37°C shaking incubator. The bacteria were then spread on ampicillin plates before individual colonies were selected, plasmid extracted and sequenced.

The correct plasmid was then transformed into BL21 (DE3) competent cells following the manufacturer's protocol (New England Biolabs). Single colony was resuspended in 10 mL LB with ampicillin in a 37°C shaking incubator overnight. The bacterial culture was then added to 1 L LB with ampicillin and 100 μ M Isopropyl β -D-1-thiogalactopyranoside (IPTG; Thermo Fisher Scientific) as soon as the optical density reached between 0.6~0.8. The culture was incubated overnight in an 18°C shaking incubator and high-speed centrifugation was used to pellet the bacteria, which was then stored at -20°C for 24 h. The pellet was resuspended and lysed on ice for 30 min in a buffer containing DNase I, RNase A, lysozyme and protease inhibitor. After thorough sonication, the mixture was centrifuged twice at 4200 rpm for 1 h at 4°C. The supernatant containing the proteins was filtered through a 0.45 μ m filter. The proteins were purified with HisTrapTM FF column (Cytiva) using the Econo Gradient Pump (BIORAD) according to the manufacturer's protocol, followed by dialysing overnight to remove imidazole. The purified protein was quantified and the purity and identity were examined by SDS-PAGE and western blot.

For increasing purity and minimising miscellaneous proteins, His-MSI2 protein was further purified by ion exchange. The protein solution was first desalted using PD-10 desalting columns (Cytiva) following the manufacturer's protocol. In particular, the column was equilibrated with washing buffer (50 mM MES, 10 mM NaCl, pH 5.6, degassed) and the protein was eluted with the same buffer. Ion exchange was kindly monitored and assisted by Dr Brian Jackson using HiTrap SP HP cation exchange chromatography column (Cytiva), and the elution buffer used contains 50mM MES, 1M NaCl, pH 5.6, degassed. The eluted protein fractions were checked on SDS-PAGE and

the fraction with the highest purity was quantified and identified by western blot.

2.2.20 Electromobility shift assay (EMSA)

The RNA bait of the first 337 nt of CHIKV 5' end was prepared by amplifying the 5' end region of the ICRES CHIKV (forward primer: TAATAC GACTCACTATAGGGATGGCTGCGTGAGACACACG; reverse primer: CGCACT GCGCATCGGGCAGA) and *in vitro* transcribed. The 5' phosphate of the un-labelled RNA was removed by incubating with quick CIP (New England Biolabs) at 37°C for 10 min. The reaction mixture was purified with RNA Clean & Concentrator Kits (Zymo Research) and the RNA was resuspended in RNase-free H₂O. The RNA was then 5' end radiolabelled with ATP, [γ -³²P] (Perkin Elmer) using T4 polynucleotide kinase (Thermo Fisher) and incubated at 37°C for 1 h. The reaction mixture was purified with RNA Clean & Concentrator Kits (Zymo Research) and the RNA was resuspended in RNase-free H₂O. The RNA-protein binding reaction was performed by denaturing the radiolabelled RNA with 0.5x Tris-EDTA (TE) buffer (5 mM Tris, 0.5 mM, pH 8.0), 3.3 x folding buffer (333 mM HEPES pH 8.0, 13.4 mM MgCl₂, 333 mM NaCl), RNasin Plus RNase Inhibitor (Promega) and un-labelled RNA competitor at 95°C for 4 min, followed by incubation at 37°C for 25 min to refold. The protein reaction mixture was prepared by mixing purified His-MSI2 protein with 3.3x folding buffer, 0.5 x TE buffer, 100% glycerol, RNasin Plus RNase Inhibitor (Promega) and yeast tRNA as non-specific competitors. The mixture was incubated at 37°C for 2 min before adding to the radiolabelled RNA mixture and further incubated at 37°C for 15 min. The RNA-protein sample was mixed with native loading dye and ran on a native PAGE gel at 135 V for ~2 h. The gel was fixed for 30 min and dried with gel dryer (BIORAD). The gel was exposed onto Hyperfilm™ ECL™ (Merck) overnight or for several days depending on the radioactivity of the samples and visualised by Xograph.

2.2.21 Western blot and Coomassie blue staining

Cells were washed twice with 1 mL ice-cold PBS, and total lysates were prepared by

lysing with IP lysis buffer (25 mM Tris-HCl pH 7.4, 150 mM NaCl, 1 mM EDTA, 1% NP-40, 5% glycerol) on a rocking platform for 20 min at 4°C. Cell debris was removed by centrifuging at top speed for 10 min at 4°C and protein concentration was determined using the Pierce BCA Protein Assay kit (Thermo Scientific). Samples were mixed with 2 x SDS buffer (4% SDS, 20% glycerol, 30% 2-mercaptoethanol, 0.004% bromophenol blue and 0.125 M Tris HCl, pH=6.8) and then boiled for 5 min at 95°C. An equal amount of each sample was loaded to 10% SDS-PAGE gels as well as the protein marker (11-245 kDa; New England Biolabs). After running at 130 V for 1.5 h, the samples were transferred onto polyvinylidene fluoride microporous membranes (PVDF) at 15 A, 500 mA for 1 h. The membrane was then blocked with 10 mL of 50% Odyssey blocking buffer (LI-COR) diluted in 1 x Tris-buffered saline with 0.1% Tween20 (TBST) for 1 h at room temperature. Primary antibodies were diluted in the blocking buffer and incubated overnight at 4°C. The membrane was washed once with 1 x TBST for 5 min before incubation with secondary antibodies for 2 h under room temperature. The membrane was then washed three times with 1 x TBST and proteins were visualised by the Odyssey scanner (LI-COR).

For Coomassie blue staining, after running the SDS-PAGE gels, it was incubated at room temperature in the de-stain solution (40% methanol and 10% acetic acid) for 5 min to remove background, followed by 1 h room temperature incubation in Coomassie blue solution (0.1% Coomassie Brilliant Blue R-250, 40% methanol and 10% acetic acid). The gel was then de-stained at room temperature overnight until the gel background was almost transparent. All incubations were done on a rocking platform.

2.2.22 Statistical analysis

Statistical analysis was carried out with one-way ANOVA and Dunnett's multiple comparisons test on GraphPad Prism version 8.4.0. *P* values of ≤ 0.05 (*), ≤ 0.01 (**), ≤ 0.001 (***) were used to represent degrees of significance between each drug treatment/knockdown/mutant to wild-type.

CHAPTER 3: PHENOTYPIC ANALYSIS OF THE ROLE OF MSI2 IN CHIKV REPLICATION USING A SMALL MOLECULE INHIBITOR

3.1 Introduction

CHIKV genome replication takes place within the membranous replication complex termed spherule (268). As the centre for virus RNA synthesis, a number of host cytosolic proteins are recruited to the spherules, where they play indispensable roles in CHIKV genome replication (188). As described in 1.4.2, preliminary work leading to this study identified host cell encoded MSI2 as a binding factor to CHIKV 5' RNA, however its role and necessity during CHIKV replication had yet to be investigated. Consequently, this study sought to investigate 1). The effect of MSI2 functional inhibition on CHIKV replication and 2). The associations between MSI2 and CHIKV proteins.

1). The RNA-recognition motif (RRM) 1 of MSI2 was selectively blocked by a small molecule inhibitor designated Ro 08-2750 (Ro), thereby precluding its association with RNAs. The expression of MSI2 in the two cell lines (Huh7 and RD) used in this study was first examined by western blot. According to UniprotKB, the 'canonical' MSI2 has a molecular weight of ~35 kDa, whereas several isoforms with different amino acid length have been recorded in human cells (262). In particular, MSI2 isoforms designated A0A6Q8PF05 and B4DHE8 contain both RRM1s and have similar molecular weight (~34 kDa and ~36 kDa respectively) to the canonical MSI2 (262). Therefore, if these MSI2 RRM1-containing isoforms also share the same epitope for the monoclonal MSI2 antibody, it was expected to detect multiple MSI2 bands in the western blot analysis.

The effect of MSI2 inhibition was assayed with infectious CHIKV in order to investigate the effect on the complement virus lifecycle, and sub-genomic replicon and *trans*-complementation assay systems to narrow down to the role of MSI2 to specific stages of the CHIKV replication cycle. While the sub-genomic replicon provides a biologically safe and recapitulative approach to investigate virus replication, the *trans*-complementation system offers a sensitive model for uncoupling the measurement of CHIKV genome replication from translation (269). Data generated from these assays would shed light on the requirement of MSI2 for CHIKV replication.

2). In order to investigate potential interactions between MSI2 and CHIKV encoded proteins, this study employed RNA affinity purification and co-immunoprecipitation (Co-IP). RNA affinity purification is based on the high-affinity binding of streptavidin agarose beads to biotin, which is chemically tagged to the 3' end of RNA. Single protein or protein complexes bound to the RNA can then be eluted from the beads and analysed by western blot or mass spectrometry (MS) (270). Due to the two RNA-binding domains in MSI2, it was considered that MSI2, as well as other virus proteins, directly binds to the CHIKV 5' RNA. Therefore, RNA affinity purification was expected to show the presence of MSI2 and nsP3 as it did in the preliminary study. Co-IP is a technique to detect physiologically relevant protein-protein interactions. Following the lysis of CHIKV infected cells, MSI2 was captured by its corresponding antibody bound to magnetic protein G beads. Cellular or CHIKV proteins bound directly or indirectly to MSI2 via protein complexes interacting with MSI2 were eluted and analysed by western blot (Figure 3.1). This study hypothesised that MSI2 would associate with CHIKV nsPs involved in virus genome replication, such as nsP1 and/or nsP3, which were also significantly enriched from the preliminary RNA affinity purification analysis.

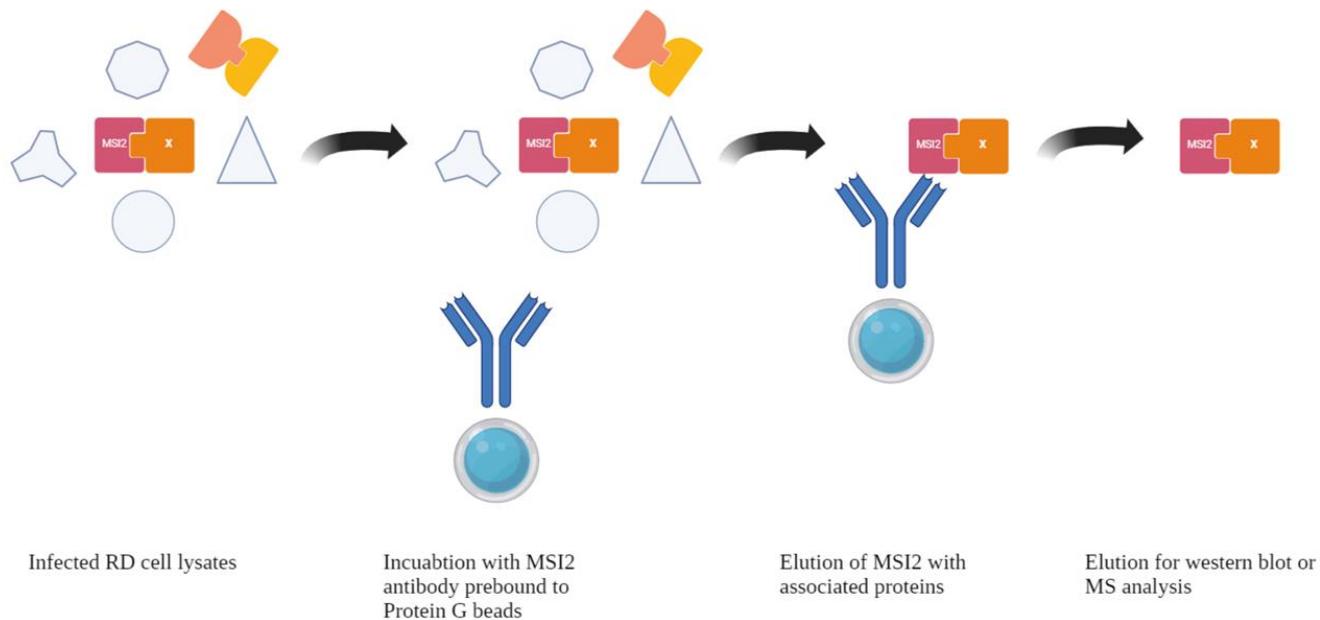


Figure 3.1. Schematic illustration of the procedures of co-immunoprecipitation and RNA affinity purification.

Whole cell protein lysates from CHIKV-infected lysates cells were incubated with MSI2 antibody prebound to protein G beads. Cellular or virus proteins associated with MSI2 were precipitated by the beads, followed by elution for western blot or mass spectrometry analysis.

3.2 Results

3.2.1 Inhibition of the MSI2 RNA-recognition motif 1 impaired CHIKV genome replication

In order to investigate if MSI2 was required for CHIKV replication, we used the small molecule inhibitor Ro to block MSI2 RNA binding capacity and measured its effect on productive virus replication. Ro was initially found to function as the non-peptide inhibitor of nerve growth factor, whose binding to its receptors is inhibited presumably by conformational change (271, 272). A recent study demonstrated that the RNA-binding activity of RRM 1 of MSI2 was selectively competed by Ro and the oncogenic

activity of MSI2 was subsequently suppressed (273). RRM1 of MSI2 is crucial for associating with RNA, and blockage of such functional domains would prohibit its ability to participate in cellular processes or virus replication.

Prior to the assay, Ro cytotoxicity in RD and Huh7 cells was analysed by MTT assay to determine the maximum non-toxic dose, which is ideally above 60% cell survival. The cell lines that we used for these assays were RD (human muscle) and Huh7 (human liver) cells, both of which represent physiological target of CHIKV infection with high permissiveness (274), and express MSI2 as confirmed by western blot (Figure 3.2.1). Notably, the MSI2 antibody detected a cluster of bands in both cell lines, suggesting the presence of MSI2 isoforms with the same epitope but different molecular weights resulted from alternative splicing (256). While the cytotoxic effect of Ro on both cells was similar, RD was more resistant, as it took 7.5 μ M for RD cells to cause less than 70% cell death, whereas for Huh7 cells it took 5 μ M (Figure 3.2.2). According to a published study, 5 μ M Ro treatment inhibited MSI2 in human myeloid leukaemia cells (273). Thus, it was considered that 7.5 μ M Ro for RD cells and 5 μ M for Huh7 cells was sufficient for MSI2 inhibition without inducing cytotoxicity. Dimethyl sulfoxide (DMSO) has no detectable effect and was therefore used as the negative control.

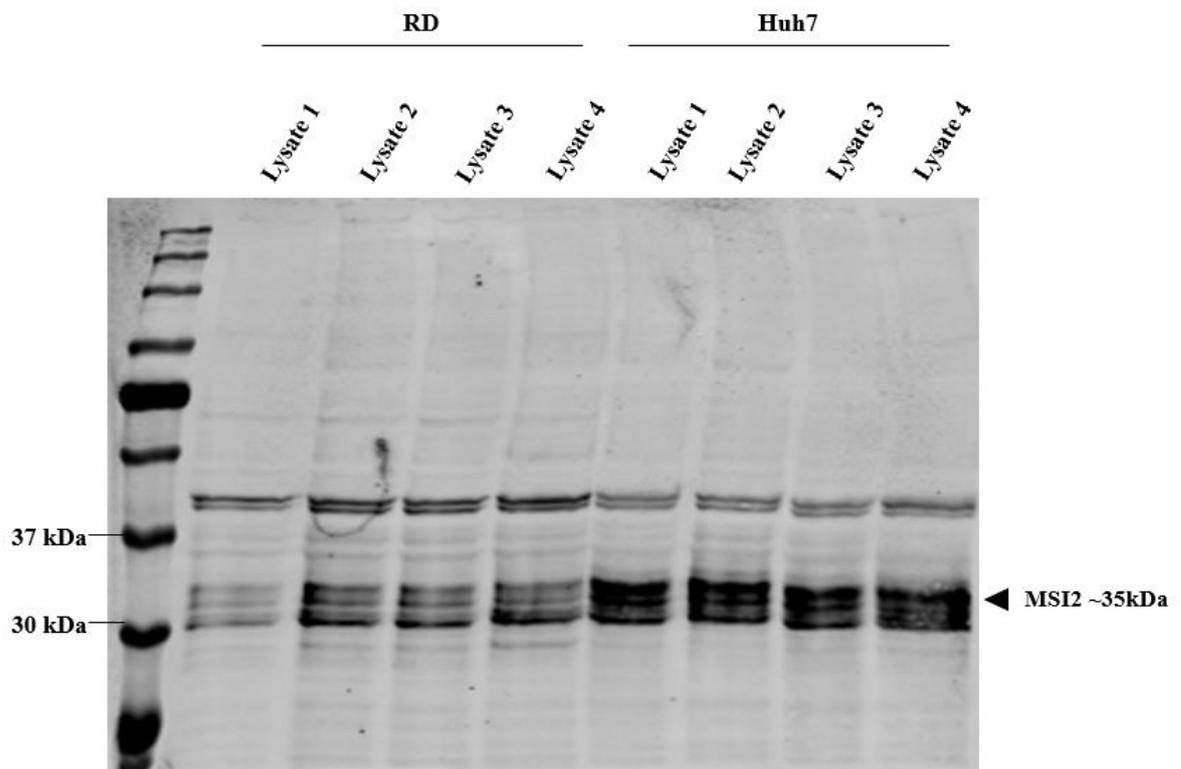


Figure 3.2.1. Detection of MSI2 expression in RD and Huh7 cells.

A total of four lysates for each cell type were used and the membrane was probed with MSI2 antibody. The cluster of bands slightly above the 30 kDa marker represents MSI2.

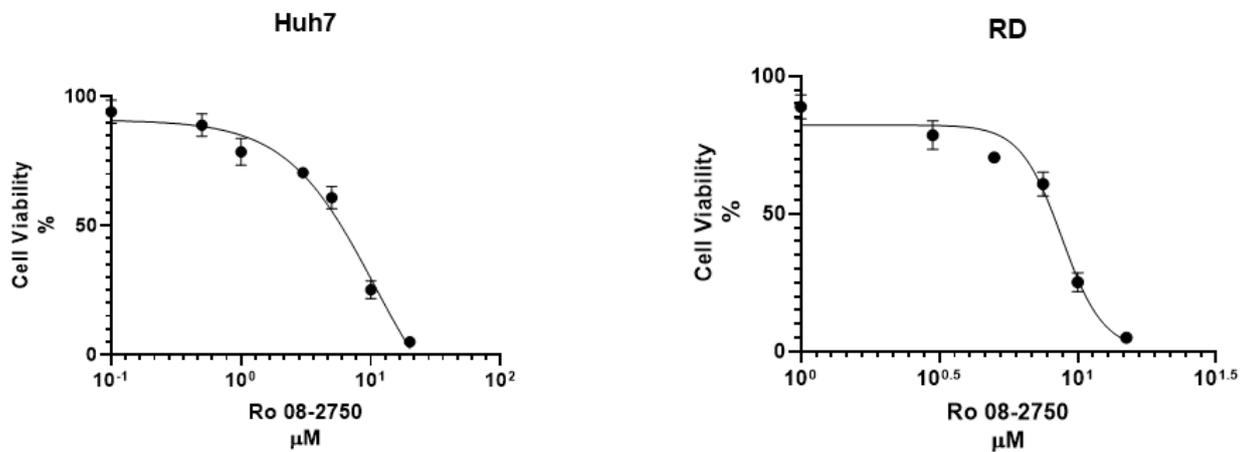


Figure 3.2.2. MTT assay of Ro 08-2750 in Huh7 and RD cells.

For Huh7 cells, the dosage of Ro used was 0.1, 0.5, 1, 3, 5, 10, 20 μM. For RD cells, the dosage of Ro used was 1, 3, 5, 7.5, 10, 20 μM. The maximal dosage to maintain at least 60% cell viability is 7.5 μM for Huh7 and 5 μM for RD. The average of three independent experiments (n=3) is shown. Error bars represent standard deviation from the mean.

To investigate the effect of Ro-mediated MSI2 inhibition on CHIKV replication, infectious CHIKV was used to assay its effect on the full-virus replication cycle and a dual-luciferase sub-genomic replicon system to assay its effect on CHIKV genome replication and translation isolated from other stages of the replication cycle. In the sub-genomic replicon system, a *Renilla* gene is fused in frame within the nsP3 encoding region, and the CHIKV structural encoding genes in ORF-2 are replaced by the firefly luciferase gene (Fig. 3.2.3A & B, top panel). For infection, RD and Huh7 cells were pre-treated with 7.5 and 5 μM of Ro for 1 h before virus infection or replicon transfection. Both cells were infected with ICRES at MOI of 0.1 in the presence of Ro. Virus samples in the supernatant were harvested for plaque assay at 8 and 24 h post infection. To maintain consistent MSI2 inhibition, Ro was co-incubated with virus throughout the experiment. The result was intriguing in two aspects: 1). MSI2 inhibition led to a

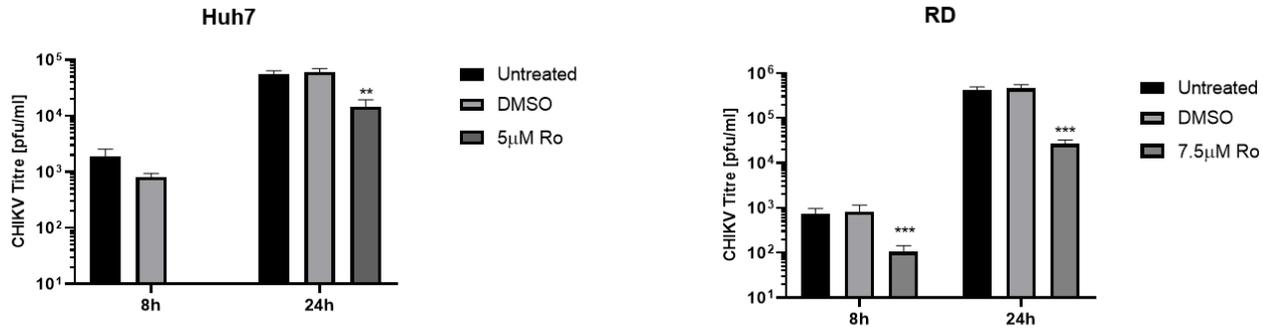
statistically significant 0.5-log lower fold change in virus production at 24 h post infection for both cells; 2). at the early time point, virus production was also significantly reduced for RD cells, whereas no productive CHIKV replication was detected for Huh7 cells at 8 h post infection (Figure 3.2.3A). These results suggested that MSI2 was required for CHIKV replication in terms of the full replication cycle. However, they did not show which stages of the CHIKV lifecycle MSI2 affected.

To narrow the effect of MSI2 down to translation and replication of CHIKV replication, CHIKV dual-luciferase sub-genomic replicon was used. The sub-genomic replicon system not only provides a virus-free approach for phenotypic analysis, but also uncouples genomic translation and replication from other stages of CHIKV lifecycle, such as entry and packaging. For this assay, Huh7 cells were not used due to lack of parallel data from the infection assay. RD cells were transfected with 1 µg dual-luciferase sub-genomic replicon in the presence of Ro. The luciferases samples were obtained from lysing the monolayer for bioluminescence analysis at 8 and 24 h post transfection. To maintain consistent MSI2 inhibition, Ro was co-incubated with the replicon throughout the experiment. The expression of the genomic ORF-1 was assayed by the *Renilla* luciferase signal to represent translation, while the expression of the sub-genomic ORF-2 was assayed by the firefly luciferase signal to represent replication. At both early and late time points, treatment with Ro led to approximately 0.5-log lower fold change in the signal of both luciferases in RD cells (Figure 3.2.3B). Therefore, the result indicated that inhibition of the MSI2 RNA binding capacity significantly impaired CHIKV replication and/or translation, as the reduction in luciferase signals could be resulted from either impaired translation of nsPs, or impaired replication of genomic and sub-genomic RNA.

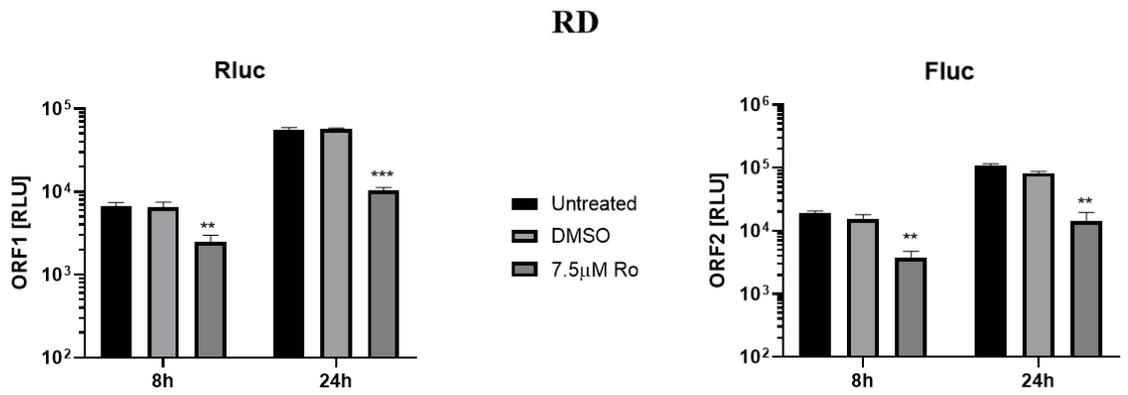
To further investigate whether MSI2 played a role in CHIKV genomic translation or replication, *trans*-complementation system was also employed which uncouples the translation of viral replicase from genomic RNA replication. In this system, the level of CHIKV genome replication was assayed by the signal of the firefly luciferase, while the

level of CHIKV sub-genome replication was assayed by the signal of gaussia luciferase. Consistent with the sub-genomic replicon result, Ro treatment led to a significant ~ 0.3 -log fold reduction in the firefly luciferase signal, while a significant ~ 0.3 -log fold reduction was also observed in the gaussia luciferase signal at both 8 and 24 h post transfection (Figure 3.2.3C). Therefore, this result demonstrated that both CHIKV genomic and sub-genome replication were impaired due to MSI2 inhibition in RD cells. However, it remained obscure that whether the impaired subgenomic replication was a collateral result of impaired genome replication, or MSI2 equally contributed to both processes.

A)



B)



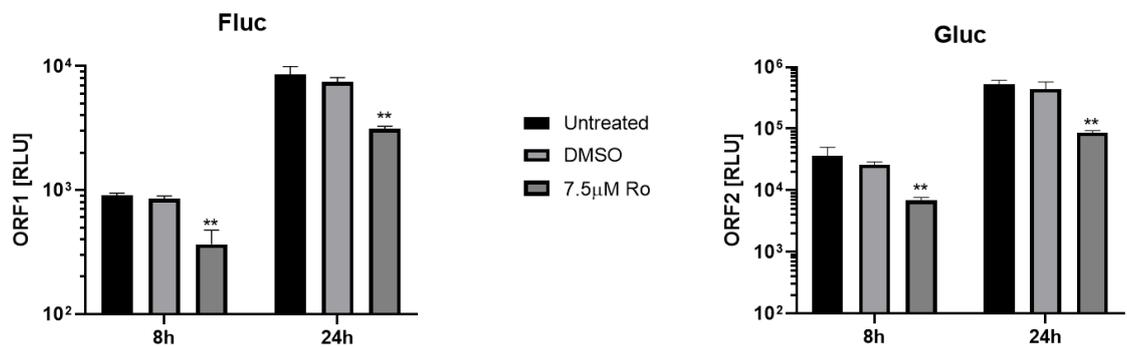
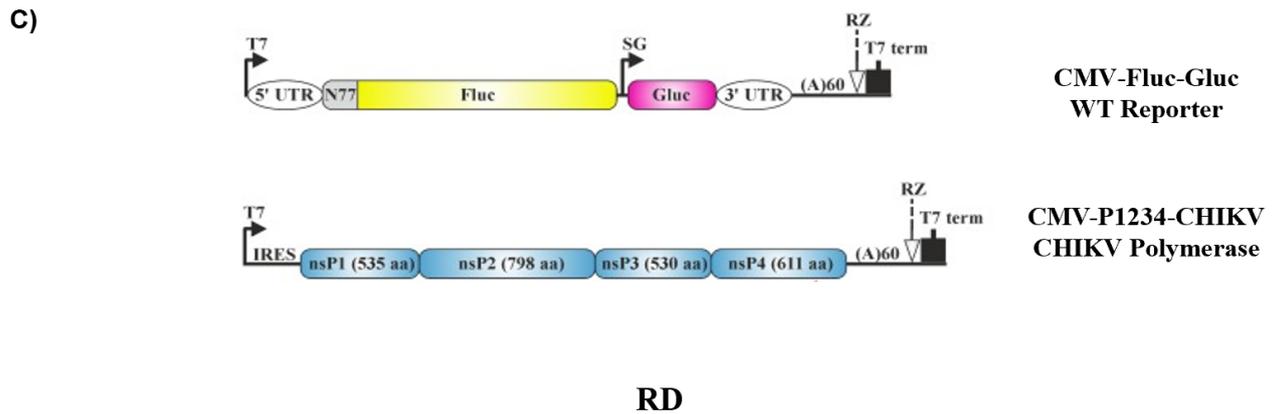


Figure 3.2.3. Inhibition of the RNA-binding activity of MSI2 reduced CHIKV replication.

A). Top: Schematic representation of the CHIKV genome. Bottom: Huh7 and RD cells were infected with CHIKV (MOI=2) in the presence of 5 and 7.5 μ M Ro, respectively. The virus produced was collected at 8 h and 24 h post-infection and titred by plaque assay. **B).** Top: Schematic representation of CHIKV dual-luciferase sub-genomic replicon. Bottom: RD cells were transfected with the CHIKV dual-luciferase sub-genomic replicon in the presence of 7.5 μ M Ro. The luciferase signal was determined at 8 h and 24 h post-transfection. **C).** Top: Schematic representation of CHIKV trans-complementation system (140). Bottom: RD cells were transfected with the CHIKV trans-complementation replicon system in the presence of 7.5 μ M Ro. The luciferase signal was determined at 8 h and 24 h post-transfection. Fluc values indicate the expression of ORF-1 and Gluc of ORF-2. DMSO served as a carrier control. The average of three independent experiments ($n=3$) is shown. Error bars represent standard deviation from the mean. *, $P<0.05$; **, $P<0.01$; ***, $P<0.001$.

Unexpectedly, phenotypic analysis in Huh7 cells not only produced uninterpretable data using the *trans*-complementation system, but also failed to generate virus under Ro treatment (Figure 3.2.3A). The possible reason was discussed in more detail in 3.3.2 below. Instead, the genomic copies of intracellular positive- and negative-sense ICRES RNA were quantified by qRT-PCR as an alternative, independent measurement of the effect of Ro treatment on CHIKV replication. Total RNA was extracted from Huh7 cells and was CHIKV-specifically reverse transcribed for qPCR analysis. Due to Ro treatment, the synthesis of both positive- and negative-sense strand RNA showed significant ~0.5-log fold decrease at 8 and 24 h post infection (Figure 3.2.4). In particular, the presence of CHIKV genomic RNA of Ro-treated samples at early time points implied that the controversial lack of virus production observed in the infection assay for Huh7 cells could be due to Ro-induced egress hindrance. Together, from these data, it was inferred that the genome replication of CHIKV was dependent on the MSI2 RNA-binding activity in human cells.

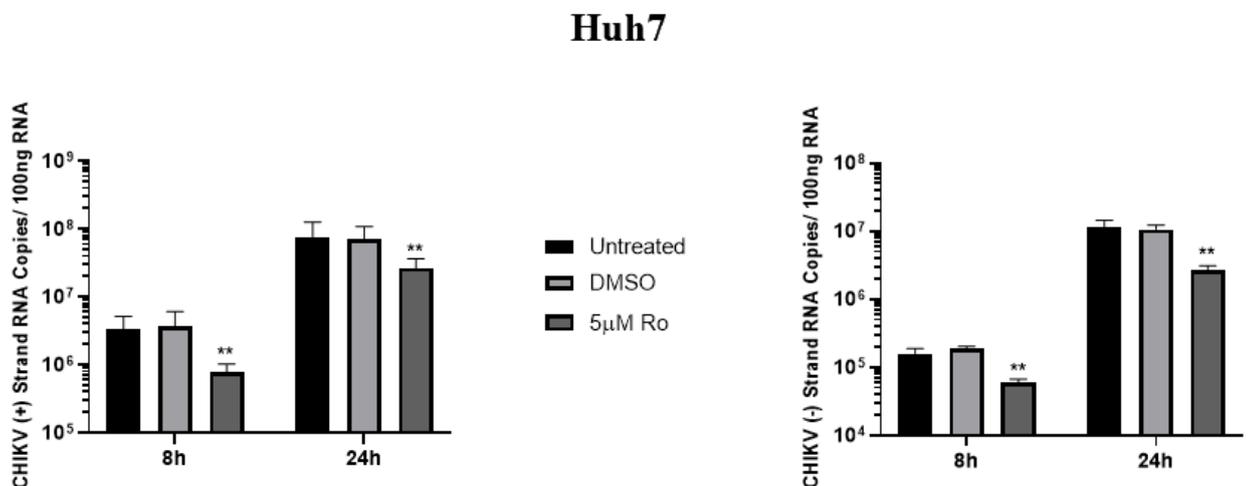


Figure 3.2.4. Positive- (left) and negative-sense (right) RNA was quantified by reverse transcription followed by qPCR under Ro 08-2750 treatment in Huh7.

DMSO served as a carrier control. The average of three independent experiments ($n=3$) is shown. Error bars represent standard deviation from the mean. *, $P<0.05$; **, $P<0.01$; ***, $P<0.001$.

3.2.2 MSI2 actively participates in the CHIKV replicase complex

As a novel factor, the involvement of MSI2 in CHIKV replication remains obscure. In the absence of literature support, it would be intriguing to investigate whether MSI2 associates with the components of CHIKV replicase complex to better understand the role of this host cell protein. For RNA affinity purification in this study, western blot analysis was used to identify MSI2 and virus proteins. The 5' end 337 nt CHIKV RNA bait was biotinylated, and a firefly luciferase gene fragment of the same length was biotinylated in a parallel and used as a non-specific negative control within this assay. The biotinylated RNA was assayed by denaturing MOPs gel electrophoresis to ensure RNA integrity after biotinylation (Figure 3.2.5A). However, the presence of MSI2 bound to the RNA bait failed to be detected. To rule out the possibility of buffer interferences, detection of nsP3 was also performed in the pulldown sample. Similarly, the result of nsP3 in the output sample was barely detectable even with low brightness and high contrast (Figure 3.2.5B). This result did not provide the expected consistency with that from the preliminary work, which used tandem mass tag mass spectrometry (TMT-MS) to show that both MSI2 and nsP3 interacted with the 5' end of CHIKV genomic RNA as described in 1.4.2. The obstacle raised from the incompatibility of this assay was unfortunately unable to circumvent in the context of this study. Therefore, an alternative method, Co-IP, was used to establish the rationale for investigating the role of MSI2 in CHIKV replication.

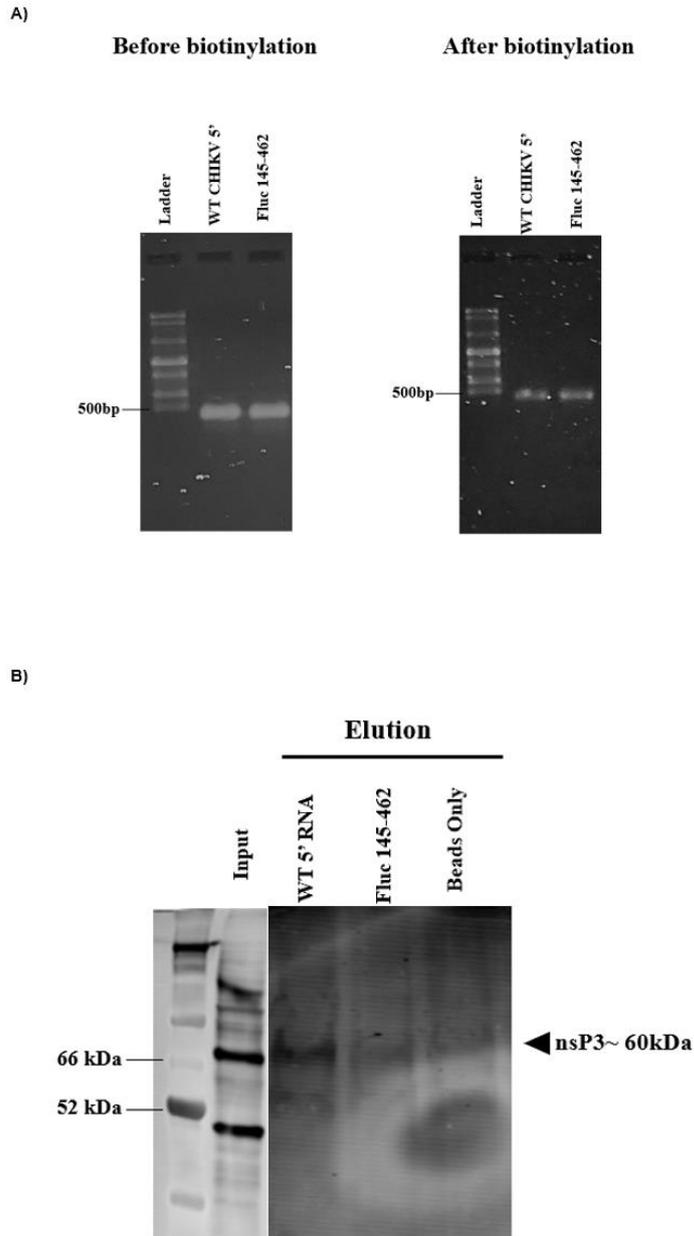


Figure 3.2.5. Biotinylation of CHIKV 5' RNA for pulldown using CHIKV-infected cell lysates.

A). The integrity of the RNA baits was checked by denaturing MOPs gel electrophoresis before and after 3' end biotinylation. **B).** The proteins pulled down from the infected cell lysates by CHIKV 5' RNA were probed against the nsP3 antibody. Except for input, the brightness and contrast for the right half of the membrane containing elution samples were adjusted to reveal the barely visible nsP3 bands.

It has been shown that cytoplasmic relocalisation of host proteins by interacting with nsP2 or nsP3 is essential during infection (241, 275). For example, DHX9, an RNA helicase required for nucleic acid unwinding, is recruited to the CHIKV replication complex by the HVD of nsP3 to specifically facilitate virus translation (106). Hence, MSI2 is speculated to be recruited to the site of virus genome replication via interactions with CHIKV nsPs. According to the list of proteins identified from the preliminary RNA affinity purification and TMT-MS proteomic analysis, nsP1 and nsP3 were significantly enriched alongside MSI2. I therefore speculated if MSI2 had a direct or indirect association with these two viral nsPs. RD cells were infected at an MOI of 10 to acquire CHIKV-infected cell lysates, and the immunoprecipitation of nsP1 and nsP3 was detected by western blot. As expected, both nsP1 and nsP3 were detected following Co-IP by using MSI2 antibody-bound beads (Figure 3.2.6, top panel). To further confirm the results, reverse Co-IP was performed by capturing MSI2 with nsP1 or nsP3 antibody-bound beads. Consistently, MSI2 was found to be associated with both nsPs (Figure 3.2.6, middle panel). However, because whole cell lysates were used in this experiment, it remained unknown that whether the interaction between MSI2 and the two nsPs was independent of other factors, or required intermediate proteins to link them together. Therefore, these results indicate that during CHIKV replication MSI2 either directly interacts with nsP1 and nsP3, or alternatively MSI2 indirectly interacts with nsP1 and nsP3 via other host and/or virus proteins. Either way, these results provided evidence that MSI2 could potentially be hijacked by viral proteins and relocalised to the replicase complex.

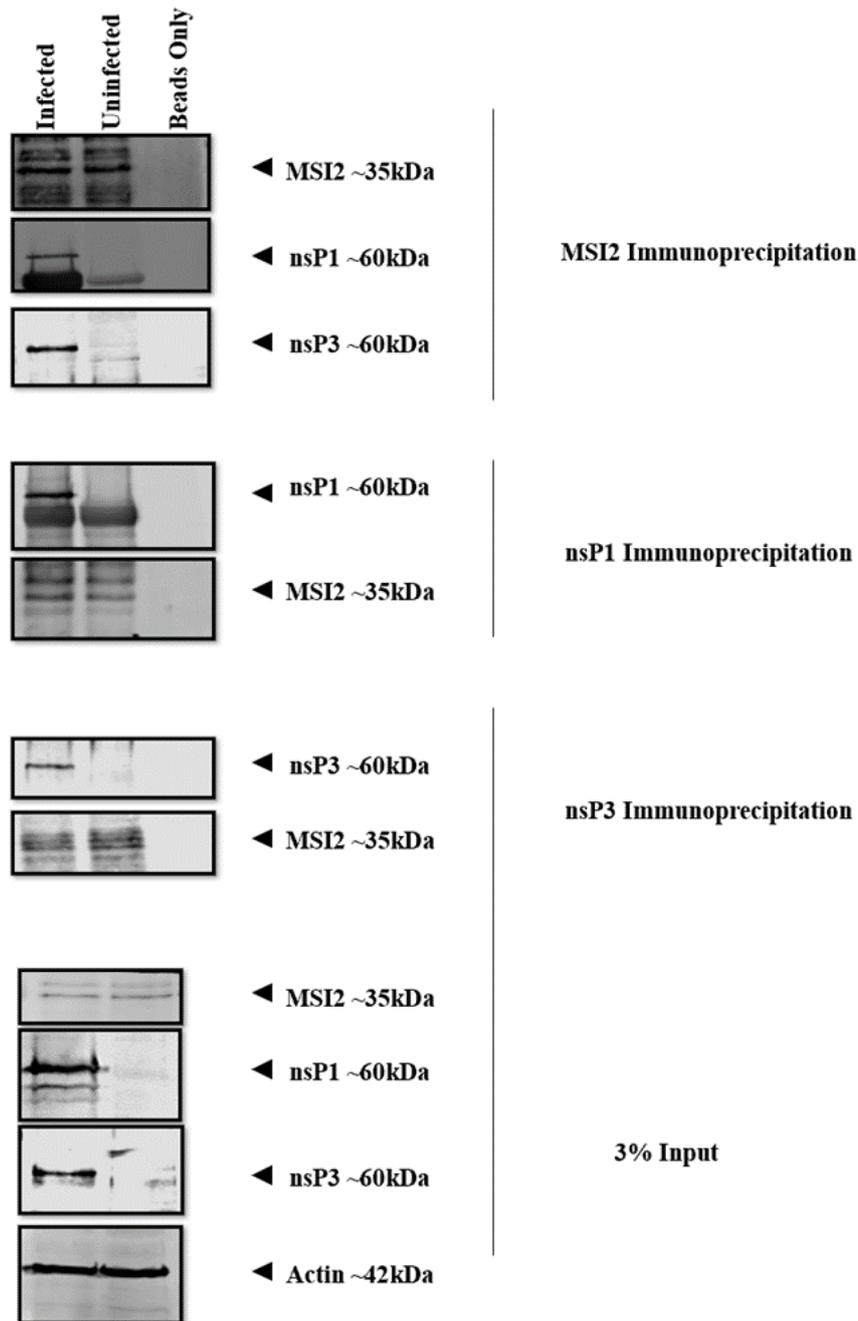


Figure 3.2.6. Forward and reverse co-immunoprecipitation of MSI2 with nsP1 and nsP3 using CHIKV infected RD cell lysates.

Top: nsP1 and nsP3 were positively precipitated from CHIKV infected RD cell lysate by MSI2. Middle: MSI2 was reverse precipitated from the infected cell lysates by nsP1 and nsP3, respectively. Bottom: 3% input samples indicating the expression of MSI2 in both infected (nsP1 and nsP3 bands) and uninfected cells (blank).

3.3 Discussion

In this chapter, preliminary results were acquired indicating the necessity of MSI2 for efficient CHIKV replication. Ro inhibition of the two RNA binding domains of MSI2 severely affected CHIKV production in human cells (Figure 3.2.3A). Although Ro is a known inhibitor of MSI2 with literature support, the current analysis did not biochemically show that MSI2 RNA-binding activity was inhibited due to Ro-treatment. An RNA pulldown experiment using a consensus target RNA of 'GUA' repeats, as stated in a previous study, in the presence and absence of Ro would be useful to confirm the inhibition of MSI2-RNA interaction (276). Additionally, a drug response curve over time should have also been conducted to show the timescale for the effectiveness of Ro on MSI2 inhibition. Nonetheless, Ro treatment inhibited Huh7 cells from producing progeny viruses at 8 h but not 24 h post infection, indicating that compared to RD cells, Huh7 cells might be more sensitive to *in vitro* applications. Results from the sub-genomic replicon and *trans*-complementation assay, together with the genomic copies of both sense strand RNA, further demonstrated that the MSI2 RNA binding activity enhanced CHIKV genome replication (Figure 3.2.3B & C and Figure 3.2.4). In addition, Co-IP with CHIKV infected cell lysates showed that MSI2 might directly interact with nsP1 and nsP3 or indirectly via each other or other host and virus proteins (Figure 3.2.6). The recruitment of the cytoplasmic MSI2 by nsPs constituted a promising starting point for investigating the requirement of MSI2 in CHIKV replication. Given that MSI2 was affinity purified using the CHIKV 5' RNA as bait, our data evoked the hypothesis of essential protein binding sites within the 5' UTR. While Ro has been shown to be highly specific to MSI2, it did not account for any off-target effects. Biological inhibitors might possess potential off-targets effects on other cellular and viral proteins, leading to the over/underestimation of cell viability which would affect result interpretation (277-279). Therefore, these results provided a convincing rationale for the experiments using MSI2-specific RNA silencing described in Chapter 4.

3.3.1 Limitations of RNA pulldown and co-immunoprecipitation

The extremely low level of RNA-bind protein recovery in our RNA pulldown was unexpected (Figure 3.2.5B). Multiple optimisations were attempted with longer incubation time, adjusted salt concentration of buffers, different antibodies and increased concentration of cell lysates. Unfortunately, none of them seemed to produce quality results. It remained a possibility that although the integrity of RNA after biotinylation was acceptable (Figure 3.2.5A), subtle RNA degradation occurred during the incubation stage presumably due to unsuitable buffer composition. On the other hand, if the RNA was intact, the pulldown may also biasedly favour the more abundant proteins present in the cell lysate rather than our target MSI2 (280). Therefore, further studies are suggested to use RNA immunoprecipitation to 'capture' RNA by protein, which circumvents the issue with quantity as RNA can be specifically amplified via reverse transcription and qPCR.

For Co-IP, it was inevitable that some bands appear to be messy due to concurrent elution of antigens and antibodies, which caused band overlapping in western blot (281). In this study, as proteins were precipitated from whole cell lysates, I was unable to deduce the exact mechanism of interaction between MSI2 and nsP1 and nsP3. The interaction might be direct, or there might be intermediate factors which link MSI2 to CHIKV nsPs. For example, DHX9 can directly bind to nsP2 and nsP3 independent of other CHIKV proteins or genomic RNA (106). It is plausible that such direct binding partners could also interact with MSI2, thereby diverting it to facilitate virus replication. The HVD of nsP3, owing to its proline-rich motif, is known for interacting with several host proteins, including G3BP1/2 and nucleosome assembly protein 1-like 1/2 (84). Therefore, co-transfection of plasmids expressing tagged MSI2 and different nsPs would be useful to examine whether their interactions are direct. Moreover, further experiments detecting immunofluorescence-labelled MSI2, nsPs and dsRNA with confocal microscopy would reveal the subcellular localisation of MSI2 to the CHIKV replication complex.

3.3.2 Issues with the incompatibility of Huh7 with the *trans*-complementation systems

The precise reason for the inability of the *trans*-complementation system to produce interpretable results in Huh7 cells was uncertain. RD and Huh7 cells are both the targets for CHIKV, and infections of Huh7 cells have always been successful. For studies concerning the sub-genomic replicon of the hepatitis C virus, it has been reported that the limitation of HCV inhibition in Huh7 cells resulted from their low resistance to the IFN- γ signalling pathway and high sensitivity to compounds with antiviral effects (282). Consistently, Huh7 was more sensitive to Ro than RD in our MTT assay (Figure 3.2.2). Despite the dosage of Ro being kept at a non-lethal level, the cytotoxicity might affect factors and pathways associated with cell cycle progression, on which the replication of sub-genomic replicons depends (283). Ro treatment could potentially downregulate the level of nucleoside triphosphate pools in Huh7 cells, which, in turn, limited its permissiveness to CHIKV replication (284). The exact effect of Ro on normal cellular processes will require further investigation. In the meantime, substitutive cell lines such as human hepatoma cells LH86 can be used to test the efficiency and compatibility of CHIKV replicon replication (285). Nonetheless, for data comparability and accurate interpretation, Huh7 cells were excluded from all replicon assays in this study.

CHAPTER 4: PHENOTYPIC AND REVERSE GENETIC ANALYSIS OF THE ROLE OF MSI2 in CHIKV GENOME REPLICATION USING RNA SILENCING

4.1 Introduction

4.1.1 MSI2-specific RNA interference

While the data from chapter 3 were consistent with MSI2 playing a role in CHIKV genome replication, the inhibition of MSI2 function depended on the small molecule inhibitor Ro, which might have off-target effects as discussed in 3.3. To avoid that, MSI2-specific RNA interference (RNAi) was used to investigate the role of MSI2 in CHIKV genome replication. In most eukaryotic cells, RNAi is triggered by dsRNA molecules to initiate gene silencing via mRNA cleavage or translation repression (286). Non-coding RNAs known as small interfering RNAs (siRNAs) specifically regulate the endonucleolytic cleavage of a certain mRNA (287). Dicer is an RNase III endonuclease which produces the siRNAs by cleavage of cytosolic dsRNA (288). The siRNAs are then associated with a dsRNA-binding protein known as PACT, which interacts with a multi-protein complex containing Dicer, TAR-RNA-binding protein (TRBP) and Argonaute 2 (AGO2) (289). These proteins regulate the transfer of siRNAs to the RNA-induced silencing complex (RISC), where the siRNAs are further cleaved by AGO2 to break the double-stranded bond and liberate the 'passenger' strand. The single-stranded 'guide' strand remained in the RISC leads to its activation to regulate the specific target recognition and via complementary base-pairing by the 'guide' strand (290). Commercial siRNAs targeting specific cellular genes are synthesised with chemically modified nucleobases to increase stability, which is a common issue with *in vitro* RNA (291). Furthermore, the choice of transfection reagents should also be cautiously considered due to their different efficacy in different cell types. Comparative screening of transfection reagents showed that Lipofectamine RNAiMAX is the most efficient one

with low cytotoxicity (292). Therefore, this reagent was used for all siRNA transfections into RD and Huh7 cells performed in this study.

MSI2-encoding genes have been found to be extensively post-transcriptional modified, leading to the formation of various isoforms with different molecular weights (256). This can be confirmed by our western blot result in Figure 3.2.1, where a cluster of bands appeared around the estimated molecular weight (~35 kDa) of MSI2. These bands were specifically reactive to the monoclonal MSI2 antibody, suggesting that they shared the same epitope. Therefore, efficient downregulation of MSI2 requires siRNAs ideally targeting as many isoforms as possible. In this study, a pool of five commercial MSI2-specific siRNAs were used to assess the effect of MSI2 knockdown on CHIKV replication.

To obtain parallel data from independent knockdown approaches, MSI2 downregulation was also achieved by introducing MSI2 short hairpin RNAs (shRNAs) to cells via lentiviral vectors. Unlike siRNAs, shRNAs are delivered via lentivirus transduction to the nucleus where they are integrated to the cellular genome (293). Components of lentiviral vectors harbouring the target gene of knockdown were first assembled in permissive HEK 293T progenitor cells before transduction to the cell of interest. Once successful lentivirus infection is established, the siRNA will be stably synthesised and processed from the shRNA to achieve constant knockdown of target protein expression (294). The lentivirus contains a puromycin-resistance gene, which can be used for the selective culture of shRNA-positive cells. This technique would provide a useful validation to the results from siRNA transfection, as the data of phenotypic analysis from these two different approaches were expected to be comparable. Similar to siRNA, a pool of five commercial lentivirus vector plasmids encoding MSI2-specific shRNAs were used to assess the effect of MSI2 knockdown on CHIKV replication.

4.1.2 The predicted MSI2 binding sites within the CHIKV 5' UTR

Using systematic evolution of ligands by exponential enrichment, it has been identified that the RRM1 of MSI2 preferentially binds to the sequence motif (G/A) U₁₋₃AGU (295). Analysis of CHIKV nt 1-303 RNA sequence initially used as the bait for RNA affinity purification and proteomic analysis, identified a potential MSI RNA binding site (₆₃AUUAA₆₇) at nucleotide positions 63-67 – a single stranded region located between SL47 and the AUG start codon (83) (Figure 4.1.1). MSI2 was hypothesised to participate in CHIKV genome replication by directly binding to this sequence. In order to investigate the effect on CHIKV genome replication when the predicted MSI2-RNA interaction was disrupted, a CHIKV mutant was designed, in which the putative MSI2 binding site was mutated (₆₃AUUAA₆₇ > ₆₃CAACU₆₇), the mutant was designed 'binding site mutant' (BSM). MSI2 BSM was then incorporated into both the CHIKV infectious clone and *trans*-complementation system. The BSM mutant was considered to block MSI2 binding while not interfering with the folding of adjacent RNA structures that have been shown to regulate genome replication of the virus (88). This was confirmed by *in silico* RNA structure prediction by free energy minimisation (Figure 4.1.2).

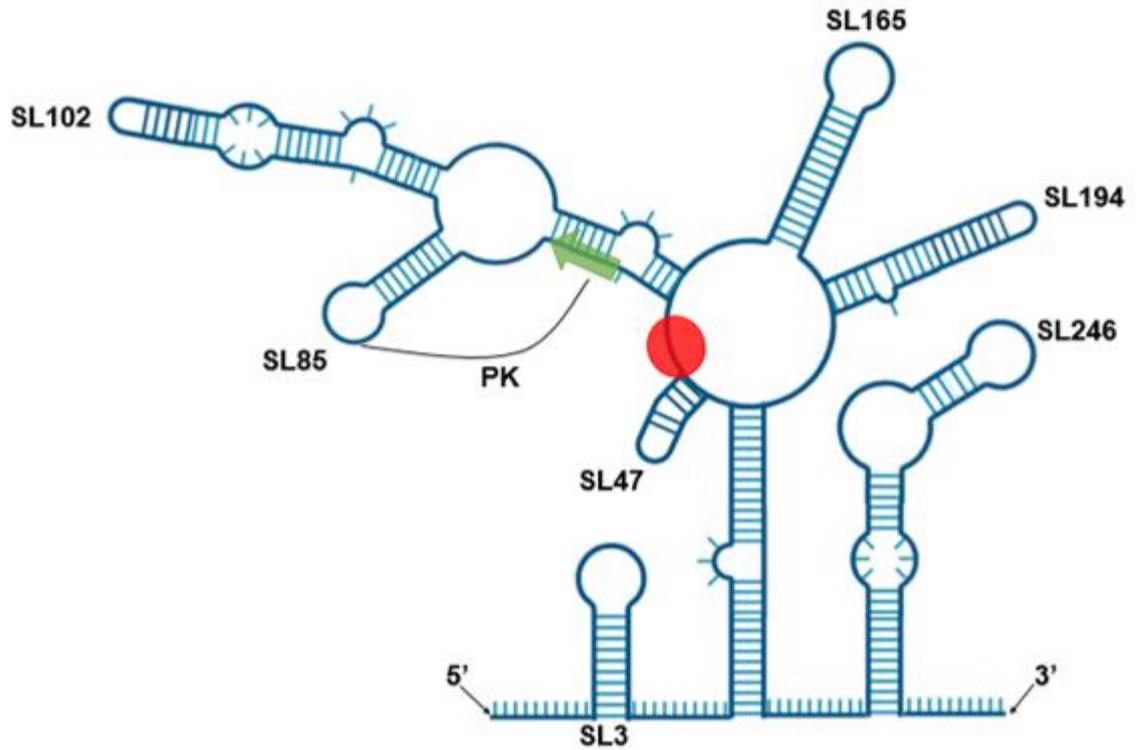


Figure 4.1.1. Schematic representation of the stem-loop structures and site of mutagenesis within the CHIKV 5' end (83).

The AUG start codon of ORF-1 was marked by the green arrow. The speculated MSI2 binding sites were marked by the red circle. The stem-loops were designated SL3, SL47, SL85, SL102, SL165, SL194 and SL246. Pseudoknot = PK.

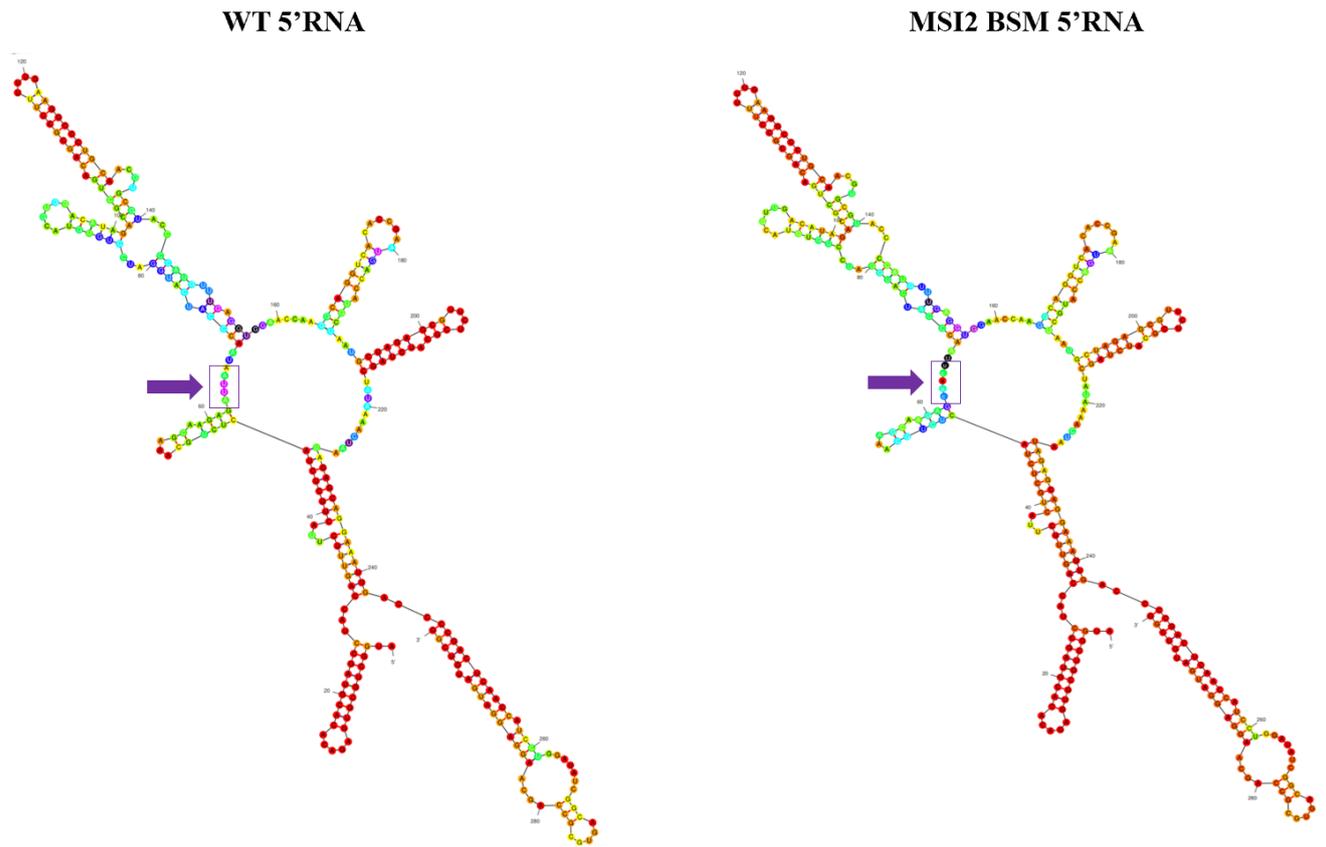


Figure 4.1.2. RNA folding model generated by M Fold predictions.

CHIKV 5' end RNA sequence of WT (left) and MSI2 BSM (right) were input to M Fold algorithm to generate RNA structures with minimised free energy. The mutated sequence was highlighted with arrows and squares. The colour represents the confidence level of the base-pairing condition of each nucleotide. Warm colours indicate the nucleotide is highly likely to be paired/unpaired at its position as predicted by the algorithm. Cool colours indicate the nucleotide is less likely to be paired/unpaired at its position as predicted by the algorithm.

4.2 Results

4.2.1 The effect of MSI 1+2 siRNA knockdown on CHIKV genome replication in Huh7 cells

Despite the potential off-target effect of Ro that might interfere with the interpretation of the results from chapter 3, the inhibition of RRM1 presumably also affected other MSI proteins which possess this motif besides MSI2. Thus, the utilisation of targets-specific siRNA should not be confined solely to MSI2. MSI1, the paralog of MSI2, is involved in the replication of the Zika virus (263, 264). The RNA-binding domains of MSI1 and MSI2 are evolutionarily conserved, thereby suggesting their functional redundancy (259). It gives rise to the arguable possibility that MSI1 participates in CHIKV replication in the same way as MSI2. As MSI1 share the highest (75%) homology with MSI2 compared to other members of the Musashi family, siRNA for both MSI1 and MSI2 were selected to assay their knockdown effects on CHIKV genome replication (256). When MSI1 expression in Huh7 cells was examined by western blot, unlike the expression of MSI2 seen in Figure 3.2.1, MSI1 was displayed as a single band rather than band clusters, indicating that protein isoforms as a result of post-transcriptional regulation events, such as alternative splicing, were more limited for MSI1 than MSI2 (Figure 4.2.1A). Interestingly, RD cells expressed extremely low level of MSI1 that the protein bands were almost undetectable, whereas high level of MSI1 expression was observed in Huh7 cells (Figure 4.2.1A). Therefore, siRNA knockdown involving MSI1 was performed in Huh7 cells only.

While the MSI siRNA was designed to specifically knockdown cellular MSI expression, it was important to optimise its transfection efficiency, confirm the level of MSI knockdown, its timescale and effects on cell viability in in Huh7 and RD cells. MSI knockdown would ideally need to last for 24 h after siRNA transfection, so that virus and replicons could be introduced to 'MSI-free' cells and resultant phenotypes analysed. Therefore, 50 pmol of MSI siRNA was transfected 24 h after cell seeding and evaluated the knockdown efficiency 24 h post siRNA transfection. A non-targeting

scrambled siRNA was used as a negative control to preclude siRNA off-target effects. The results clearly demonstrated that both MSI1 and MSI1+2 siRNA treatments led to the efficient knockdown of MSI1 expression, measured by western blot in Huh7 cells, at 24 hrs post siRNA transfection (Figure 4.2.1B). Similarly, MSI2 siRNA treatment led to the efficient knockdown of MSI2 in both cell lines, indicating that the cluster of bands indeed represented different MSI2 isoforms (Figure 4.2.1C). Notably, cell morphology was constantly monitored during the transfection and no morphological changes or cytotoxicity was observed. This can be further justified by the same expression level of the actin control for each treatment, which indicated that the efficient knockdown was achieved when the same quantity of total protein lysate was used. Therefore, this siRNA approach was used for all further siRNA knockdown experiments.

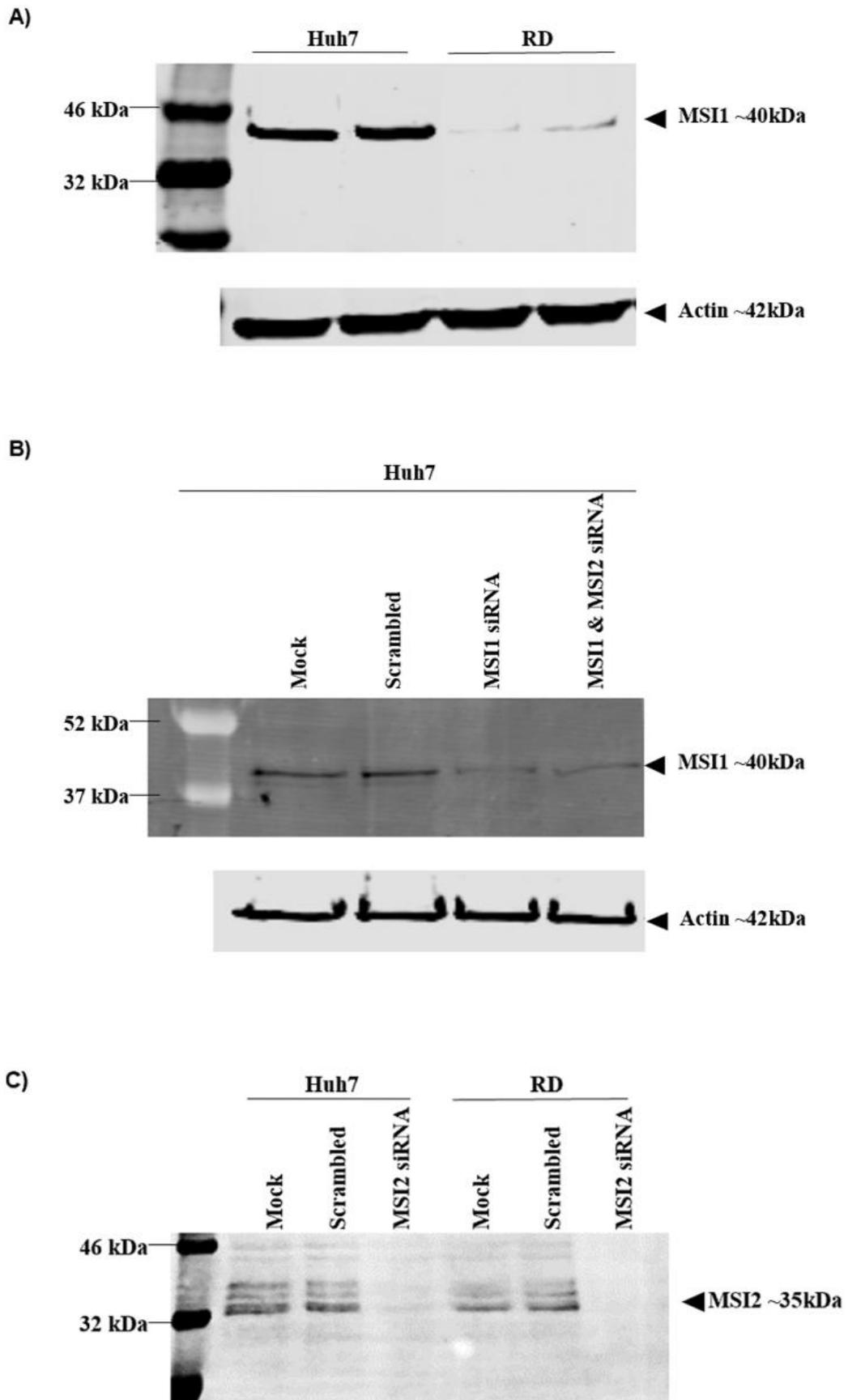


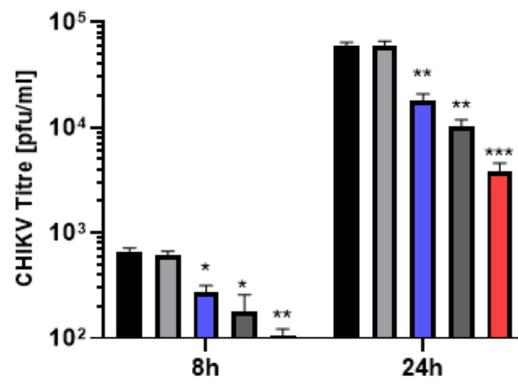
Figure 4.2.1. Evaluation of the efficiency of MSI knockdown by western blot in Huh7 and RD cells 24 h after siRNA transfection.

A). The expression of MSI1 in Huh7 and RD cells. **B).** The level of MSI1 expression in Huh7 cells was compared by mock, scrambled siRNA, MSI1 siRNA and MSI1+2 siRNA treatments. **C).** The level of MSI2 expression in Huh7 and RD cells was compared by mock, scrambled siRNA and MSI2 siRNA treatments. Actin bands indicate the equal quantity of protein samples used for each treatment.

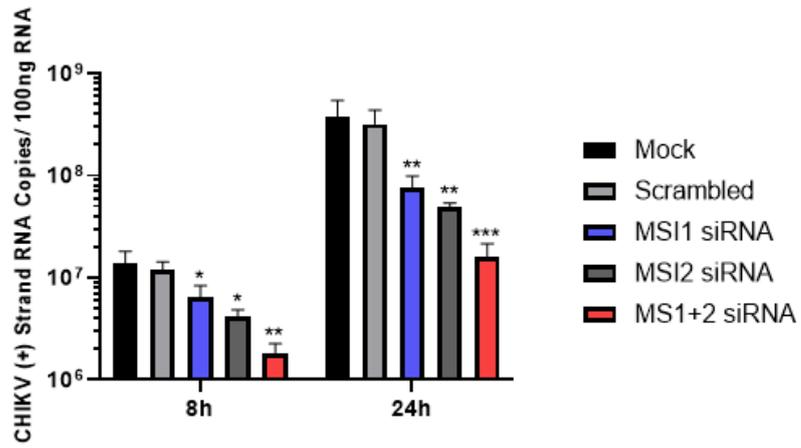
After the optimisation of siRNA transfection, the effect of MSI1, MSI2 and MSI1+2 siRNA knockdown on CHIKV genome replication in Huh7 cells was assayed by ICRES infection. Huh7 cells of each treatment were infected by CHIKV at MOI of 0.1. Progeny virus at early (8 h) and late (24 h) time points were collected and proactive CHIKV replication measured by plaque assay in BHK-21 cells. The genomic copies of intracellular positive- and negative-sense ICRES RNA were quantified by qRT-PCR. For infectious virus production at 8 and 24 h post infection, MSI1 siRNA treatment led to a significant ~ 0.4 -log fold decrease. MSI2 siRNA treatment led to a significant ~ 0.7 -log fold decrease, whereas MSI1+2 siRNA treatment led a significant ~ 1.1 -log fold decrease (Figure 4.2.2A). For the genomic copies of both positive- and negative-sense RNA at 8 h post infection, MSI1 siRNA treatment resulted in a significant ~ 0.3 -log fold reduction; MSI2 siRNA treatment resulted in a significant ~ 0.6 -log fold reduction, whereas MSI1+2 siRNA treatment resulted in a significant ~ 1.1 -log fold reduction. At 24 h post infection, MSI1 siRNA treatment resulted in a significant ~ 0.6 -log fold reduction; MSI2 siRNA treatment resulted in a significant ~ 0.8 -log fold reduction, whereas MSI1+2 siRNA resulted in a significant ~ 1.4 -log fold reduction. Despite the strong inhibition of CHIKV replication, Huh7 cells treated with MSI1+2 siRNA were still able to maintain a low level of CHIKV replication. This implied that either there was functional compensation from other Musashi proteins, or MSI1 and MSI2 were required but not indispensable for CHIKV replication. Therefore, these results demonstrated that both MSI1 and MSI2 were required for CHIKV genome replication in Huh7 cells. However, MSI2, but not MSI1, was detected in the proteomic screening

from the preliminary RNA affinity purification. Comparison of the level of replication repression in Figure 4.2.2 also indicated that CHIKV replication was more severely impaired when MSI2 expression was downregulated. Therefore, the following assays were conducted using RNAi of MSI2 in RD cells.

A)



B)



C)

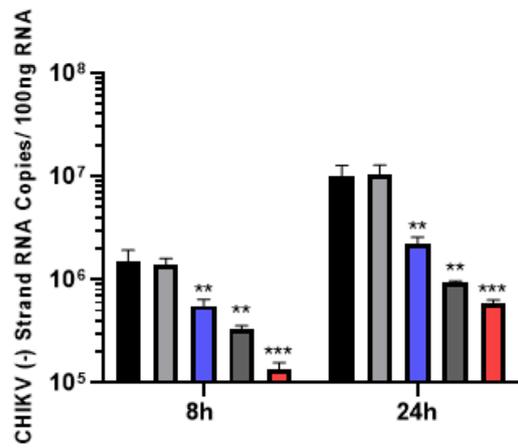


Figure 4.2.2. Comparison of the effect between MSI1, MSI2 and MSI1+2 siRNA knockdowns on CHIKV replication in Huh7 cells.

Phenotypic analysis was conducted regarding **A).** infectious virus production, the genomic copies of **B).** positive- and **C).** negative-sense strand RNA. The average of three independent experiments ($n=3$) is shown. Error bars represent standard deviation from the mean. *, $P<0.05$; **, $P<0.01$; ***, $P<0.001$ (One-way ANOVA).

4.2.2 MSI2 positively facilitates CHIKV replication in RD cells

To further validate the requirement of MSI2 for CHIKV genome replication in RD cells, the effect of MSI2 siRNA knockdown was assayed by ICRES infection. RD cells treated with mock, scrambled siRNA and MSI2 siRNA were infected by ICRES at MOI of 0.1. The infected cells were lysed at 24 h post infection and CHIKV protein expression was assayed by western blot. Progeny virus at early (8 h) and late (24 h) time points were collected and proactive CHIKV replication measured by plaque assay in BHK-21 cells. The genomic copies of intracellular positive- and negative-sense ICRES RNA were quantified by qRT-PCR. Following densitometric analysis using the LICOR software, the expression of nsP1 and capsid in MSI2 siRNA treated cells was reduced by 38% and 47% compared to mock, respectively (Figure 4.2.3A). Similar to Huh7 cells, MSI2 siRNA knockdown in RD cells led to a significant ~ 0.5 -log and ~ 1.2 -log fold decrease in infectious virus production compared to mock at 8 h and 24 h post infection, respectively (Figure 4.2.3B). For the genomic copies of positive-sense RNA, MSI2 siRNA knockdown in RD cells led to a significant ~ 0.8 -log and ~ 1.0 -log fold reduction compared to mock at 8 h and 24 h post infection, respectively (Figure 4.2.3C). For the genomic copies of negative-sense RNA, MSI2 siRNA knockdown in RD cells led to a significant ~ 0.3 -log and ~ 1.0 -log fold reduction compared to mock at 8 h and 24 h post infection, respectively (Figure 4.2.3D). Therefore, these results suggested that MSI2 was also required for CHIKV genome replication in RD cells. However, in order to obtain the parallel results analysis using sub-genomic replicon and the *trans*-complementation system as shown in Figure 3.2.3B&C, RD cells treated with MSI2 siRNA would need to be processed with an additional transfection of replicons. This

might damage the cells due to the cytotoxicity of transfection reagents. Consequently, MSI2 shRNA treated RD cells were used to maintain cell viability via a single transfection of replicons.

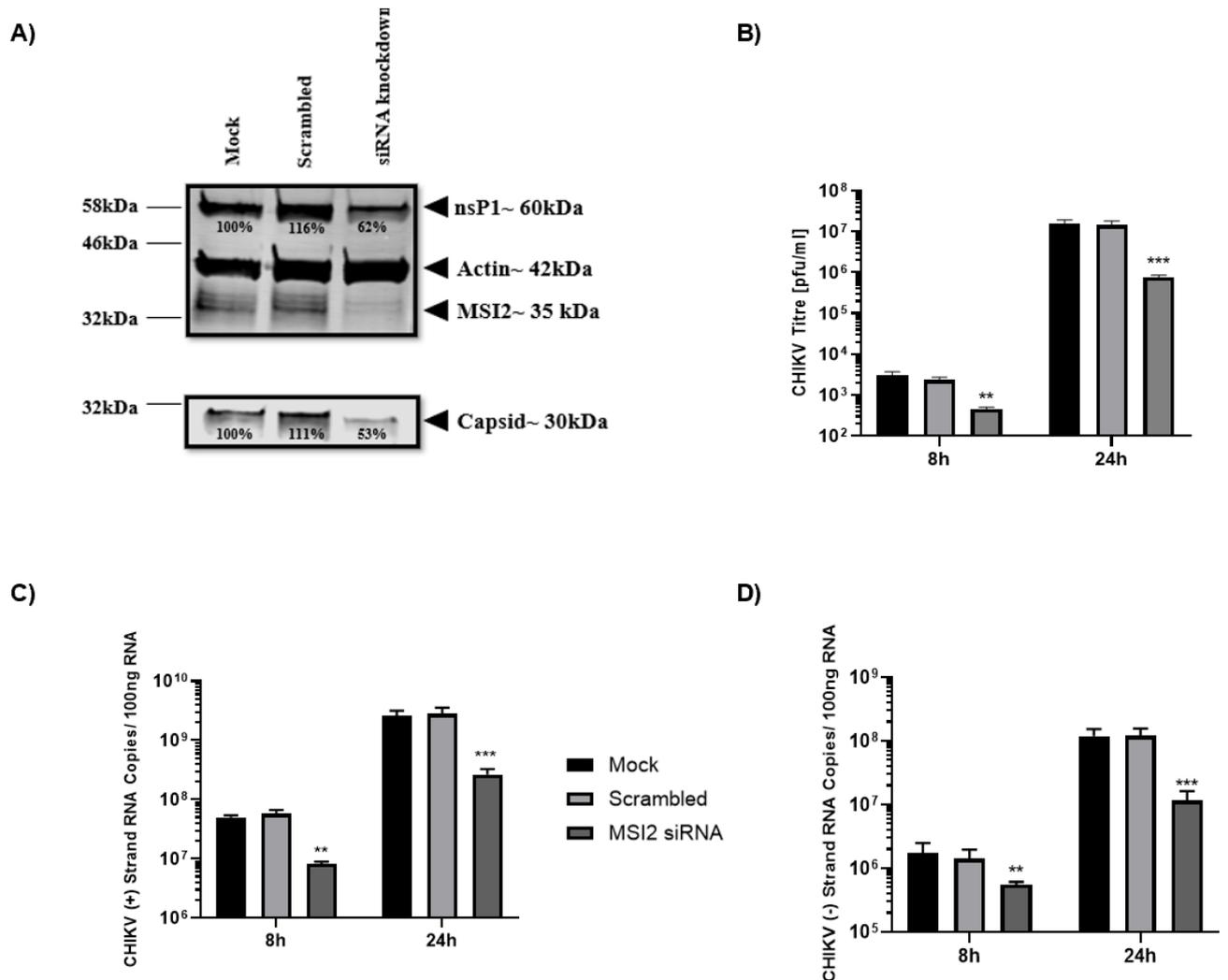


Figure 4.2.3. Replication phenotype of MSI2 siRNA knockdown compared to mock and scrambled samples in RD cells.

Phenotypic analysis was conducted regarding **A)** Virus protein expression at 24 h post infection. Numbers in percentage below the bands indicate their relative intensity compared to mock; **B)** Virus production; **C)** Genomic copies of CHIKV positive-sense strand RNA; **D)** Genomic copies of CHIKV negative-sense strand RNA. The average of

three independent experiments ($N=3$) is shown. Error bars represent standard deviation from the mean. *, $P<0.05$; **, $P<0.01$; ***, $P<0.001$ (One-way ANOVA).

4.2.3 Downregulation of MSI2 by shRNA impaired CHIKV genome replication in RD cells

As an independent approach, MSI2 shRNA knockdown did not require transfection since the shRNA was transduced into cells through lentivirus. Persistent downregulation of MSI2 by shRNA in Huh7 cells has been unsuccessful due to excessive cell death after transduction, hence only RD cells were used for MSI2 shRNA experiment. Components of the lentiviral particles, including the ones containing MSI2 shRNA and puromycin-resistant genes, were transfected to HEK 293T cells for lentivirus production, which were used to infect RD cells. RD cells were puromycin treated to select those which were successfully transduced. Western blot analysis showed that three different passages of MSI2 shRNA⁺ RD cells were all able to establish a consistent downregulation of MSI2 (Figure 4.2.4A). Given that the level of MSI2 knockdown was comparable to siRNA treatment even at 24 h post ICRES infection as shown in Figure 4.2.3A, MSI2 shRNA treatment was assumed to produce similar phenotypes from ICRES infection.

To confirm that, mock, scrambled shRNA⁺ and MSI2 shRNA⁺ RD cells were infected by ICRES at MOI of 0.1. Progeny virus at early (8 h) and late (24 h) time points were collected and proactive CHIKV replication measured by plaque assay in BHK-21 cells. The genomic copies of intracellular positive- and negative-sense ICRES RNA were quantified by qRT-PCR. Comparable to siRNA treatment, infectious virus production was significantly decreased by ~ 1.3 -log and ~ 1.1 -log fold in MSI2 shRNA⁺ RD cells compared to mock at 8 h and 24 h post infection, respectively (Figure 4.2.4B). For the genomic copies of positive-sense RNA, the number of copies was significantly decreased by ~ 1.1 -log and ~ 1.2 -log fold in MSI2 shRNA⁺ RD cells compared to mock at 8 h and 24 h post infection, respectively (Figure 4.2.4C). For the genomic copies of

negative-sense RNA, the number of copies was significantly decreased by ~ 0.8 -log and ~ 1.3 -log fold in MSI2 shRNA⁺ RD cells compared to mock at 8 h and 24 h post infection, respectively (Figure 4.2.4D). Noteworthy, the fold change of shRNA-treated samples was slightly greater than siRNA-treated ones shown in 4.2.2. This suggested that despite the convincing demonstration of phenotypes by MSI2 siRNA, the long-lasting MSI2 knockdown by shRNA was more efficient than the transient siRNA that the inhibition of CHIKV replication was increased. Therefore, these results not only showed that MSI2 shRNA was more efficient at downregulating the expression of MSI2 than siRNA, but also further validated the requirement of MSI2 for CHIKV replication.

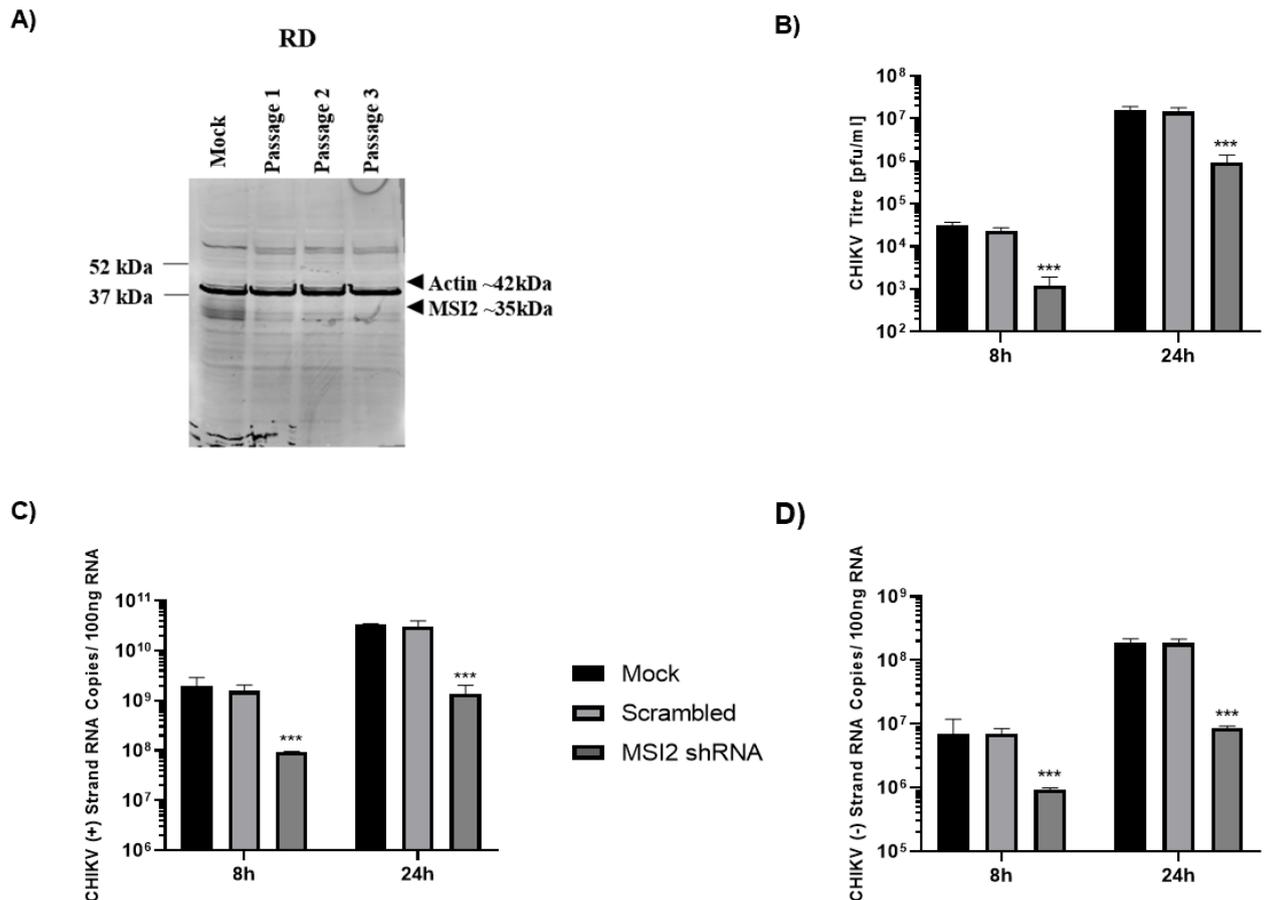


Figure 4.2.4. Replication phenotype of MSI2 shRNA knockdown compared to mock and scrambled samples in RD cells.

A). Western blot showing the efficient and steady knockdown of MSI2 in different RD passages after being transduced with MSI2 shRNA-lentivirus. **B).** The effect of MSI2 shRNA knockdown analysed by virus production. **C).** The effect of MSI2 shRNA knockdown was analysed by genomic copies of CHIKV positive-sense strand RNA. **D).** The effect of MSI2 shRNA knockdown was analysed by genomic copies of CHIKV negative-sense strand RNA. The average of three independent experiments (N=3) is shown. Error bars represent standard deviation from the mean. *, P<0.05; **, P<0.01; ***, P<0.001 (One-way ANOVA).

Next, MSI2 shRNA treated RD cells were used for the sub-genomic replicon transfection to confine the role of MSI2 to CHIKV genomic translation and replication.

RD cells of mock, scrambled shRNA⁺ and MSI2 shRNA⁺ were transfected with 1 µg dual-luciferase sub-genomic replicon. The luciferase samples were obtained from lysing the monolayer for bioluminescence analysis at 8 and 24 h post transfection. The expression of the genomic ORF-1 was assayed by the *Renilla* luciferase signal to represent translation, while the expression of the sub-genomic ORF-2 was assayed by the firefly luciferase signal to represent replication. For translation, the *Renilla* luciferase signal was significantly reduced by ~0.5-log and ~0.7-log fold in MSI2 shRNA⁺ RD cells compared to mock at 8 h and 24 h post transfection, respectively. For replication, the Firefly luciferase signal was significantly reduced by ~0.5-log and ~0.8-log fold in MSI2 shRNA⁺ RD cells compared to mock at 8 h and 24 h post transfection, respectively (Figure 4.2.5A). Therefore, the result demonstrated that MSI2 downregulation significantly impaired CHIKV replication and/or translation, as the reduction in luciferase signals could be resulted from either impaired translation of nsPs, or impaired replication of genomic and sub-genomic RNA.

To further narrow the role of MSI2 down to solely CHIKV genome replication, *trans*-complementation system was also employed which uncouples the translation of viral replicase from genomic RNA replication. A schematic representation of the *trans*-complementation system can be found in Figure 2.2B. In this system, the level of CHIKV genome replication was assayed by the signal of the firefly luciferase, while the level of CHIKV sub-genome replication was assayed by the signal of gaussia luciferase. For genome replication, the firefly luciferase signal was significantly reduced by ~0.9-log and ~1.0-log fold in MSI2 shRNA⁺ RD cells compared to mock at 8 h and 24 h post transfection, respectively. For sub-genome replication, the gaussia luciferase signal was significantly reduced by ~1.1-log and ~0.8-log fold in MSI2 shRNA⁺ RD cells compared to mock at 8 h and 24 h post transfection, respectively (Figure 4.2.5B). Therefore, this result demonstrated that both CHIKV genomic and sub-genome replication were impaired due to MSI2 knockdown in RD cells. However, as the genomic copies of negative-sense strand RNA was significantly decreased in MSI2 siRNA-treated Huh7 (Figure 4.2.2C) and RD (Figure 4.2.3D), as well as MSI2 shRNA⁺ RD

cells (Figure 4.2.4D), it was proposed that MSI2 predominantly participates in CHIKV genome replication.

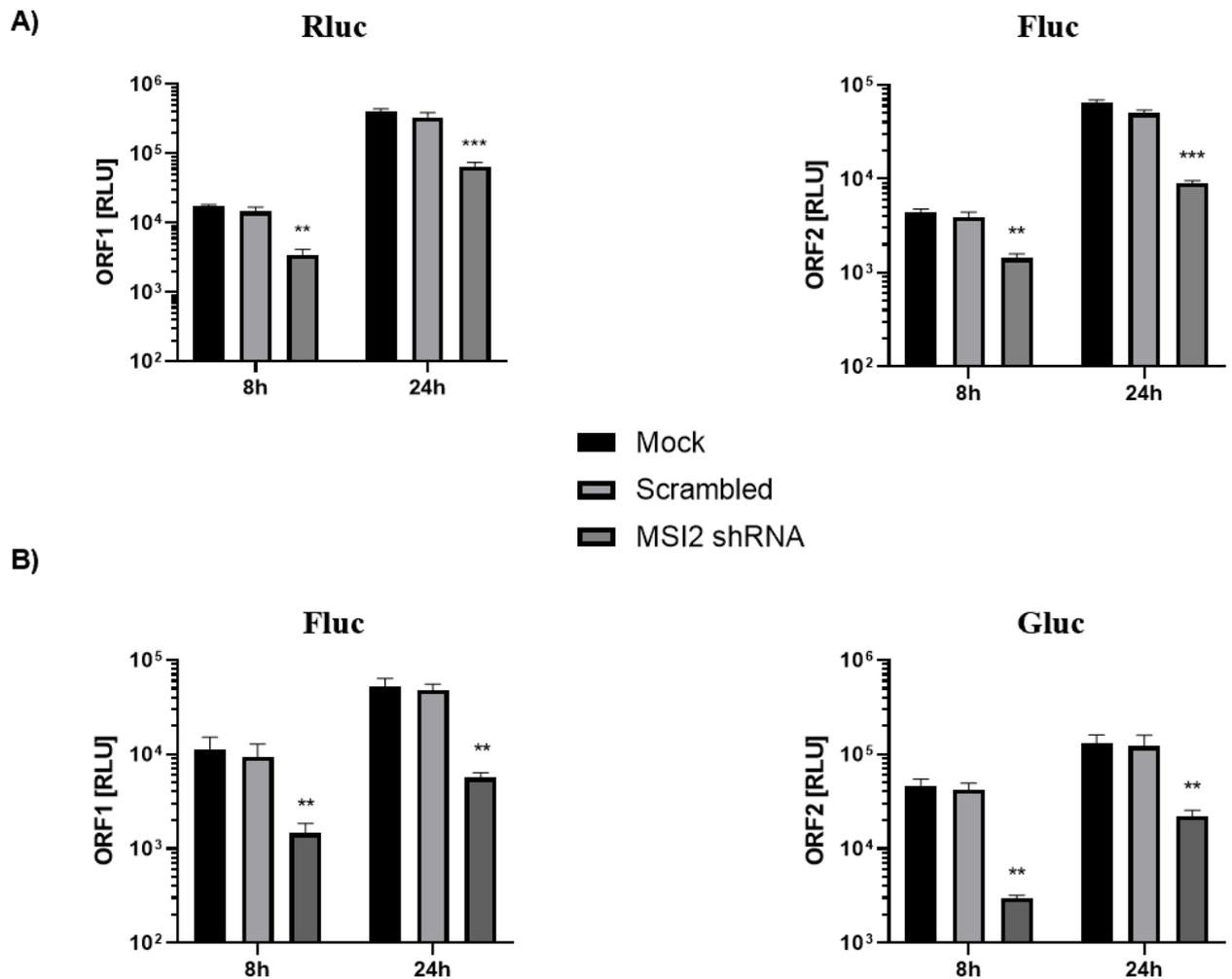


Figure 4.2.5. MSI2 is required for the efficient CHIKV genome replication in RD cells.

A). The expression of ORF-1 and -2 following MSI2 shRNA knockdown was measured by the dual-luciferase subgenomic replicon system at 8 h and 24 h post transfection.

B). The expression of ORF-1 and -2 following MSI2 siRNA knockdown was measured by the trans-complementation replicon system at 8 h and 24 h post transfection. The average of three independent experiments (n=3) is shown. Error bars represent standard deviation from the mean. *, P<0.05; **, P<0.01; ***, P<0.001 (One-way ANOVA).

4.2.4 Mutagenesis of predicted MSI2 binding sequence on CHIKV genome significantly impaired virus replication in RD cells

Since the results from 4.2.2 and 4.2.3 demonstrated that MSI2 was required for CHIKV genome replication, the predicted MSI2-RNA interaction was presumably vital for this process. MSI2 has the propensity to bind to the sequence motif (G/A) U₁₋₃AGU (257), which can be found at the single-stranded sequence nt 63-67 upstream of SL47. This region was therefore mutated (₆₃AUUAA₆₇ > ₆₃CAACU₆₇), and cloned back to the full-virus, the *trans*-complementation system. This mutation would impede the MSI2-RNA interaction and consequently inhibit CHIKV genome replication. Wild-type ICRES and MSI2 BSM RNA were electroporated into BHK-21 cells. Progeny virus produced after 24 h were collected and proactive CHIKV replication measured by plaque assay in BHK-21 cells. Unexpectedly, the mutant virus was absolutely lethal for full-virus production (Fig. 4.2.6A). For the *trans*-complementation assay, it was also very surprising to observe the dramatic contradiction that for genome replication, the firefly luciferase signal was only significantly reduced by ~0.1-log and ~0.3-log fold in MSI2 BSM compared to WT at 8 h and 24 h post transfection, respectively. For sub-genome replication, the gaussia luciferase signal was only significantly reduced by ~0.2-log and ~0.3-log fold in MSI2 BSM compared to WT at 8 h and 24 h post transfection, respectively (Fig. 4.2.6B). These results were consistent with the prediction that region nt 63-67 of the CHIKV 5' UTR indeed being the binding site for MSI2, and showed that its primary sequence was required for CHIKV genome replication. However, they did not explain why MSI2 BSM displayed lethal phenotype for infectious virus.

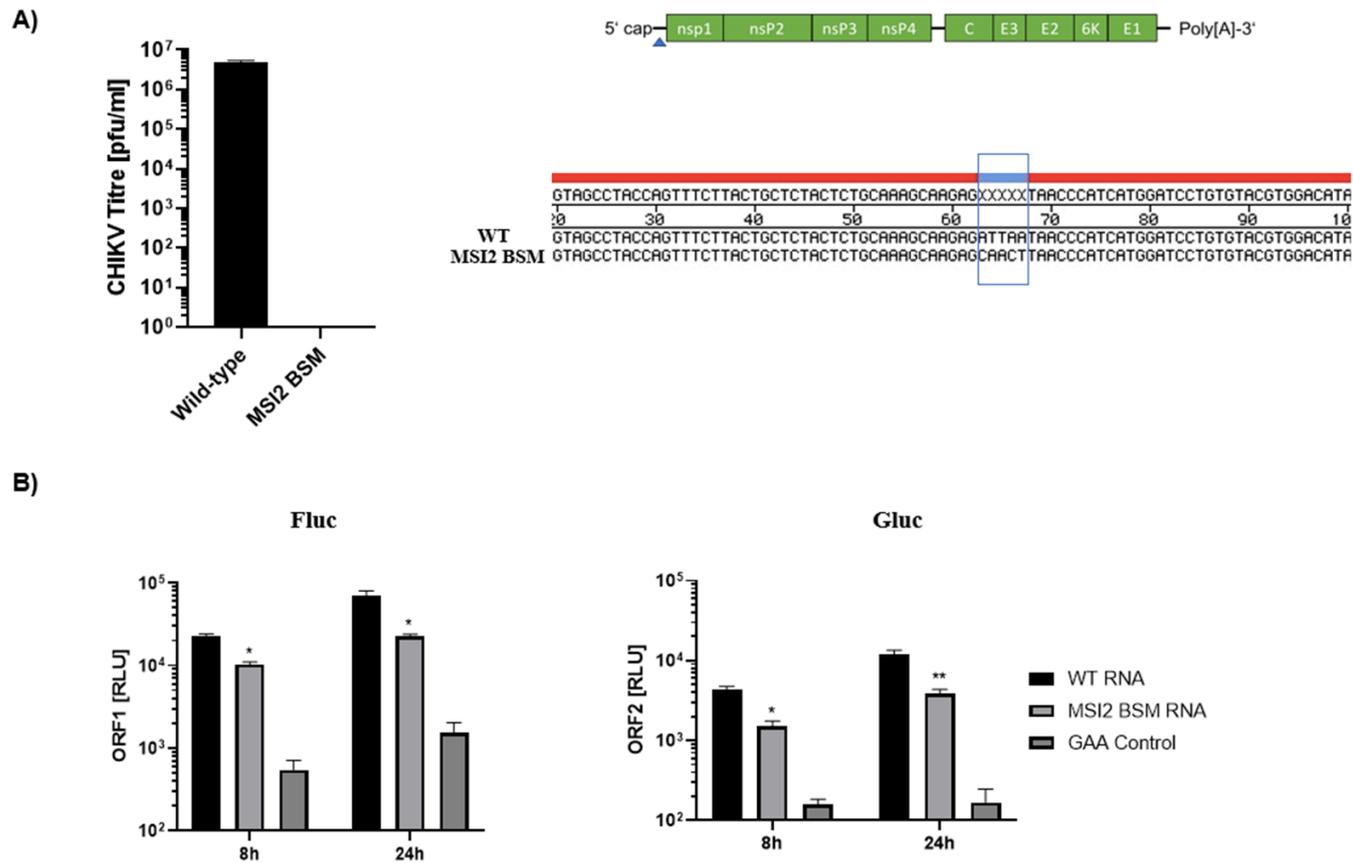


Figure 4.2.6. Mutagenesis of the MSI2 binding sequence on the CHIKV genome significantly impairs replication.

The difference between WT and MSI2 BSM CHIKV is compared in **A)**. Infectious virus production at 24 h post electroporation. The CHIKV genome with the approximate site of mutations indicated by a triangle. The Sanger sequencing data showing the $_{63}AUUAA_{67} > _{63}CAACU_{67}$ mutation in the MSI2 BSM RNA compared to WT (highlighted with blue square); **B)**. trans-complementation replicon system in RD cells at 8 h and 24 h post transfection. The average of three independent experiments ($n=3$) is shown. Error bars represent standard deviation from the mean. *, $P<0.05$; **, $P<0.01$; ***, $P<0.001$ (One-way ANOVA).

4.3 Discussion

The genomic RNA replication of alphaviruses begins by the synthesis of the complementary negative-sense RNA, followed by the synthesis of genomic and sub-genomic RNAs. In this chapter, the role of MSI2 CHIKV genome replication was analysed. MSI2 interacts with RNA via the RRM1, which is also conserved in its paralog MSI1 (257). Therefore, the effect of downregulation of MSI1 and MSI2 using siRNA was first examined to investigate their requirement in CHIKV replication respectively and collectively. To test this, only Huh7 cells were used, as the expression of MSI1 in RD cells was barely detectable by western blot analysis (Figure 4.2.1A). For the siRNAs used in this study, the 'MSI-free' cellular environment due to knockdown were able to last up to 48 h (24 h after the virus infection/replicon transfection). Results from infectious virus production and genomic copies of both sense strand RNA demonstrated that while both MSI1 and MSI2 siRNA knockdown significantly inhibited CHIKV replication, MSI1+2 siRNA knockdown further increased the inhibition but was still able to maintain replication (Figure 4.2.2). Compared to MSI1, MSI2 siRNA knockdown also led to a stronger inhibition of CHIKV replication, presumably due to the multiple MSI2 isoforms which had overlapping functions. Notably, Huh7 cells treated with MSI2 siRNA generated significantly reduced virus at 8 h compared to mock, which indicated that the absence of corresponding data for Ro treatment was indeed caused by the higher drug sensitivity of Huh7 cells (discussed in 3.2.1). Together, these data suggested that MSI1 and MSI2 were both required, but not indispensable, for efficient CHIKV replication. However, since the preliminary study using RNA affinity purification and proteomic analysis did not identify the interaction between MSI1 and the 5' region of CHIKV genome, the role of MSI1 in CHIKV replication was not further investigate in detail.

Using two independent approaches to achieve MSI2 downregulation in RD cells, the results reinforced the role of MSI2 in CHIKV replication, as MSI2 knockdown led to significant reductions in infectious virus production and the synthesis of positive- and

negative-sense strand RNA (Figure 4.2.3 & 4.2.4). By isolating genomic translation and replication from other stages of the CHIKV infectious cycle, the significant reduction of both firefly and *Renilla* luciferase from the sub-genomic replicon demonstrated that MSI2 was involved in the replication and/or translation of the CHIKV genome (Figure 4.2.5A). The phenotypes shown by the *trans*-complementation systems further narrowed the role of MSI2 down to CHIKV genome replication (Figure 4.2.5B). As the CHIKV sub-genome replication, which was represented by the gaussia luciferase signal, was also significantly reduced, MSI2 might also be involved in the replication of the sub-genomic RNA. For the data from using replicon systems, parallel experiments should have been conducted to show the corresponding virus protein expression using western blot and the genomic RNA copies using qRT-PCR. Together, these conclusive results were consistent with those in Chapter 3, indicating that MSI2 was specifically required for CHIKV genome replication. To our knowledge, this was a novel finding which investigated the role of MSI2 in CHIKV genome replication.

4.3.1 Possible reasons for incompatibility of lentivirus transduction with Huh7 cells

The liver is one of the target organs of CHIKV infections, which are typically detected with high titre of viral RNA (296). Even though Huh7 cells are considered a model cell line for both *in vivo* representation and *in vitro* applications of the CHIKV infectious cycle, the virus replication level in Huh7 is not as high as in RD cells (274). Our data also showed that the virus titre was at least 2-log fold higher in RD cells, especially at the later time point (Figure 4.2.3B-D). Such reduced permissiveness to CHIKV infection, together with the sensitivity to chemical compounds (Ro and Lipofectamine RNAiMAX) described in 3.3.2, provides a plausible explanation as to the lack of phenotypes with the *trans*-complementation systems in Huh7 cells. On the other hand, shRNA transduction via lentiviral vectors to Huh7 cells could have been more prone to insertional mutagenesis, resulting in unwanted activation of the immune pathways (297). This would substantially interfere with the fusion of shRNA to the genome,

which even if successful, would likely not be permissive to CHIKV replication as activated RIG-I-mediated antiviral pathways in Huh7 cells significantly attenuate virus replication such as hepatitis E virus (298). Therefore, as far as this study concerns, RD cell were a better choice for CHIKV phenotypic studies to generate comparable and parallel results due to their higher versatility to various *in vitro* biological manipulations. For future reference, human foetal astrocyte cells (SVG-A) and mouse myoblast cells (C2C12), both of CHIKV showed high level of replication for CHIKV virus and sub-genomic replicons, can be used for CHIKV infectivity comparison between vertebrate hosts (274).

4.3.2 The proposed role of MSI2 during CHIKV genome replication

Given that the RRM1 of MSI2 preferentially binds to the sequence motif (G/A) U₁₋₃AGU (257, 295), the predicted binding site ₆₃AUUAA₆₇, which is located at the single-stranded sequence nt 63-67 upstream of SL47, was mutated to ₆₃CAACU₆₇ to analyse whether it affected CHIKV replication. It was unexpected that while downregulation of MSI2 expression strongly reduced virus replication, mutagenesis of its potential binding site within the CHIKV 5' UTR resulted in a completely lethal phenotype (Figure 4.2.6A). However, the data from the *trans*-complementation system showed that neither genomic or sub-genome replication were severely reduced (Figure 4.2.6B). As shown in Figure 2.2B, the reporter plasmid encoding genes for firefly and gaussia luciferase of the *trans*-complementation assay system only retained the 5' nt 303, the 3' UTR and the SGP of the CHIKV genome. Given that MSI1 has been found to interact with the 3' end of ZIKA virus genome, there might be MSI2 binding sites within the 3' UTR of CHIKV as well (264). Mutagenesis the MSI2 binding site within the 5' UTR alone may not be sufficient to disrupt the overall association between MSI2 and CHIKV RNA. This would explain the reduced inhibition of genome replication obtained from MSI2 BSM compared to MSI2 siRNA/shRNA knockdown. Nonetheless, it was difficult to interpret the lethal phenotype seen in Figure 4.2.6A, as evidence was needed to confirm the direct interaction between MSI2 and ₆₃AUUAA₆₇ within the CHIKV 5' end.

For alphaviruses, it is vital to balance the mutually antagonistic translation-replication process as they use the same RNA template. A key step is for the viral RdRp to specifically interact with the 3' end of its cognate RNA for the initiation of negative-sense strand synthesis (299). However, increasing evidence shows that *cis*-acting elements located within the 5' UTR of many positive-sense RNA viruses are also required to promote or suppress virus replication, presumably via genome circularisation (85, 248, 251, 300, 301). Notably, inhibition of genomic translation has been found in flaviviruses when the genome is in circularised form, which might involve a dynamical structural-dependent mechanism within the 5' end (302, 303). In the case of alphaviruses, host encoded hnRNP A1 is required for the synthesis of Sindbis virus negative-sense RNA by directly interacting with the promoter elements within the 5' end (304). Interestingly, this cellular protein has also been shown to inversely regulate the replication and translation of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in a methylation-dependent mechanism (305). A recent study stated that methylation of Sindbis virus RNA leads to a reduction in virus replication and infectivity in insect cells (306). Based on the aforementioned information and our results, it was hypothesised that if MSI2 directly interacts with the 5' UTR of CHIKV genomic RNA via the proposed binding sequence (₆₃AUUAA₆₇). MSI2 presumably functions synergistically with other cellular proteins to dynamically rearrange the genomic structure of CHIKV to switch from translation to initiation of negative-sense strand synthesis. Furthermore, apart from mammalian cells, orthologs of MSI have been found in *Drosophila* (307, 308). This might suggest that a similar mechanism also exists in mosquito cells, involving host-specific *trans*-activating factors to maintain fitness in both vertebrate and invertebrate hosts. Therefore, to support our hypothesis, the interaction between MSI2 and the 5' region of CHIKV genomic RNA, and the effect of MSI2 BSM on this interaction, were biochemically analysed in Chapter 5.

CHAPTER 5: BIOCHEMICAL ANALYSIS OF THE NOVEL INTERACTION BETWEEN MSI2 AND CHIKV 5' RNA

5.1 Introduction

5.1.1 Methodological rationale for investigating CHIKV RNA-MSI2 interaction

Following the phenotypic and reverse genetic analysis in Chapter, MSI2 was hypothesised to function as a molecular switch for CHIKV translation and genome replication. Therefore, it was important to identify and analyse if there was specific an interaction between CHIKV 5' RNA and MSI2 at the predicted binding site (₆₃AUUAA₆₇) using biochemical approaches. The electromobility shift assay (EMSA), also known as the gel shift assay, is a straightforward and sensitive technique to detect interactions between nucleic acids and proteins (309). Dye or isotopically labelled nucleic acids are incubated in appropriate reaction conditions with the proteins, which are then separated by polyacrylamide or agarose gel electrophoreses. The principle of EMSA is that under native or non-denaturing gel electrophoresis, protein-nucleic acid complexes migrate slower than unbound nucleic acids due to higher molecular weights of the RNA/protein complex (310). Consequently, the RNA/protein bands migrate less far than those representing unbound RNA, while proteins and un-labelled nucleic acids will not be visualised on the gel (Figure 5.1.1). The gel was then fixed to perpetuate the position and signal of the dye or isotope before X-ray film exposure.

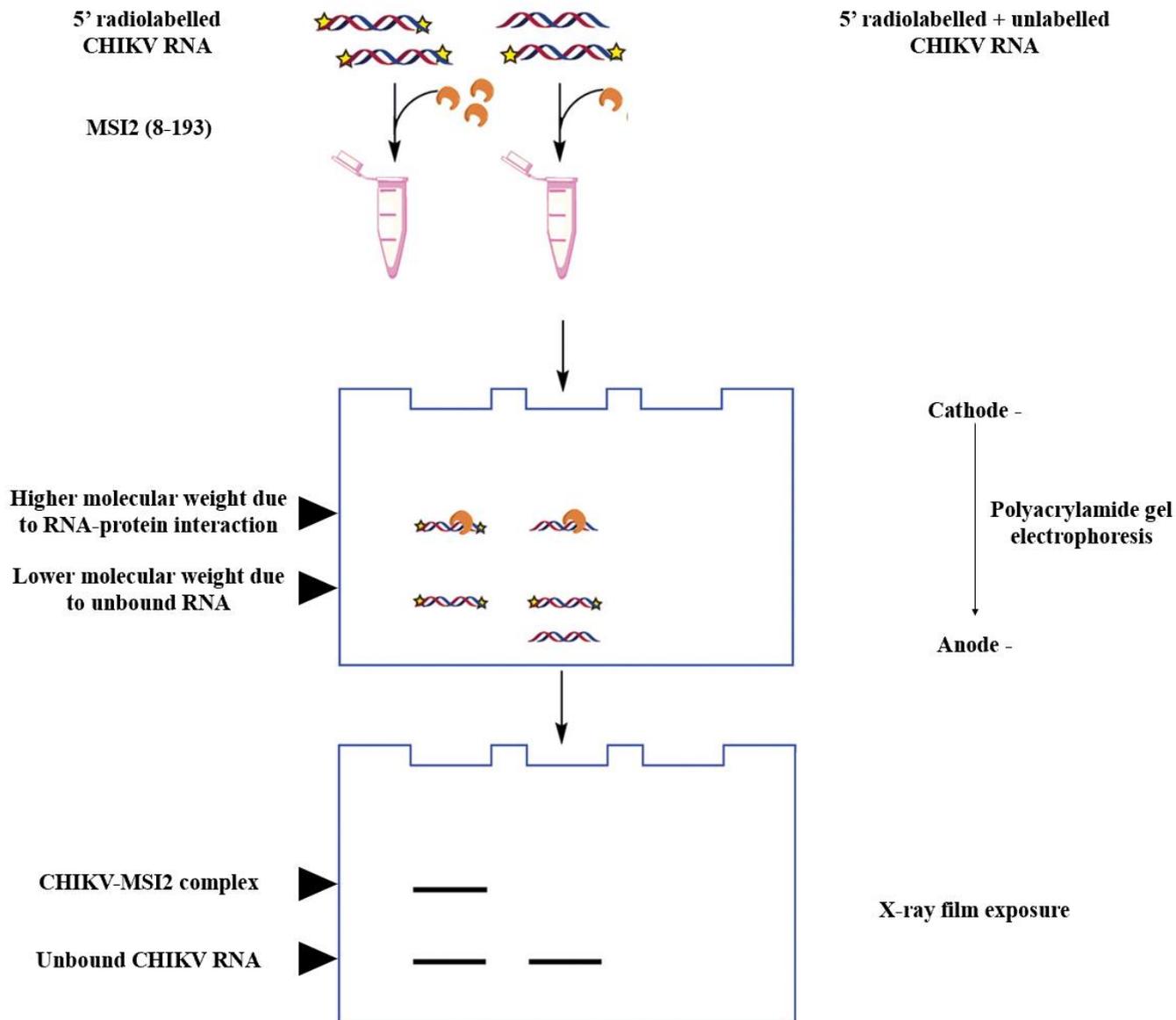


Figure 5.1.1. The schematic illustration of EMSA (310).

The 5' end of CHIKV 5' RNA bait was radiolabelled (star) before mixing with proteins. Unlabelled RNA was added as specific competitors. The samples are separated on native polyacrylamide. The bands at the bottom represent free or unbound labelled RNA, and the upper shifted band represents the association with MSI2 (8-193) which increases the molecular weight of labelled nucleic acids. Un-labelled RNA cannot be exposed and thus invisible on the X-ray film despite their presence.

5.1.2 Hypothesis underpinning the biochemical analysis

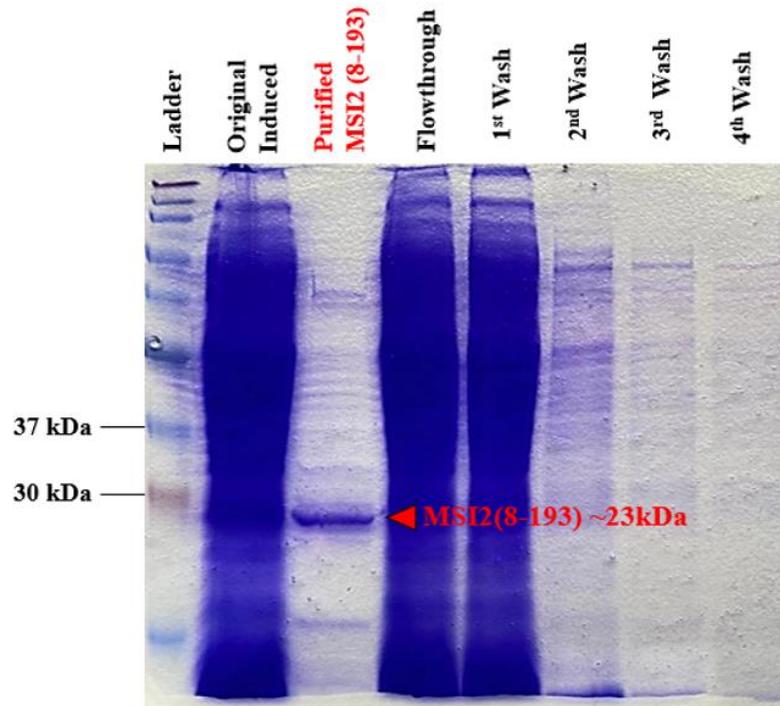
Stable secondary structures within the alphavirus 5' UTR are critical for modulating regulation, as mutagenesis studies indicate that this is both structural- and sequence-dependent (86, 89, 90). The complementary 3' UTR of the negative-sense strand might be directly associated with the viral RdRp for initiating the synthesis of the positive-sense strand (311). Interestingly, these secondary structures vary between species despite sharing the conserved sequence motifs, suggesting that for replication their interaction with the cognate 3' UTR is distinct and intraspecies incompatible (312). This indicates the involvement of virus-specific host cell factors during alphavirus replication. Results outlined in Chapter 3 and 4 demonstrated that MSI2 is required for efficient CHIKV genome replication by participating in the initiation of negative-sense strand synthesis, as silencing of MSI2 expression significantly reduced the genomic copies of the negative-sense strand. Therefore, it was postulated that MSI2 regulates this process by directly interacting with the CHIKV 5' UTR single-stranded sequence nt 63-67, which is flanked by stem-loops essential for virus replication (Figure 4.1.2) (86, 312). In order to test our hypothesis, a range of EMSA competition binding experiments were conducted. The kinetics of interaction were also compared between MSI2 and wild-type CHIKV 5' RNA and mutant RNA in which either the putative MSI2 binding site was mutated or the flanking RNA structures were mutated. The EMSA binding affinity was quantified by comparing the optical density of shifted bands. This biochemical analysis would have a significant impact on dissecting the mechanism of MSI2 during CHIKV replication.

5.2 Results

5.2.1 Purification of MSI2 protein containing the two RNA binding domains

To establish *in vitro* molecular interactions, the CHIKV 5' RNA and MSI2 protein were individually expressed and purified. The RNA can be easily synthesised by PCR amplification of the CHIKV 5' end and *in vitro* transcription, while the MSI2 protein sequence was cloned into expression vectors. Construct pET-22HT-MSI2 (8-193) was a gift from Sean Ryder (Addgene plasmid # 60356; [http://n2t.net/addgene: 60356](http://n2t.net/addgene:60356); RRID: Addgene_60356). It expressed the two RNA binding domains of human MSI2 rather than the full-length protein and was therefore designated MSI2 (8-193). This truncated version of MSI2 has been shown to accurately represent the binding mechanism of full-length MSI2 with its cellular targets, including RNA and fatty acids (266). This construct was purified as described in 2.2.19. The efficiency of purification was assessed by SDS-PAGE (Figure 5.2.1A). MSI2 (8-193) was expressed to a high concentration (~1.5 mg/ml). The protein band for the ~23 kDa MSI2 (8-193) was visually undetectable in flowthrough or wash samples, suggesting efficient His-column purification. When analysed by western blotting, the purified protein interacted with the MSI2 antibody and its expected apparent molecular weight (~23 kDa) was justified compared to the protein ladder (Figure 5.2.1B). This confirmed that the protein was indeed MSI2 (8-193).

A)



B)

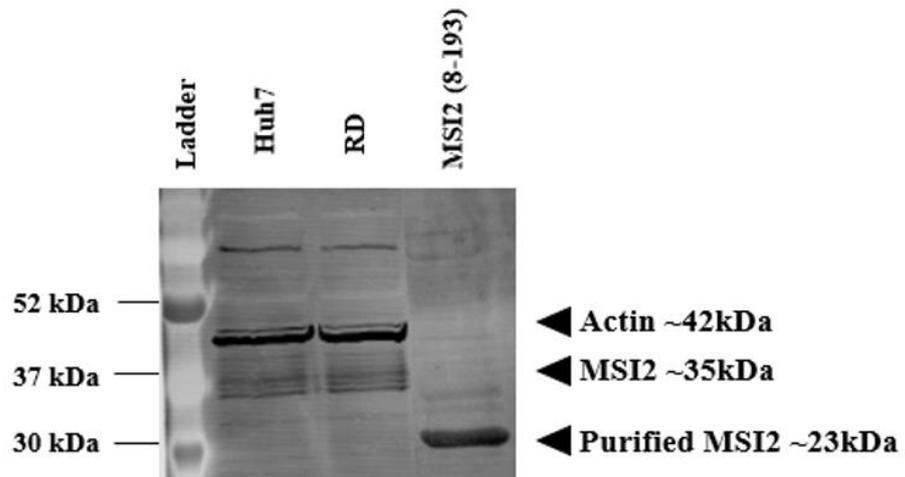


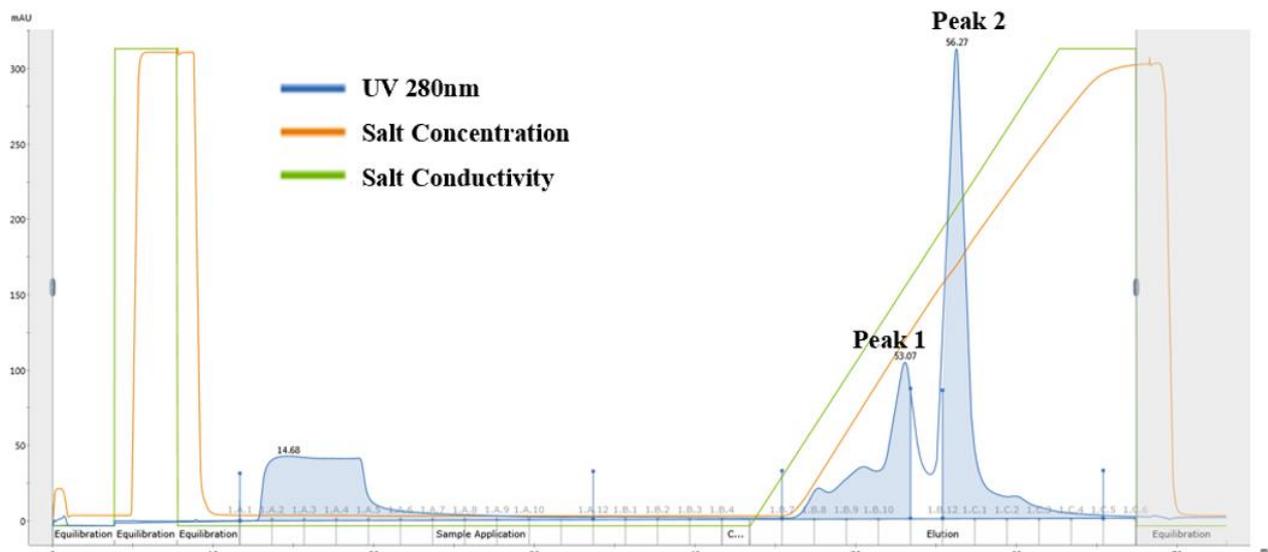
Figure 5.2.1. Expression and purification of MSI2 (8-193).

A). SDS-PAGE of purified protein compared to the original induced sample, flowthrough and washes during purification. The expected MSI2 band was highlighted in red, with an apparent molecular weight of ~23 kDa. **B).** Western blot confirming the identity of purified protein. Untreated RD and Huh7 cell lysates were used as the

positive control. The membrane was probed with β -actin and MSI2 antibodies. The presence of cellular MSI2 and purified MSI2 (8-193) were labelled according to their apparent molecular weight.

To eliminate the non-specific co-purifying proteins, the purified MSI2 (8-193) was subjected to ion exchange chromatography. *In silico* analysis of the MSI2 (8-193) amino acid sequence indicated a pI value of ~ 7.34 . The acidic compound MES was used as a buffer to conduct cation exchange. The purified protein was de-salted using a PD-10 column to a low-salt MES buffer before the exchange process, so that it could firmly bind to the exchange column and fractionally elute as the salt concentration was increased. The fraction with the most concentrated MSI2 (8-193) would theoretically be eluted at a high salt concentration due to the strong affinity binding of protein to the cation column induced by molecular charge difference, thus creating an absorbance peak in the corresponding chromatogram. However, there were two absorbance peaks recorded during the salt gradient, which was an indication that other proteins apart from MSI2 (8-193) in the mixture were equally, but to a less extent, sensitive to salt gradient (313) (Figure 5.2.2A). A total of eleven elutions were obtained, and these fractionated protein samples were quantified and analysed by SDS-PAGE to confirm purity. Four out of the eleven purified MSI2 (8-193) demonstrated high concentration (> 0.5 mg/ml), with elution 6 being the most concentrated and pure with minimum non-specific protein bands (Figure 5.2.2B). Therefore, elution 6 was quantified and used for downstream experiments.

A)



B)

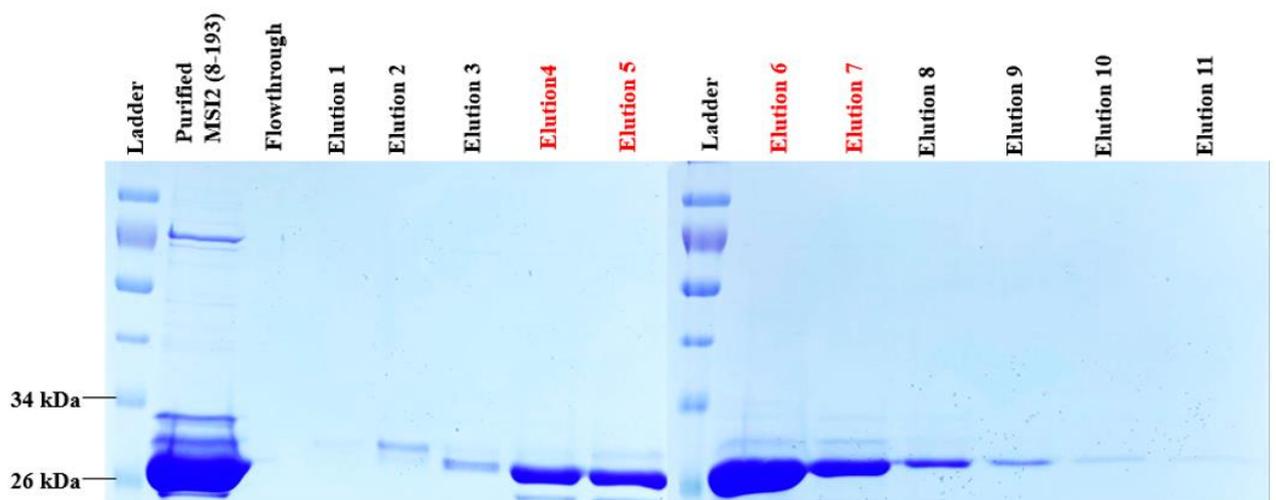


Figure 5.2.2. Further purification of MSI2 (8-193) using ion exchange.

A). Chromatogram profile of ion exchange of MSI2 (8-193). Absorbance readings were recorded at 280nm UV indicated by blue lines; salt concentration and conductivity were indicated by orange and green lines, respectively. **B).** SDS-PAGE of the eleven elutions from the ion exchange. Purified MSI2 (8-193) before the exchange process was used as positive control and the flowthrough sample was used as the negative control. The four elutions with relatively high purity and protein concentration were highlighted in red.

5.2.2 MSI2 binds specifically to the 5' region of the CHIKV genome

As MSI2 was identified by using the 5' end CHIKV genome as bait, their specific interaction was investigated by EMSA. The 5' end radiolabelled RNA was denatured and then allowed to refold at 37°C to the structured RNA molecule (83). The RNA and purified MSI2 protein were then incubated together with yeast tRNA as the non-specific competitor, followed by analysis on native polyacrylamide gel electrophoresis. EMSA results were visualised on X-ray films exposed from the native polyacrylamide gels. In Figure 5.2.3, as the amount of MSI2 (8-193) added to the interaction was increased, the high-intensity band representing MSI2-RNA complex gradually shifted up, suggesting that radiolabelled RNA became heavier as it associated with more MSI2 (8-193). Interestingly, the shifted bands were smeary, which might indicate that CHIKV 5' RNA directly binds to multiple MSI2 (8-193) molecules so that MSI2 (8-193) interacted with CHIKV in both monomers and oligomers.

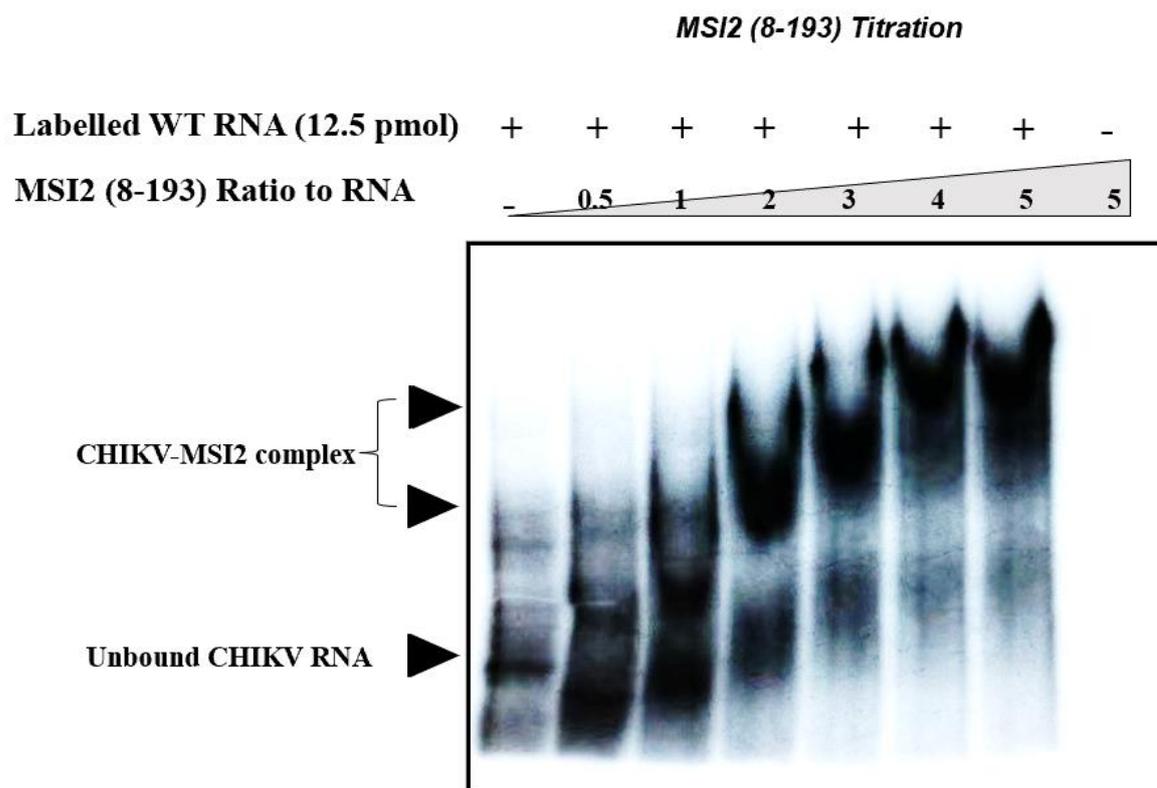


Figure 5.2.3. MSI2 (8-193) titration of labelled WT CHIKV 5' RNA.

A representative result out of three biological repeats was shown. 12.5 pmol RNA was radiolabelled and incubated with an increasing amount of MSI2 (8-193). Labelled RNA only sample was used as positive control and protein only was used as the negative control. Yeast tRNA was applied to all samples as non-specific competitors.

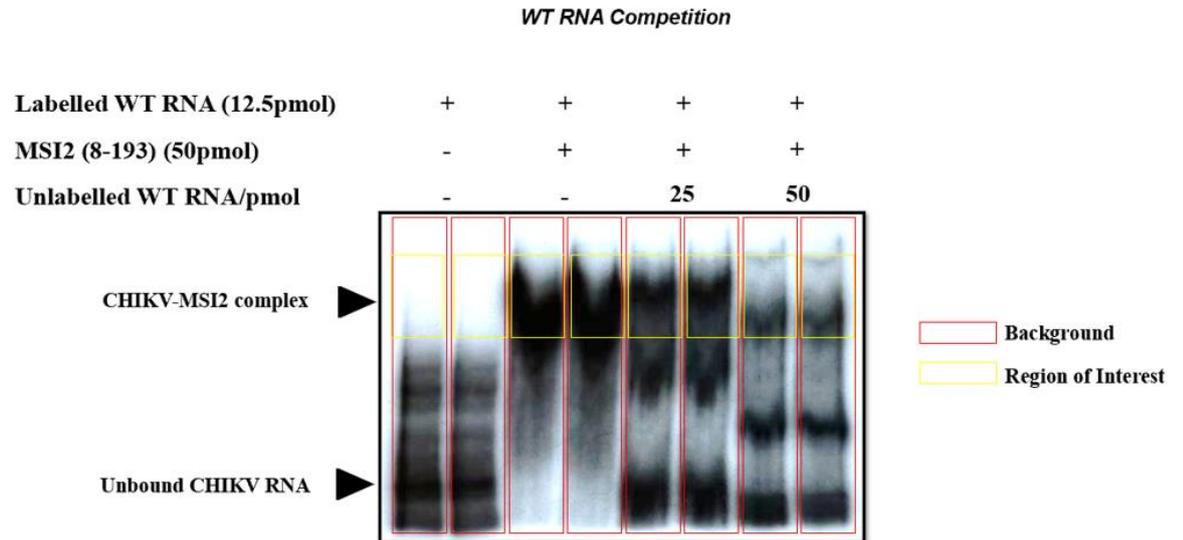
To further investigate the specificity of the interaction, an un-labelled RNA competition assay was performed. Un-labelled RNA was considered to interact with the protein specifically and equally as labelled RNA. The addition of the un-labelled competitor led to dramatic changes to the pattern of shifted bands (Figure 5.2.4A). Instead of forming one predominant band around the top part of the gel, two more bands of lower molecular weight were visualised, demonstrating the competitive binding of un-labelled RNA to the protein. The bottom band represented out-competed unbound

labelled RNA. When four times more un-labelled RNA was added to the reaction mix, the two upper bands prominently shifted down, compared to un-competed and two times more un-labelled RNA, as a consequence of increased opportunities for un-labelled RNA-protein interactions. These results, together with the protein titration shown in Figure 5.2.3, provide convincing evidence that MSI2 specifically interacts with the CHIKV 5' region.

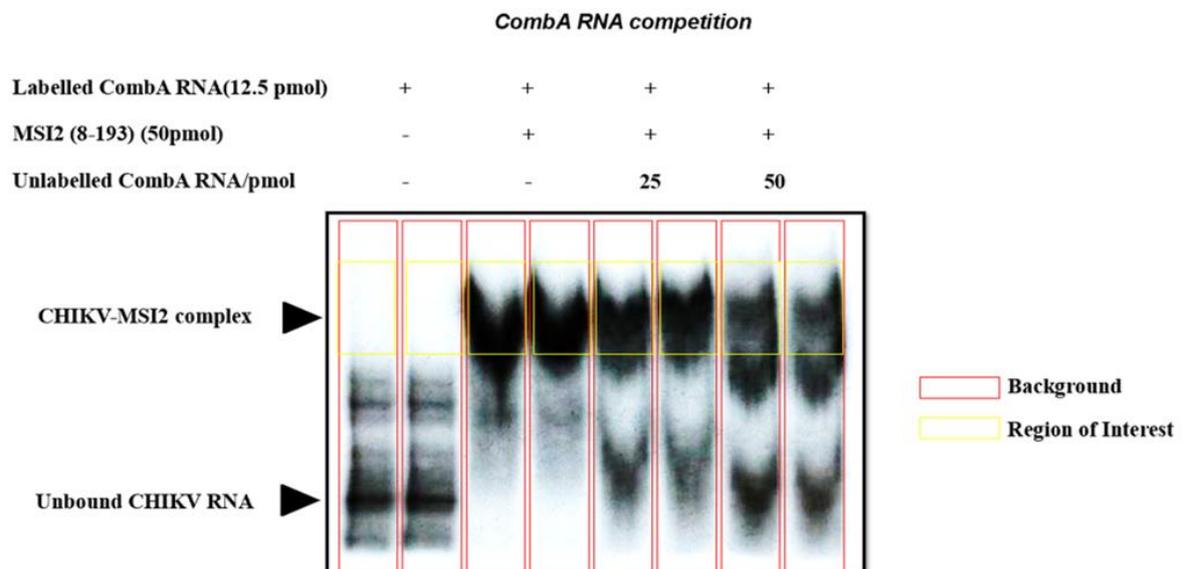
The 5' end of CHIKV encodes several stem-loops that are essential for viral replication in vertebrates and invertebrates and flank the proposed MSI2 binding site (Figure 4.1.2) (83). I therefore sought to investigate whether these secondary structures affected the RNA-MSI2 interaction. Previous work in our group by Catherine Kendall generated a mutated version of CHIKV ICRES in which all the RNA stem-loops were disrupted. This mutant, designated combination mutant A (CombA), was unable to produce progeny virus in either human or mosquito cells. The 5' nt 1-330 of CombA was amplified by PCR and used as a template for *in vitro* transcription and subsequent RNA was radiolabelled to perform EMSA with MSI2 (8-193). Surprisingly, although CombA mutations did not abrogate interactions with MSI2 (8-193), competition assay revealed that, unlike the wild-type interaction, the MSI2-RNA interaction was not out-competed by the addition of un-labelled competitor (Figure 5.2.4B). To statistically quantify and analyse the interaction, band shift densitometry was performed to compare the results from WT and CombA RNA competition assays. The uppermost bands were chosen for the analysis. Competition with WT RNA showed a significant reduction in band intensity. In contrast, CombA RNA competition had no significant effects (Figure 5.2.4C). Parallel competition assay also showed that when un-labelled WT and CombA were added to labelled WT RNA, their competition ability was comparable (Figure 5.2.5). This demonstrated that the secondary structures within the 5' end had no effect on MSI2-RNA binding, whereas their primary sequence may provide additional binding platforms for MSI2. The relaxation of stem-loops might facilitate the protein to interact more stringent with the RNA, making the effect of the competition assay less conspicuous. Together, these results not only corroborated the finding from

preliminary IP and proteomic analysis that MSI2 was significantly enriched, but also confirmed the specific and direct interaction between CHIKV 5' end and MSI2.

A)



B)



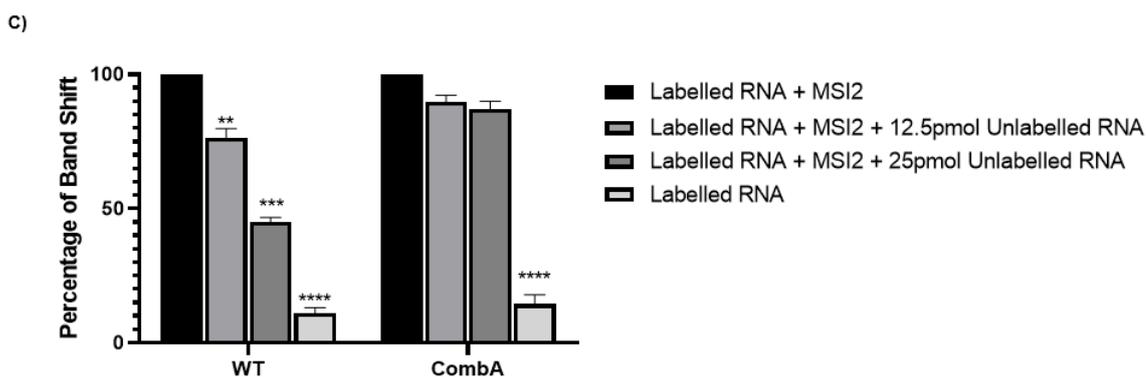


Figure 5.2.4. WT and CombA RNA competition assay.

A). A representative result out of three biological repeats was shown. Un-labelled WT RNA was increasingly added to the reaction mix containing MSI2 (8-193) and radiolabelled WT RNA. Each setting was performed in duplicate. For densitometry analysis, the background for each lane was highlighted in red and the region of interest for band shifting was highlighted in yellow. Labelled RNA only was used as the positive control. Yeast tRNA was applied to all samples as non-specific competitors. **B).** A representative result out of three biological repeat was shown. Un-labelled CombA RNA was increasingly added to the reaction mix containing MSI2 (8-193) and radiolabelled CombA RNA. Each setting was performed in duplicate. For densitometry analysis, the background for each lane was highlighted in red and the region of interest for band shifting was highlighted in yellow. Labelled RNA only was used as the positive control. Yeast tRNA was applied to all samples as non-specific competitors. **C).** Densitometry analysis of band shifting resulted from un-labelled RNA competition. The average of three independent experiments ($n=3$) is shown. The average of three independent experiments ($n=3$) is shown. Error bars represent standard deviation. *, $P<0.05$; **, $P<0.01$; ***, $P<0.001$ (One-way ANOVA).

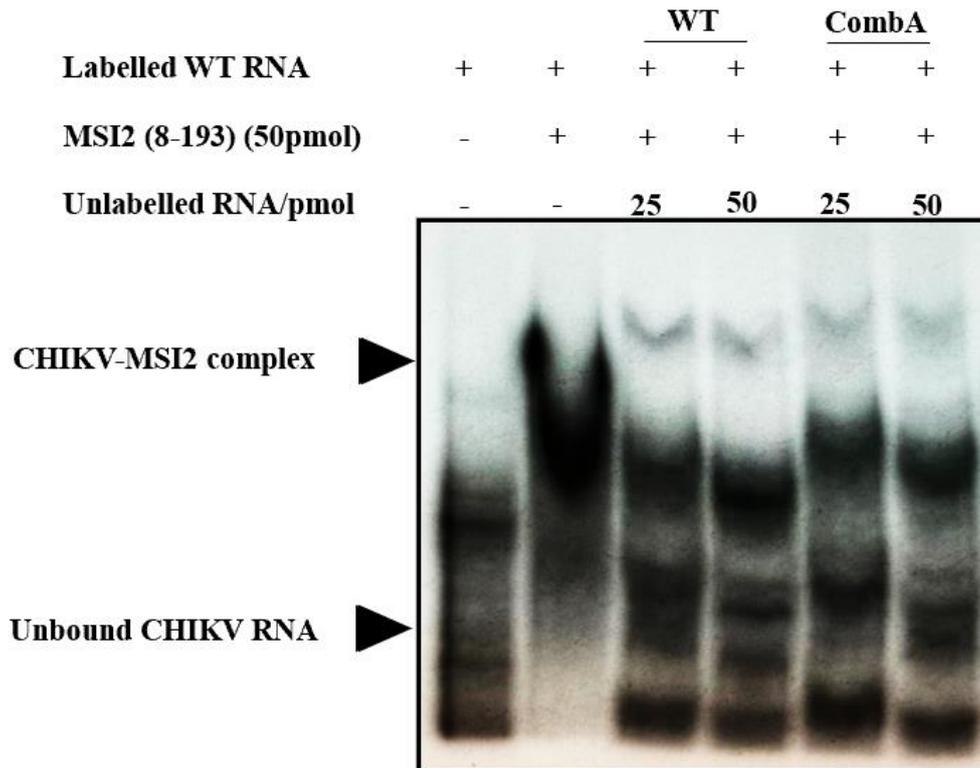


Figure 5.2.5. The ability of un-labelled WT and CombA RNA to compete with labelled WT RNA for MSI2 interaction.

A representative result out of three biological repeats was shown. An equal amount of un-labelled WT and CombA RNA were added for comparison. Labelled RNA only was used as the positive control. Yeast tRNA was applied to all samples as a non-specific competitor.

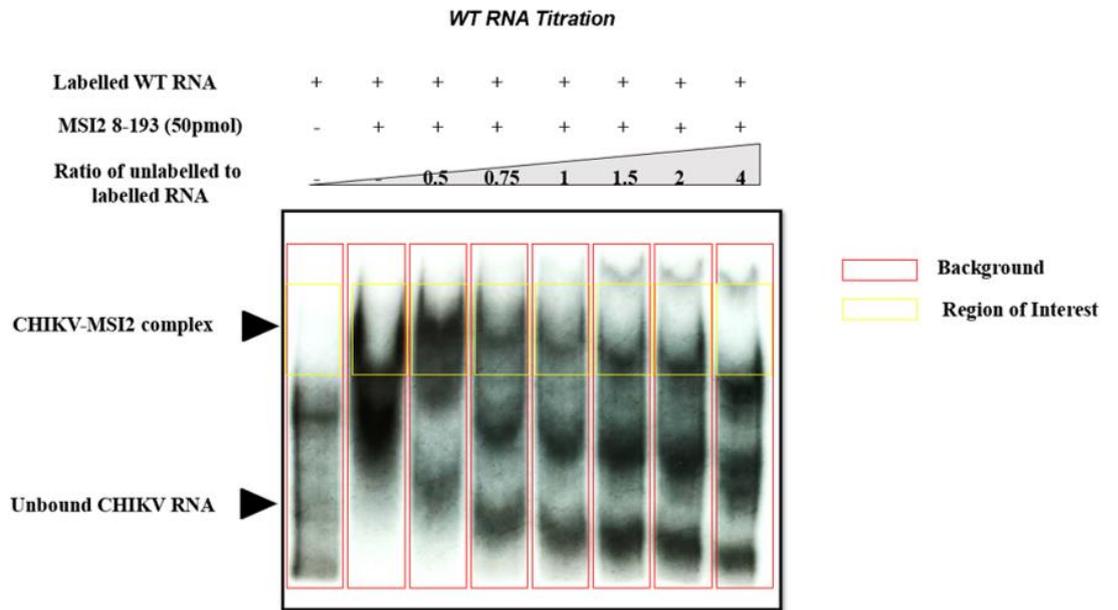
5.2.3 MSI2 specifically binds to the single-stranded region upstream of the AUG start codon

Since inconsistency in the results of full-virus and trans-complementation assay for the MSI2 BSM shown in Figure 4.2.5 of Chapter 4 suggested that MSI2 may function as a molecular switch for CHIKV, an alternative titrated version of the EMSA competition assay was performed. This assay used the un-labelled WT and MSI2 BSM RNA respectively to titrate against radiolabelled WT RNA. The MSI2 BSM RNA had the

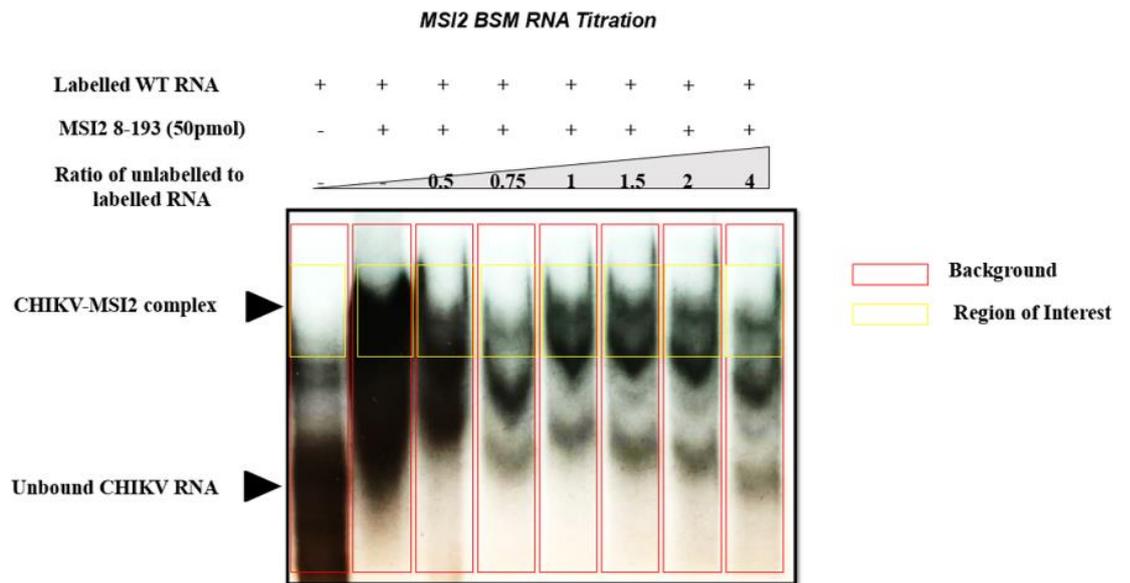
potential MSI2 binding site ${}_{63}\text{AUUAA}_{67}$ mutated and showed lethal phenotype for full-virus and reduced replication for *trans*-complementation assay (Figure 4.1.2). As expected, the bands for the MSI2-RNA complex gradually shifted down as more un-labelled wild-type RNA competitor was added (Figure 5.2.6A), consistent with the result shown in 5.2.2 above. In contrast, titration with increasing concentrations of MSI2 BSM RNA did not result in the same level of band shift, indicating that it competed less well for MSI2 binding than WT CHIKV RNA (Figure 5.2.6B).

In order to get a more comprehensive understanding of how the BSM mutation affected MSI2-RNA binding, the difference in band shifting between un-labelled wild-type and MSI2 BSM RNA was analysed by quantifying the relative band intensity. The result showed that mutagenesis of the putative binding site reduced the RNA's ability to compete for MSI2 binding by ~ 3 fold (Figure 5.2.6C). The addition of un-labelled MSI2 BSM RNA also failed to induce the gradual decline of band shift as seen in WT, suggesting a severe impairment in MSI2 interaction. Together, these results demonstrated that MSI2 directly interacts with the 5' end of the CHIKV genome, and a single-strand sequence ${}_{63}\text{AUUAA}_{67}$ upstream of SL47 served as potentially one of the binding sites.

A)



B)



C)

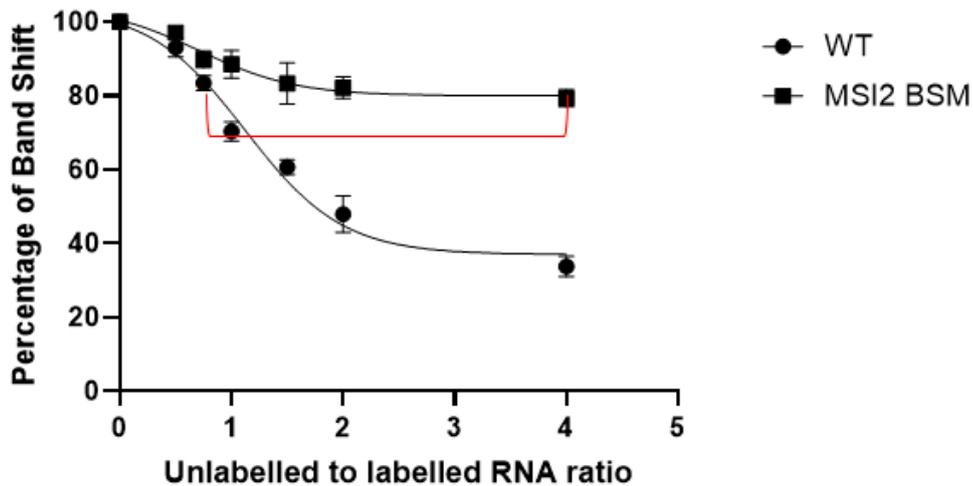


Figure 5.2.6. WT and MSI2 BSM RNA competition titration assay.

A). A representative result out of three biological repeats was shown. Un-labelled WT RNA was increasingly added to the reaction mix containing MSI2 (8-193) and radiolabelled WT RNA. For densitometry analysis, the background for each lane was highlighted in red and the region of interest for band shifting was highlighted in yellow. Labelled RNA only was used as the positive control. Yeast tRNA was applied to all samples as non-specific competitors. **B).** A representative result out of three biological repeats was shown. Un-labelled MSI2 BSM RNA was increasingly added to the reaction mix containing MSI2 (8-193) and radiolabelled WT RNA. For densitometry analysis, the background for each lane was highlighted in red and the region of interest for band shifting was highlighted in yellow. Labelled RNA only was used as the positive control. Yeast tRNA was applied to all samples as non-specific competitors. **C).** The decline of band intensity resulted from un-labelled RNA competition. WT RNA data was indicated as circle and MSI2 BSM as square. The red line indicates the difference in the amount of un-labelled RNA required to cause band shift between WT and MSI2 BSM. The average of three independent experiments ($n=3$) is shown.

5.3 Discussion

Despite the fact that the CHIKV 5' UTR has been demonstrated to play crucial roles in regulating translation, replication and innate immunity evasion during CHIKV replication, its specific interactions with cellular and viral factors remain obscure (85, 312, 314-317). In this chapter, the results from the EMSA biochemical analysis confirmed MSI2 as a host cell protein which specifically binds to the single-stranded sequence domain (nt 63-67) upstream of the AUG start codon within the CHIKV 5' end. As MSI2 was shown to be essentially required for efficient CHIKV genome replication in Chapter 3 and 4, our finding provides valuable insights into dissecting the mechanism of the virus lifecycle of how it hijacks host cell proteins to promote replication.

5.3.1 The proposed mechanism of MSI2 during CHIKV genome replication

These results are the first to show that CHIKV 5' RNA binds directly to MSI2. As the concentration of MSI2 (8-193) was increased, the RNA-MSI2 band further shifted up, indicating that MSI2 presumably interacted with the CHIKV 5' region in a multi-protein complex form (Figure 5.2.3). Although the protein construct used in this study only expressed the two RNA binding domains of MSI2 rather than full length, it was considered to possess the complete function of MSI2 as allosteric inhibitors selectively impeded the RNA-MSI2 association by inducing conformational changes of the RNA binding domains (266). The single-stranded region ${}_{63}\text{AUUAA}_{67}$ within the CHIKV 5' UTR was specifically recognised by the protein, and mutations of this sequence interfered with the interaction (Figure 5.2.6). Given that the binding sequence is the single-stranded sequence ${}_{63}\text{AUUAA}_{67}$ flanked by RNA stem-loops critical to the initiation of negative-sense strand synthesis, it may be supportive evidence to our hypothesised that MSI2 was also essentially involved in this process.

Interestingly, disruption of all RNA stem-loops had minimal effect on the RNA-protein interaction (Figure 5.2.5). This CombA mutant virus was phenotypically lethal through an exclusive genome replication-related mechanism (83). As mentioned in Chapter 4, MSI2 was required for the efficient initiation of negative-sense strand synthesis as MSI2 knockdown significantly reduced the genomic RNA copies of the negative-sense strand. Based on these findings, it was speculated that the stem-loops may play critical roles after the initiation of replication regulated by MSI2. However, the significant inability of un-labelled CombA RNA to compete with WT labelled RNA suggested that CombA might form stronger interaction with MSI2 than WT due to the relaxation of secondary structures, or the mutations potentially generated more protein binding sites within the 5' end (Figure 5.2.4). In other words, the natural sequence and structure of the stem-loops presumably confine high-affinity RNA-protein interaction so that MSI2 can be promptly dissociated from the viral genome if it functions as a molecular switch as hypothesised in 4.3.1 of Chapter 4.

Results in this chapter demonstrate that the RNA-MSI2 interaction was direct and could form independent of any host or viral proteins. CHIKV exploits several host cell proteins from the cellular transcription and translation machinery to benefit its replication. For example, nsP2 interacts with hnRNP K, resulting in its relocation from the nucleus to cytoplasm to facilitate virus replication (318); nsP3 recruits Ras GTPase-activating protein-binding protein 2 (G3BP2) to virus-induced spherules to inhibit the formation of stress granules (174). These host cell proteins are hijacked and sequestered to the replication complex by CHIKV nsPs to divert normal cellular processes to virus replication and inhibit innate immune response. However, direct interactions between cellular protein and CHIKV RNA seen in this study are less well documented. Therefore, it was hypothesised that MSI2 might function as a platform for other host and viral proteins to associate with the CHIKV genome. For instance, MSI2 may bind to CHIKV nsPs as well as cellular proteins responsible for genome circularisation so that the nsPs could initiate negative-sense strand synthesis. Furthermore, possessing binding sites in the positive-sense genome may also suggest

the involvement of MSI2 in progeny RNA packaging. For influenza viruses, packaging signals can be found within the 5' and 3' UTR for different segments (319). A conserved and shared MSI2 binding site for both replication initiation and packaging in the CHIKV genome would reflect how viruses evolve to achieve efficient genome arrangement (320).

5.3.2 Evaluation of EMSA

EMSA is an extremely sensitive technique for *in vitro* analysis of nucleic acid-protein interactions. To create conditions similar to the cellular environment, CHIKV 5' RNA was folded in a physiologically relevant buffer to its predicted structures, while MSI2 (8-193) was processed to be as pure as possible (Figure 5.2.2). Despite the overall promising results, there were a few factors which affected the interpretation of the results. The main issue was the instability of RNA, which needs RNase-free environments and low temperatures to minimise degradation. After *in vitro* transcription, the CHIKV 5' end RNA went through 5' phosphate removal, purification, 5' end radiolabelling and another purification before incubating with MSI2 (8-193). This time-consuming process inevitably led to the loss of RNA quantity during the experimental process. Each experiment was therefore performed with three biological repeats to ensure accuracy. The data shown in this chapter also indicated that MSI2 might bind to the CHIKV 5' RNA as monomers as well as oligomers. This would potentially affect the accuracy of the densitometry analysis as the uppermost shifted bands did not account for RNA bound to MSI2 monomers. Moreover, decreasing the electrostatic property of RNA by adjusting the salt concentration in the buffers would be useful to further confirm the specificity of MSI2-RNA interaction. On the other hand, the exposure intensity of the radioactive gel was very difficult to control due to the decaying radioactivity of ATP, [γ -32P] (321). This resulted in differences in exposure intensity between experiments, which was overcome by the quantitative analysis of band intensity relative to the internal control of each individual result (Figure 5.2.4 & 5.2.5). Alternative methods using chemiluminescent tags for RNA labelling could be performed for further studies to avoid the issue with exposure.

Chapter 6: DISCUSSION AND CONCLUSION

6.1 Summary on the role of MSI2 in CHIKV genome replication

Preliminary study in our group identified the host cell protein MSI2 as the binding partner of the CHIKV 5' 1-303nt bait RNA using RNA affinity purification and TMT-MS. Here, using both pharmacological and genetic approaches, for the first time this study demonstrated that the host cell protein MSI2 had a significant pro-viral effect during CHIKV genome replication. Biochemical analysis further identified a direct interaction between MSI2 and ₆₃AUUAA₆₇ at a single stranded region located between SL47 and the AUG start codon within the 5' UTR of CHIKV genome. Disruption of this interaction or inhibiting the expression/RNA-binding activity of MSI2 significantly reduced CHIKV genome replication. Taken together, the results from this study demonstrate that MSI2 interacts with the 5' region of the CHIKV genome. They also support the hypothesis that this MSI2-RNA interaction might be involved in switching the CHIKV replication cycle from ORF-1 translation to the initiation of negative-sense strand synthesis.

As translation and replication of positive-sense RNA viruses share the same genomic template, dissecting the functions and structures of their promoter-containing UTRs is fundamental for understanding how these two mutually incompatible processes are regulated through a molecular switch (250, 322, 323). Translation initiates at the AUG start codon at the 5' end of the positive-sense RNA virus genome, while replication initiates from the 3' end. Notably, mutagenesis which disrupted the secondary structures within the 5' UTR of alphaviruses had no effect on translation, whereas replication was dramatically precluded (83, 86, 324). This implies that during translation and replication, viral genome must circularise or communicate between the 5' and 3' ends (325, 326), and *trans*-activating host/viral cell factors must interact synergistically with the viral genome to ensure the efficient switching between the two mutually exclusive processes. Increasing evidence also demonstrates that blockage of

cis-acting interactions within or between the 5' and 3' UTR have a stronger deleterious effect on alphavirus and flavivirus replication in invertebrate cells, which further highlight the importance of viral UTRs in adapting to host alteration and maintaining optimal fitness (66, 89, 327).

Although our results demonstrated that the ${}_{63}\text{AUUAA}_{67} > {}_{63}\text{CAACU}_{67}$ mutation hindered the binding capacity of MSI2 to CHIKV 5' RNA by approximately 3-fold, it did not completely abrogate RNA-protein interaction (Figure 5.2.6B & C). This can be supported by the *trans*-complementation assay that the expression level of both ORFs for the MSI2 BSM was only reduced by ~ 0.3 -log fold (Figure 4.2.6B), whereas MSI2 shRNA knockdown showed a ~ 1.1 -log fold reduction (Figure 4.2.5B). The shifted bands were smeary, suggesting CHIKV 5' RNA molecule might interact with both MSI2 monomers and oligomers (328). However, ${}_{63}\text{AUUAA}_{67} > {}_{63}\text{CAACU}_{67}$ was lethal for the infectious virus (Figure 4.2.6A), suggesting that ${}_{63}\text{AUUAA}_{67}$ might be the only MSI2 binding site within the CHIKV 5' end. Based on our results, MSI2 was proposed to regulate CHIKV genome replication by functioning as a molecular switch, which requires synergistic interactions with other virus or host factors, e.g., DHX9. Reduced binding affinity between MSI2 to RNA might result in inefficient switching from translation to replication. For picornaviruses, a recent study using virus infection real-time imaging has demonstrated the shutdown of virus translation was independent of the initiation of replication (329). This could suggest that if the MSI2-RNA interaction was impaired, CHIKV genome translation and replication would both terminate due to the inability to switch. Since the *trans*-complementation system uncouples genome replication from translation, the data from Figure 4.2.6B might not accurately represent the failed molecular switching between translation and replication shown in the infectious virus, where the genome replication was in fact completely inhibited.

6.2 Model of the proposed mechanism of MSI2 regulating the initiation of CHIKV negative-sense strand synthesis

In Chapter 3 and 4, phenotypic analysis using independent approaches demonstrated that MSI2 was specifically required for CHIKV genome replication in human cells. In Chapter 5, results from EMSA analysis confirmed the direct interaction between MSI2 and the 5' UTR of the CHIKV genome via $_{63}\text{AUUAA}_{67}$ located at the single-stranded regions upstream of the AUG start codon, while the secondary stem-loop structures were not found to be involved in this interaction. Based on these results and a previous phenotypic study on the facilitative role of DHX9 on CHIKV translation but not genome replication (106), a hypothesised model was proposed explaining how translation and genome replication of CHIKV are controlled by alternating interactions between MSI2, DHX9 and the CHIKV genome (Figure 6.1). Following the release of the CHIKV genome into the cytoplasm, nsP3 produced from inefficient genome translation recruits DHX9 to the CHIKV 5' end where it functions as a helicase to unwind RNA secondary structures, leading to the efficient initiation of ORF-1 translation of the viral replicase proteins (106). As the virus lifecycle proceeds and the level of nsPs increase, DHX9 is proteolytically cleaved by interacting with nsP2, which has protease activity. Following that, the 5' RNA structures are restored, which might promote the dissociation of eukaryotic translation factors to reveal the MSI2 binding site. MSI2 then recognises and binds to the CHIKV 5' UTR via $_{63}\text{AUUAA}_{67}$. CHIKV nsPs which directly or indirectly interact with MSI2 are brought to close contact with the virus genome to initiate genome circularisation and the synthesis of negative-strand RNA. Based on this model, disruption of the MSI2-RNA interaction might lead to inefficient genome circularisation, which subsequently impedes the initiation of negative-sense strand synthesis and production of nascent genomic RNA. Moreover, the pro-viral association of other cellular proteins to the CHIKV genome may also be affected due to mutagenesis of the MSI2 binding site. It is worth noting that the genomic 5'-3' interaction in this proposed mechanism might be host-dependent, as the genetic modulation of the stability of the RNA secondary structures has been shown to have a

more significant impact in invertebrate cells (330). In conclusion, our model not only provides one of the possible mechanisms of how CHIKV translates and replicates, but also explains the reason for the lethal phenotype seen in the infectious virus.

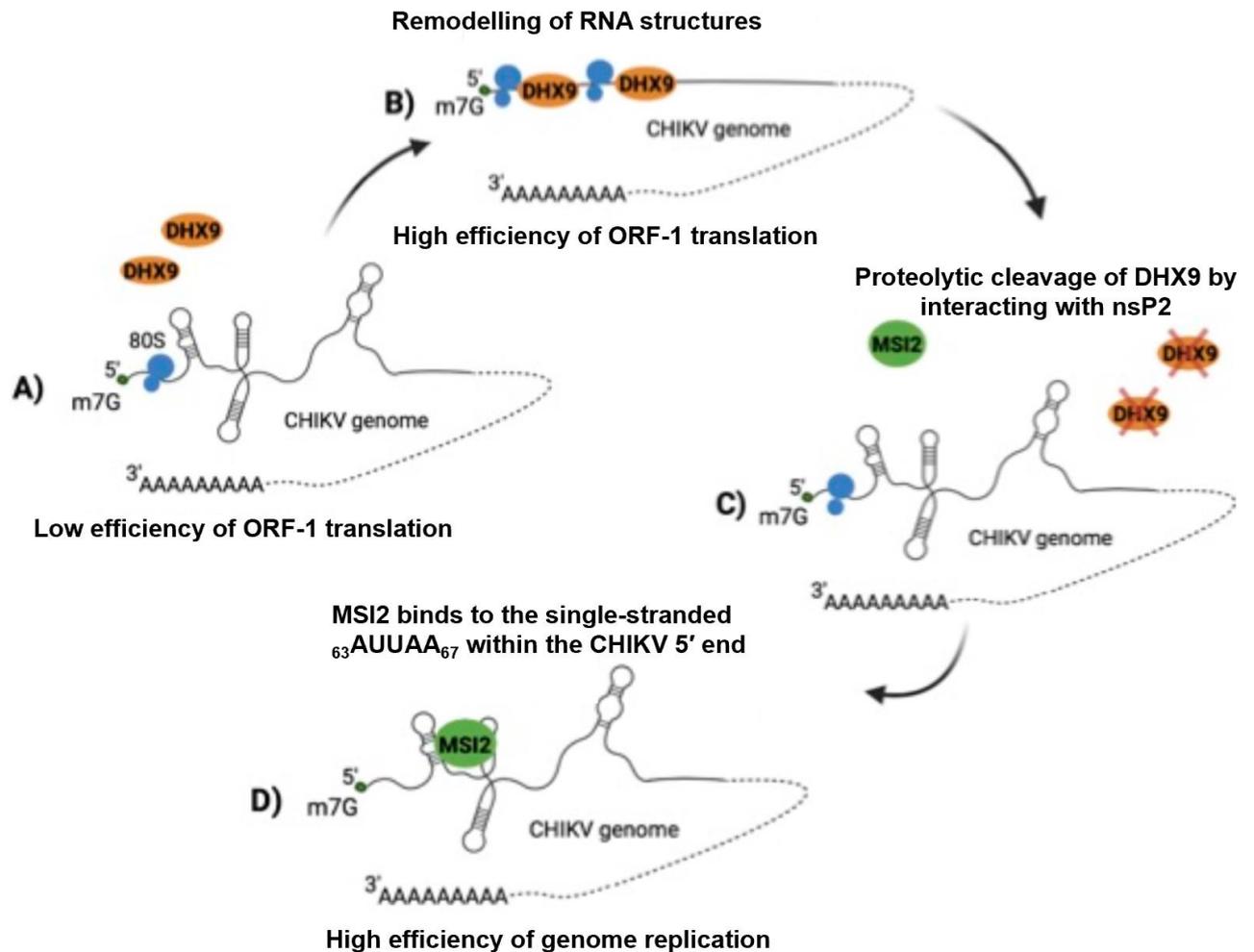


Figure 6.1. Proposed mechanism by which MSI2 and DHX9 influence CHIKV translation/replication switch.

A). RNA structure in the 5' region of the CHIKV genome inhibits ORF-1 translation. **B).** DHX9 interacts with the 5' region and remodels RNA structure, increasing the translation of ORF-1 nsPs. **C).** nsP2 promotes DHX9 degradation in proteasomes. **D).** RNA structures are restored, MSI2 interacts with the 5' UTR via ${}_{63}\text{AUUAA}_{67}$. CHIKV nsPs which directly or indirectly interact with MSI2 are subsequently brought close to the virus genome to promote CHIKV genome replication.

6.3 Future perspectives

The current study predominantly focused on MSI2 of the Musashi protein family. Despite its direct involvement in the virus lifecycle, MSI2 is an important cellular protein responsible for regulating the expression of mRNA at the translational level, thus downstream molecules of MSI2-related cellular pathways would be interesting to explore. For example, MSI2 affects the stability of IL-6 signal transducer mRNA of the IL-6 signalling pathway by facilitating its degradation (258). IL-6 is a pleiotropic cytokine involved in many important cellular processes, such as immune response and haematopoiesis (331). Further studies investigating the IL-6 expression and cellular location, or the effect of IL-6 knockdown would be beneficial to unveil the mechanism of CHIKV replication. Interestingly, multiple MSI2 isoforms, presumably due to alternative splicing, were detected in both Huh7 and RD cells (Figure 3.2.1). The expression of all isoforms was knocked down when treated with MSI2-specific si/shRNA (Figure 4.2.1C & 4.2.4A), which was indicative of functional redundancy. It would be worth investigating the sequence of individual isoforms and whether they are equally exploited by CHIKV during infection.

As one of the paralogs of MSI2, MSI1 was not included due to its low level of expression in RD cells. MSI1 has been shown to interact with the 3' UTR of Zika virus to enhance virus replication (263). Although MSI1 and MSI2 share high sequence similarity which leads to functional redundancy, their interactions with the CHIKV genome might or might not be compatible in cells which express high levels of both proteins, such as Huh7 cells. Given MSI1 was equally required for CHIKV replication as MSI2 in Huh7 cells (Figure 4.2.2), a reasonable starting point for future experiments would be determining whether CHIKV 5' end promiscuously interact with MSI1 and MSI2, or it interacts with each protein at different stages during the virus lifecycle. The other paralogs of the Musashi protein family, including TARDBP, DAZAP1, and HNRNP A, B, D and D-like, are less likely to play roles in CHIKV replication, as evidence has shown that they not only share low sequence homology to MSI1, but also do not interact with Zika

virus (256). Furthermore, it is also possible that the interaction between MSI2 and CHIKV nsPs is direct, as indicated in Figure 3.2.6. Besides the direct interaction of MSI2-RNA via $_{63}\text{AUUAA}_{67}$, MSI2 might be able to maintain association with CHIKV RNA by directly interacting with RNA-bound nsPs. Further experiments will be needed to investigate the interaction between MSI2 and the four CHIKV nsPs, such as immunoprecipitation of cells co-transfected with tagged-MSI2 with tagged-nsPs, and immunofluorescent analysis of colocalisation between MSI2 and nsPs.

From the aspect of CHIKV, the emphasis of this study was on the 5' nt 1-303, which was used as the RNA bait in the preliminary immunoprecipitation and proteomics analysis assay. The interaction of MSI2 and CHIKV RNA should be extended to other regions of the genome, especially the 3' UTR. Different core promoter elements for the synthesis of positive- and negative-sense strand RNA are located within the CHIKV 5' as well as 3' UTR. Besides virus replication, the 3' UTR of CHIKV plays an important role in modulating pathogenesis, host range and cell tropism (332). The 19 nt CSE is highly conserved in most alphaviruses and is immediately followed by the poly (A) tail, which is at least 12nt for efficient binding to the poly (A) binding protein (96). Mutations of this CSE have different effects on vertebrate and invertebrate cells, implying its functional diversity to interact with different host proteins (95, 333). In addition, the repeated sequence elements (RSEs) located upstream of the CSE may also contribute to vector tropism due to the differences in length and number of copies among alphaviruses (94). Therefore, MSI2 might also interact with the CHIKV 3' UTR by direct binding or cellular/viral protein intermediate to promote the initiation of virus genome replication. If so, it would be intriguing to perform mutagenesis within the 3' UTR to investigate the MSI2 binding site(s) and how it would affect CHIKV genome replication.

6.4 Conclusion

In summary, for the first time, this study biochemically and phenotypically identified MSI2 as a host cell factor involved in CHIKV genome replication. Based on our finding of the direct binding between MSI2 and the 5' UTR ₆₃AUUAA₆₇ of the CHIKV genome, together with the interaction between MSI2 and CHIKV nsPs, a model mechanism was hypothesised in which MSI2 might be required for the initiation of negative-sense strand synthesis by functioning as a molecular switch between CHIKV protein translation and genome replication. Future therapeutic interventions against CHIKV, and potentially other arboviruses, may benefit from this study by using MSI2 as a target for developing antivirals.

REFERENCES

1. Walker PJ, Siddell SG, Lefkowitz EJ, Mushegian AR, Dempsey DM, Dutilh BE, et al. Changes to virus taxonomy and the International Code of Virus Classification and Nomenclature ratified by the International Committee on Taxonomy of Viruses (2019). *Archives of virology*. 2019;164(9):2417-29.
2. Fong SW, Kini RM, Ng LFP. Mosquito Saliva Reshapes Alphavirus Infection and Immunopathogenesis. *J Virol*. 2018;92(12).
3. McPherson RL, Abraham R, Sreekumar E, Ong SE, Cheng SJ, Baxter VK, et al. ADP-ribosylhydrolase activity of Chikungunya virus macrodomain is critical for virus replication and virulence. *P Natl Acad Sci USA*. 2017;114(7):1666-71.
4. Solignat M, Gay B, Higgs S, Briant L, Devaux C. Replication cycle of chikungunya: a re-emerging arbovirus. *Virology*. 2009;393(2):183-97.
5. Powers AM, Brault AC, Shirako Y, Strauss EG, Kang W, Strauss JH, et al. Evolutionary relationships and systematics of the alphaviruses. *J Virol*. 2001;75(21):10118-31.
6. Lumsden WH. An epidemic of virus disease in Southern Province, Tanganyika territory, in 1952–1953 II. General description and epidemiology. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. 1955;49(1):33-57.
7. Petitdemange C, Wauquier N, Vieillard V. Control of immunopathology during chikungunya virus infection. *Journal of Allergy and Clinical Immunology*. 2015;135(4):846-55.
8. Schwartz O, Albert ML. Biology and pathogenesis of chikungunya virus. *Nature Reviews Microbiology*. 2010;8(7):491-500.

9. Diallo M, Thonnon J, Traore-Lamizana M, Fontenille D. Vectors of Chikungunya virus in Senegal: current data and transmission cycles. *The American journal of tropical medicine and hygiene*. 1999;60(2):281-6.
10. Filipe AR, Pinto MR. Arbovirus studies in Luanda, Angola: 2. Virological and serological studies during an outbreak of dengue-like disease caused by the Chikungunya virus. *Bulletin of the World Health Organization*. 1973;49(1):37.
11. Eisenhut M, Schwarz T, Hegenscheid B. Seroprevalence of dengue, chikungunya and Sindbis virus infections in German aid workers. *Infection*. 1999;27(2):82-5.
12. Rodhain F, Carteron B, Laroche R, Hannoun C. Human arbovirus infections in Burundi: results of a seroepidemiologic survey, 1980-1982. *Bulletin de la Societe de pathologie exotique et de ses filiales*. 1987;80(2):155-61.
13. Kuniholm MH, Wolfe ND, Huang CY-h, Mpoudi-Ngole E, Tamoufe U, Burke DS, et al. Seroprevalence and distribution of Flaviviridae, Togaviridae, and Bunyaviridae arboviral infections in rural Cameroonian adults. *The American journal of tropical medicine and hygiene*. 2006;74(6):1078-83.
14. Guilherme J, Gonella-Legall C, Legall F, Nakoume E, Vincent J. Seroprevalence of five arboviruses in Zebu cattle in the Central African Republic. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. 1996;90(1):31-3.
15. Pastorino B, Muyembe-Tamfum J, Bessaud M, Tock F, Tolou H, Durand J, et al. Epidemic resurgence of Chikungunya virus in democratic Republic of the Congo: identification of a new central African strain. *Journal of medical virology*. 2004;74(2):277-82.
16. Peyrefitte CN, Bessaud M, Pastorino BA, Gravier P, Plumet S, Merle OL, et al. Circulation of Chikungunya virus in Gabon, 2006–2007. *Journal of medical virology*.

2008;80(3):430-3.

17. Jentes ES, Robinson J, Johnson BW, Conde I, Sakouvougui Y, Iverson J, et al. Acute arboviral infections in Guinea, west Africa, 2006. *The American journal of tropical medicine and hygiene*. 2010;83(2):388.
18. Sergon K, Njuguna C, Kalani R, Ofula V, Onyango C, Konongoi LS, et al. Seroprevalence of chikungunya virus (CHIKV) infection on Lamu Island, Kenya, October 2004. *The American journal of tropical medicine and hygiene*. 2008;78(2):333-7.
19. Kaschula V, Van Dellen, AF**, De Vos V. Some infectious diseases of wild vervet monkeys (*Cercopithecus aethiops pygerythrus*) in South Africa. *Journal of the South African Veterinary Association*. 1978;49(3):223-7.
20. Pistone T, Ezzedine K, Schuffenecker I, Receveur M-C, Malvy D. An imported case of Chikungunya fever from Madagascar: use of the sentinel traveller for detecting emerging arboviral infections in tropical and European countries. *Travel medicine and infectious disease*. 2009;7(1):52-4.
21. Van den Bosch C, Lloyd G. Chikungunya fever as a risk factor for endemic Burkitt's lymphoma in Malawi. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. 2000;94(6):704-5.
22. Moore Dá, Causey O, Carey D, Reddy S, Cooke A, Akinkugbe F, et al. Arthropod-borne viral infections of man in Nigeria, 1964–1970. *Annals of Tropical Medicine & Parasitology*. 1975;69(1):49-64.
23. Weinbren M, Haddow AJ, Williams M. The occurrence of chikungunya virus in Uganda I. Isolation from mosquitoes. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. 1958;52(3):253-62.

24. Pistone T, Ezzedine K, Boisvert M, Receveur MC, Schuffenecker I, Zeller H, et al. Cluster of chikungunya virus infection in travelers returning from Senegal, 2006. *Journal of travel medicine*. 2009;16(4):286-8.
25. Woodruff A, Bowen E, Platt G. Viral infections in travellers from tropical Africa. *Br Med J*. 1978;1(6118):956-8.
26. Gould LH, Osman MS, Farnon EC, Griffith KS, Godsey MS, Karch S, et al. An outbreak of yellow fever with concurrent chikungunya virus transmission in South Kordofan, Sudan, 2005. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. 2008;102(12):1247-54.
27. Burt FJ, Chen W, Miner JJ, Lenschow DJ, Merits A, Schnettler E, et al. Chikungunya virus: an update on the biology and pathogenesis of this emerging pathogen. *The Lancet Infectious Diseases*. 2017;17(4):e107-e17.
28. Weaver SC, Forrester NL. Chikungunya: Evolutionary history and recent epidemic spread. *Antiviral research*. 2015;120:32-9.
29. Lanciotti RS, Valadere AM. Transcontinental movement of Asian genotype chikungunya virus. *Emerging infectious diseases*. 2014;20(8):1400.
30. Ng LF. Immunopathology of Chikungunya virus infection: lessons learned from patients and animal models. *Annual review of virology*. 2017;4:413-27.
31. Murugan SB, Sathishkumar R. Chikungunya infection: A potential re-emerging global threat. *Asian Pacific journal of tropical medicine*. 2016;9(10):933-7.
32. Josseran L, Paquet C, Zehgnoun A, Caillere N, Le Tertre A, Solet J-L, et al. Chikungunya disease outbreak, Reunion island. *Emerging infectious diseases*. 2006;12(12):1994.

33. Pinheiro TJ, Guimarães LF, Silva MTT, Soares CN. Neurological manifestations of Chikungunya and Zika infections. *Arquivos de neuro-psiquiatria*. 2016;74:937-43.
34. Renault P, Solet J-L, Sissoko D, Balleydier E, Larrieu S, Filleul L, et al. A major epidemic of chikungunya virus infection on Reunion Island, France, 2005-2006. *The American journal of tropical medicine and hygiene*. 2007;77(4):727-31.
35. Tsetsarkin KA, Vanlandingham DL, McGee CE, Higgs S. A single mutation in chikungunya virus affects vector specificity and epidemic potential. *PLoS pathogens*. 2007;3(12):e201.
36. Chatterjee PK, Vashishtha M, Kielian M. Biochemical consequences of a mutation that controls the cholesterol dependence of Semliki Forest virus fusion. *J Virol*. 2000;74(4):1623-31.
37. Arankalle VA, Shrivastava S, Cherian S, Gunjekar RS, Walimbe AM, Jadhav SM, et al. Genetic divergence of Chikungunya viruses in India (1963-2006) with special reference to the 2005-2006 explosive epidemic. *Journal of General Virology*. 2007;88(7):1967-76.
38. Leparc-Goffart I, Nougairede A, Cassadou S, Prat C, De Lamballerie X. Chikungunya in the Americas. *The Lancet*. 2014;383(9916):514.
39. Sy AK, Saito-Obata M, Medado IA, Tohma K, Dapat C, Segubre-Mercado E, et al. Molecular characterization of chikungunya virus, Philippines, 2011–2013. *Emerging infectious diseases*. 2016;22(5):887.
40. Control ECfDPa. Geographical distribution of chikungunya virus disease cases reported worldwide, 2020: European Centre for Disease Prevention and Control; 2021 [Available from: <https://www.ecdc.europa.eu/en/publications-data/geographical-distribution-chikungunya-virus-disease-cases-reported-worldwide-2020>].

41. Thiberville S-D, Boisson V, Gaudart J, Simon F, Flahault A, De Lamballerie X. Chikungunya fever: a clinical and virological investigation of outpatients on Reunion Island, South-West Indian Ocean. *PLoS neglected tropical diseases*. 2013;7(1):e2004.
42. Thiberville S-D, Moyen N, Dupuis-Maguiraga L, Nougairède A, Gould EA, Roques P, et al. Chikungunya fever: epidemiology, clinical syndrome, pathogenesis and therapy. *Antiviral research*. 2013;99(3):345-70.
43. Miner JJ, Aw Yeang HX, Fox JM, Taffner S, Malkova ON, Oh ST, et al. Brief report: chikungunya viral arthritis in the United States: a mimic of seronegative rheumatoid arthritis. *Arthritis & rheumatology*. 2015;67(5):1214-20.
44. Mahendradas P, Avadhani K, Shetty R. Chikungunya and the eye: a review. *Journal of ophthalmic inflammation and infection*. 2013;3(1):1-9.
45. Ramful D, Carbonnier M, Pasquet M, Bouhmani B, Ghazouani J, Noormahomed T, et al. Mother-to-child transmission of Chikungunya virus infection. *The Pediatric infectious disease journal*. 2007;26(9):811-5.
46. Gérardin P, Sampéris S, Ramful D, Boumahni B, Bintner M, Alessandri J-L, et al. Neurocognitive outcome of children exposed to perinatal mother-to-child Chikungunya virus infection: the CHIMERE cohort study on Reunion Island. *PLoS neglected tropical diseases*. 2014;8(7):e2996.
47. Mohan A, Kiran D, Manohar IC, Kumar DP. Epidemiology, clinical manifestations, and diagnosis of Chikungunya fever: lessons learned from the re-emerging epidemic. *Indian journal of dermatology*. 2010;55(1):54.
48. Borgherini G, Poubeau P, Staikowsky F, Lory M, Moullec NL, Becquart JP, et al. Outbreak of chikungunya on Reunion Island: early clinical and laboratory features in 157 adult patients. *Clinical infectious diseases*. 2007;44(11):1401-7.

49. Schilte C, Staikovskiy F, Couderc T, Madec Y, Carpentier F, Kassab S, et al. Chikungunya virus-associated long-term arthralgia: a 36-month prospective longitudinal study. *PLoS neglected tropical diseases*. 2013;7(3):e2137.
50. Marimoutou C, Vivier E, Oliver M, Boutin J-P, Simon F. Morbidity and impaired quality of life 30 months after chikungunya infection: comparative cohort of infected and uninfected French military policemen in Reunion Island. *Medicine*. 2012;91(4):212-9.
51. Chow A, Her Z, Ong EK, Chen J-m, Dimatatac F, Kwek DJ, et al. Persistent arthralgia induced by Chikungunya virus infection is associated with interleukin-6 and granulocyte macrophage colony-stimulating factor. *Journal of Infectious Diseases*. 2011;203(2):149-57.
52. Couderc T, Chrétien F, Schilte C, Disson O, Brigitte M, Guivel-Benhassine F, et al. A mouse model for Chikungunya: young age and inefficient type-I interferon signaling are risk factors for severe disease. *PLoS pathogens*. 2008;4(2):e29.
53. Gardner J, Anraku I, Le TT, Larcher T, Major L, Roques P, et al. Chikungunya virus arthritis in adult wild-type mice. *J Virol*. 2010;84(16):8021-32.
54. Medzhitov R, Janeway Jr C. Innate immune recognition: mechanisms and pathways. *Immunological reviews*. 2000;173:89-97.
55. Teng T-S, Kam Y-W, Tan JJ, Ng LF. Host response to Chikungunya virus and perspectives for immune-based therapies. *Future Virology*. 2011;6(8):975-84.
56. Stetson DB, Medzhitov R. Type I interferons in host defense. *Immunity*. 2006;25(3):373-81.
57. Her Z, Teng TS, Tan JJ, Teo TH, Kam YW, Lum FM, et al. Loss of TLR3 aggravates

CHIKV replication and pathology due to an altered virus-specific neutralizing antibody response. *EMBO molecular medicine*. 2015;7(1):24-41.

58. Kawai T, Akira S. The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. *Nature immunology*. 2010;11(5):373-84.

59. Bréhin A-C, Casadémont I, Frenkiel M-P, Julier C, Sakuntabhai A, Desprès P. The large form of human 2', 5'-Oligoadenylate Synthetase (OAS3) exerts antiviral effect against Chikungunya virus. *Virology*. 2009;384(1):216-22.

60. Teng T-S, Kam Y-W, Lee B, Hapuarachchi HC, Wimal A, Ng L-C, et al. A systematic meta-analysis of immune signatures in patients with acute chikungunya virus infection. *The Journal of infectious diseases*. 2015;211(12):1925-35.

61. Noret M, Herrero L, Rulli N, Rolph M, Smith PN, Li RW, et al. Interleukin 6, RANKL, and osteoprotegerin expression by chikungunya virus-infected human osteoblasts. *The Journal of infectious diseases*. 2012;206(3):455-7.

62. Hoarau J-J, Bandjee M-CJ, Trotot PK, Das T, Li-Pat-Yuen G, Dassa B, et al. Persistent chronic inflammation and infection by Chikungunya arthritogenic alphavirus in spite of a robust host immune response. *The Journal of Immunology*. 2010;184(10):5914-27.

63. Li Y, Bäckesjö C-M, Haldosén L-A, Lindgren U. IL-6 receptor expression and IL-6 effects change during osteoblast differentiation. *Cytokine*. 2008;43(2):165-73.

64. Lu B, Rutledge BJ, Gu L, Fiorillo J, Lukacs NW, Kunkel SL, et al. Abnormalities in monocyte recruitment and cytokine expression in monocyte chemoattractant protein 1-deficient mice. *The Journal of experimental medicine*. 1998;187(4):601-8.

65. Romano M, Sironi M, Toniatti C, Polentarutti N, Fruscella P, Ghezzi P, et al. Role

of IL-6 and its soluble receptor in induction of chemokines and leukocyte recruitment. *Immunity*. 1997;6(3):315-25.

66. Chen R, Wang E, Tsetsarkin KA, Weaver SC. Chikungunya virus 3' untranslated region: adaptation to mosquitoes and a population bottleneck as major evolutionary forces. *PLoS pathogens*. 2013;9(8):e1003591.

67. Poo YS, Nakaya H, Gardner J, Larcher T, Schroder WA, Le TT, et al. CCR2 deficiency promotes exacerbated chronic erosive neutrophil-dominated chikungunya virus arthritis. *J Virol*. 2014;88(12):6862-72.

68. Her Z, Malleret B, Chan M, Ong EK, Wong S-C, Kwek DJ, et al. Active infection of human blood monocytes by Chikungunya virus triggers an innate immune response. *The Journal of Immunology*. 2010;184(10):5903-13.

69. Kam Y-W, Simarmata D, Chow A, Her Z, Teng T-S, Ong EK, et al. Early appearance of neutralizing immunoglobulin G3 antibodies is associated with chikungunya virus clearance and long-term clinical protection. *Journal of Infectious Diseases*. 2012;205(7):1147-54.

70. Hawman DW, Stoermer KA, Montgomery SA, Pal P, Oko L, Diamond MS, et al. Chronic joint disease caused by persistent Chikungunya virus infection is controlled by the adaptive immune response. *J Virol*. 2013;87(24):13878-88.

71. Lum F-M, Teo T-H, Lee WW, Kam Y-W, Rénia L, Ng LF. An essential role of antibodies in the control of Chikungunya virus infection. *The Journal of Immunology*. 2013;190(12):6295-302.

72. Weger-Lucarelli J, Aliota MT, Kamlangdee A, Osorio JE. Identifying the role of E2 domains on alphavirus neutralization and protective immune responses. *PLoS neglected tropical diseases*. 2015;9(10):e0004163.

73. Weber C, Büchner SM, Schnierle BS. A small antigenic determinant of the Chikungunya virus E2 protein is sufficient to induce neutralizing antibodies which are partially protective in mice. *PLoS neglected tropical diseases*. 2015;9(4):e0003684.
74. Goh LY, Hobson-Peters J, Prow NA, Gardner J, Bielefeldt-Ohmann H, Suhrbier A, et al. Monoclonal antibodies specific for the capsid protein of chikungunya virus suitable for multiple applications. *Journal of General Virology*. 2015;96(3):507-12.
75. Warter L, Lee CY, Thiagarajan R, Grandadam M, Lebecque S, Lin RT, et al. Chikungunya virus envelope-specific human monoclonal antibodies with broad neutralization potency. *The Journal of Immunology*. 2011;186(5):3258-64.
76. Takada A, Kawaoka Y. Antibody-dependent enhancement of viral infection: molecular mechanisms and in vivo implications. *Reviews in medical virology*. 2003;13(6):387-98.
77. Lum F-M, Couderc T, Chia B-S, Ong R-Y, Her Z, Chow A, et al. Antibody-mediated enhancement aggravates chikungunya virus infection and disease severity. *Scientific reports*. 2018;8(1):1-14.
78. Wauquier N, Becquart P, Nkoghe D, Padilla C, Ndjoyi-Mbiguino A, Leroy EM. The acute phase of Chikungunya virus infection in humans is associated with strong innate immunity and T CD8 cell activation. *Journal of Infectious Diseases*. 2011;204(1):115-23.
79. Teo T-H, Lum F-M, Claser C, Lulla V, Lulla A, Merits A, et al. A pathogenic role for CD4+ T cells during Chikungunya virus infection in mice. *The Journal of Immunology*. 2013;190(1):259-69.
80. Poo YS, Rudd PA, Gardner J, Wilson JA, Larcher T, Colle M-A, et al. Multiple immune factors are involved in controlling acute and chronic chikungunya virus

infection. *PLoS neglected tropical diseases*. 2014;8(12):e3354.

81. Teo T-H, Chan Y-H, Lee WW, Lum F-M, Amrun SN, Her Z, et al. Fingolimod treatment abrogates chikungunya virus-induced arthralgia. *Science Translational Medicine*. 2017;9(375):eaal1333.

82. Strauss JH, Strauss EG. The alphaviruses: gene expression, replication, and evolution. *Microbiological reviews*. 1994;58(3):491-562.

83. Kendall C, Khalid H, Müller M, Banda DH, Kohl A, Merits A, et al. Structural and phenotypic analysis of Chikungunya virus RNA replication elements. *Nucleic acids research*. 2019;47(17):9296-312.

84. Meshram CD, Agback P, Shiliaev N, Urakova N, Mobley JA, Agback T, et al. Multiple Host Factors Interact with the Hypervariable Domain of Chikungunya Virus nsP3 and Determine Viral Replication in Cell-Specific Mode. *J Virol*. 2018;92(16).

85. Frolov I, HARDY R, RICE CM. Cis-acting RNA elements at the 5' end of Sindbis virus genome RNA regulate minus- and plus-strand RNA synthesis. *Rna*. 2001;7(11):1638-51.

86. Kulasegaran-Shylini R, Atasheva S, Gorenstein DG, Frolov I. Structural and functional elements of the promoter encoded by the 5' untranslated region of the Venezuelan equine encephalitis virus genome. *J Virol*. 2009;83(17):8327-39.

87. Mendes A, Kuhn RJ. Alphavirus nucleocapsid packaging and assembly. *Viruses*. 2018;10(3):138.

88. Ou J-H, Strauss EG, Strauss JH, Simons K. The 5'-terminal sequences of the genomic RNAs of several alphaviruses. *Journal of molecular biology*. 1983;168(1):1-15.

89. Nickens DG, Hardy RW. Structural and functional analyses of stem-loop 1 of the Sindbis virus genome. *Virology*. 2008;370(1):158-72.
90. Niesters H, Strauss JH. Defined mutations in the 5'nontranslated sequence of Sindbis virus RNA. *J Virol*. 1990;64(9):4162-8.
91. Niesters HGM, Strauss JH. Mutagenesis of the Conserved 51-Nucleotide Region of Sindbis Virus. *J Virol*. 1990;64(4):1639-47.
92. Fayzulin R, Frolov I. Changes of the secondary structure of the 5' end of the Sindbis virus genome inhibit virus growth in mosquito cells and lead to accumulation of adaptive mutations. *J Virol*. 2004;78(10):4953-64.
93. Michel G, Petrakova O, Atasheva S, Frolov I. Adaptation of Venezuelan equine encephalitis virus lacking 51-nt conserved sequence element to replication in mammalian and mosquito cells. *Virology*. 2007;362(2):475-87.
94. Pfeffer M, Kinney RM, Kaaden O-R. The alphavirus 3'-nontranslated region: size heterogeneity and arrangement of repeated sequence elements. *Virology*. 1998;240(1):100-8.
95. Kuhn RJ, Hong Z, Strauss JH. Mutagenesis of the 3'nontranslated region of Sindbis virus RNA. *J Virol*. 1990;64(4):1465-76.
96. Hardy RW, Rice CM. Requirements at the 3' end of the sindbis virus genome for efficient synthesis of minus-strand RNA. *J Virol*. 2005;79(8):4630-9.
97. Hardy RW. The role of the 3' terminus of the Sindbis virus genome in minus-strand initiation site selection. *Virology*. 2006;345(2):520-31.
98. Barnhart MD, Moon SL, Emch AW, Wilusz CJ, Wilusz J. Changes in cellular mRNA

stability, splicing, and polyadenylation through HuR protein sequestration by a cytoplasmic RNA virus. *Cell reports*. 2013;5(4):909-17.

99. Dickson AM, Anderson JR, Barnhart MD, Sokoloski KJ, Oko L, Opyrchal M, et al. Dephosphorylation of HuR protein during alphavirus infection is associated with HuR relocalization to the cytoplasm. *Journal of Biological Chemistry*. 2012;287(43):36229-38.

100. Fuller SD. The T= 4 envelope of Sindbis virus is organized by interactions with a complementary T= 3 capsid. *Cell*. 1987;48(6):923-34.

101. Cheng RH, Kuhn RJ, Olson NH, Rossmann MG, Choi H-K, Smith TJ, et al. Nucleocapsid and glycoprotein organization in an enveloped virus. *Cell*. 1995;80(4):621-30.

102. Paredes AM, Simon MN, Brown DT. The mass of the Sindbis virus nucleocapsid suggests it has T= 4 icosahedral symmetry. *Virology*. 1992;187(1):329-32.

103. Vogel R, Provencher S, Von Bonsdorff C-H, Adrian M, Dubochet J. Envelope structure of Semliki Forest virus reconstructed from cryo-electron micrographs. *Nature*. 1986;320(6062):533-5.

104. Wahlberg J, Boere W, Garoff H. The heterodimeric association between the membrane proteins of Semliki Forest virus changes its sensitivity to low pH during virus maturation. *J Virol*. 1989;63(12):4991-7.

105. Niranjana A, Gupta P. Recent Trends in Chikungunya Virus diagnosis.

106. Matkovic R, Bernard E, Fontanel S, Eldin P, Chazal N, Hersi DH, et al. The Host DHX9 DExH-Box Helicase Is Recruited to Chikungunya Virus Replication Complexes for Optimal Genomic RNA Translation. *J Virol*. 2019;93(4).

107. Remenyi R, Gao Y, Hughes RE, Curd A, Zothner C, Peckham M, et al. Persistent Replication of a Chikungunya Virus Replicon in Human Cells Is Associated with Presence of Stable Cytoplasmic Granules Containing Nonstructural Protein 3. *J Virol.* 2018;92(16).
108. Rupp JC, Sokoloski KJ, Gebhart NN, Hardy RW. Alphavirus RNA synthesis and non-structural protein functions. *The Journal of general virology.* 2015;96(Pt 9):2483.
109. Vasiljeva L, Merits A, Auvinen P, Kääriäinen L. Identification of a Novel Function of the Alphavirus Capping Apparatus: RNA 5'-TRIPHOSPHATASE ACTIVITY OF Nsp2. *Journal of Biological Chemistry.* 2000;275(23):17281-7.
110. Singh A, Kumar A, Yadav R, Uversky VN, Giri R. Deciphering the dark proteome of Chikungunya virus. *Scientific reports.* 2018;8(1):1-10.
111. Li C, Guillén J, Rabah N, Blanjoie A, Debart F, Vasseur J-J, et al. mRNA capping by Venezuelan equine encephalitis virus nsP1: functional characterization and implications for antiviral research. *J Virol.* 2015;89(16):8292-303.
112. Ahola T, Kääriäinen L. 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.* 1995;92(2):507-11.
113. Delang L, Li C, Tas A, Quérat G, Albulescu I, De Burghgraeve T, et al. The viral capping enzyme nsP1: a novel target for the inhibition of chikungunya virus infection. *Scientific reports.* 2016;6(1):1-10.
114. Filipowicz W, Furuichi Y, Sierra J, Muthukrishnan S, Shatkin A, Ochoa S. A protein binding the methylated 5'-terminal sequence, m⁷GpppN, of eukaryotic messenger RNA. *Proceedings of the National Academy of Sciences.* 1976;73(5):1559-63.

115. Liu H, Kiledjian M. Decapping the message: a beginning or an end. *Biochemical Society Transactions*. 2006;34(1):35-8.
116. Nallagatla SR, Toroney R, Bevilacqua PC. A brilliant disguise for self RNA: 5' cap end and internal modifications of primary transcripts suppress elements of innate immunity. *RNA biology*. 2008;5(3):140-4.
117. Sokoloski K, Haist K, Morrison T, Mukhopadhyay S, Hardy R. Noncapped alphavirus genomic RNAs and their role during infection. *J Virol*. 2015;89(11):6080-92.
118. Lampio A, Kilpeläinen I, Pesonen S, Karhi K, Auvinen P, Somerharju P, et al. Membrane binding mechanism of an RNA virus-capping enzyme. *Journal of Biological Chemistry*. 2000;275(48):37853-9.
119. Spuul P, Salonen A, Merits A, Jokitalo E, Kääriäinen 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(2):872-83.
120. Zhang K, Law Y-S, Law MCY, Tan YB, Wirawan M, Luo D. Structural insights into viral RNA capping and plasma membrane targeting by Chikungunya virus nonstructural protein 1. *Cell host & microbe*. 2021;29(5):757-64. e3.
121. Bakhache W, Neyret A, Bernard E, Merits A, Briant L. Palmitoylated cysteines in chikungunya virus nsP1 are critical for targeting to cholesterol-rich plasma membrane microdomains with functional consequences for viral genome replication. *J Virol*. 2020;94(10):e02183-19.
122. Tomar S, Narwal M, Harms E, Smith JL, Kuhn RJ. Heterologous production, purification and characterization of enzymatically active Sindbis virus nonstructural protein nsP1. *Protein expression and purification*. 2011;79(2):277-84.

123. Gottipati K, Woodson M, Choi KH. Membrane binding and rearrangement by chikungunya virus capping enzyme nsP1. *Virology*. 2020;544:31-41.
124. Jones R, Bragagnolo G, Arranz R, Reguera J. Capping pores of alphavirus nsP1 gate membranous viral replication factories. *Nature*. 2021;589(7843):615-9.
125. Russo AT, White MA, Watowich SJ. The crystal structure of the Venezuelan equine encephalitis alphavirus nsP2 protease. *Structure*. 2006;14(9):1449-58.
126. Frolov I, Agapov E, Hoffman Jr TA, Prágai BIM, Lippa M, Schlesinger S, et al. Selection of RNA replicons capable of persistent noncytopathic replication in mammalian cells. *J Virol*. 1999;73(5):3854-65.
127. de Cedrón MG, Ehsani N, Mikkola ML, García JA, Kääriäinen L. RNA helicase activity of Semliki Forest virus replicase protein NSP2. *FEBS letters*. 1999;448(1):19-22.
128. Gomatos PJ, Kääriäinen L, Keränen S, Ranki M, Sawicki DL. Semliki Forest virus replication complex capable of synthesizing 42S and 26S nascent RNA chains. *Journal of General Virology*. 1980;49(1):61-9.
129. Gorbalenya AE, Koonin EV, Donchenko AP, Blinov VM. A novel superfamily of nucleoside triphosphate-binding motif containing proteins which are probably involved in duplex unwinding in DNA and RNA replication and recombination. *FEBS letters*. 1988;235(1-2):16-24.
130. Karpe YA, Aher PP, Lole KS. NTPase and 5'-RNA triphosphatase activities of Chikungunya virus nsP2 protein. *PloS one*. 2011;6(7):e22336.
131. Rikkonen M, Peränen J, Kääriäinen L. ATPase and GTPase activities associated with Semliki Forest virus nonstructural protein nsP2. *J Virol*. 1994;68(9):5804-10.

132. Das PK, Merits A, Lulla A. Functional cross-talk between distant domains of chikungunya virus non-structural protein 2 is decisive for its RNA-modulating activity. *Journal of Biological Chemistry*. 2014;289(9):5635-53.
133. Hardy WR, Strauss JH. Processing the nonstructural polyproteins of Sindbis virus: nonstructural proteinase is in the C-terminal half of nsP2 and functions both in cis and in trans. *J Virol*. 1989;63(11):4653-64.
134. 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(11):5413-22.
135. Owen KE, Kuhn RJ. Identification of a region in the Sindbis virus nucleocapsid protein that is involved in specificity of RNA encapsidation. *J Virol*. 1996;70(5):2757-63.
136. Ten Dam E, Flint M, Ryan MD. Virus-encoded proteinases of the Togaviridae. *Journal of general virology*. 1999;80(8):1879-88.
137. de Groot RJ, Hardy WR, Shirako Y, Strauss JH. Cleavage-site preferences of Sindbis virus polyproteins containing the non-structural proteinase. Evidence for temporal regulation of polyprotein processing in vivo. *The EMBO journal*. 1990;9(8):2631-8.
138. Shin G, Yost SA, Miller MT, Elrod EJ, Grakoui A, Marcotrigiano J. Structural and functional insights into alphavirus polyprotein processing and pathogenesis. *Proceedings of the National Academy of Sciences*. 2012;109(41):16534-9.
139. Vasiljeva L, Merits A, Golubtsov A, Sizemskaja V, Kääriäinen L, Ahola T. Regulation of the sequential processing of Semliki Forest virus replicase polyprotein. *Journal of Biological Chemistry*. 2003;278(43):41636-45.

140. Bartholomeeusen K, Utt A, Coppens S, Rausalu K, Vereecken K, Arien KK, et al. A Chikungunya Virus trans-Replicase System Reveals the Importance of Delayed Nonstructural Polyprotein Processing for Efficient Replication Complex Formation in Mosquito Cells. *J Virol.* 2018;92(14).
141. Garmashova N, Gorchakov R, Frolova E, Frolov I. Sindbis virus nonstructural protein nsP2 is cytotoxic and inhibits cellular transcription. *J Virol.* 2006;80(12):5686-96.
142. Gorchakov R, Frolova E, Frolov I. Inhibition of transcription and translation in Sindbis virus-infected cells. *J Virol.* 2005;79(15):9397-409.
143. Kim KH, Rümenapf T, Strauss EG, Strauss JH. Regulation of Semliki Forest virus RNA replication: a model for the control of alphavirus pathogenesis in invertebrate hosts. *Virology.* 2004;323(1):153-63.
144. Akhrymuk I, Frolov I, Frolova EI. Sindbis virus infection causes cell death by nsP2-induced transcriptional shutoff or by nsP3-dependent translational shutoff. *J Virol.* 2018;92(23):e01388-18.
145. Akhrymuk I, Kulemzin SV, Frolova EI. Evasion of the innate immune response: the Old World alphavirus nsP2 protein induces rapid degradation of Rpb1, a catalytic subunit of RNA polymerase II. *J Virol.* 2012;86(13):7180-91.
146. Göertz GP, McNally KL, Robertson SJ, Best SM, Pijlman GP, Fros JJ. The methyltransferase-like domain of chikungunya virus nsP2 inhibits the interferon response by promoting the nuclear export of STAT1. *J Virol.* 2018;92(17):e01008-18.
147. Levy DE, Marié IJ, Durbin JE. Induction and function of type I and III interferon in response to viral infection. *Current opinion in virology.* 2011;1(6):476-86.

148. Fros JJ, Major LD, Scholte FE, Gardner J, van Hemert MJ, Suhrbier A, et al. Chikungunya virus non-structural protein 2-mediated host shut-off disables the unfolded protein response. *Journal of General Virology*. 2015;96(3):580-9.
149. Koonin EV, Gorbalenya AE, Purdy MA, Rozanov MN, Reyes GR, Bradley DW. Computer-assisted assignment of functional domains in the nonstructural polyprotein of hepatitis E virus: delineation of an additional group of positive-strand RNA plant and animal viruses. *Proceedings of the National Academy of Sciences*. 1992;89(17):8259-63.
150. Rack JGM, Perina D, Ahel I. Macrodomains: structure, function, evolution, and catalytic activities. *Annual review of biochemistry*. 2016;85:431-54.
151. Egloff M-P, Malet H, Putics A, Heinonen M, Dutartre H, Frangeul A, et al. Structural and functional basis for ADP-ribose and poly (ADP-ribose) binding by viral macro domains. *J Virol*. 2006;80(17):8493-502.
152. Malet H, Coutard B, Jamal S, Dutartre H, Papageorgiou N, Neuvonen M, 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(13):6534-45.
153. Neuvonen M, Ahola T. Differential activities of cellular and viral macro domain proteins in binding of ADP-ribose metabolites. *Journal of molecular biology*. 2009;385(1):212-25.
154. Park E, Griffin DE. The nsP3 macro domain is important for Sindbis virus replication in neurons and neurovirulence in mice. *Virology*. 2009;388(2):305-14.
155. Lulla A, Lulla V, Merits A. Macromolecular assembly-driven processing of the 2/3 cleavage site in the alphavirus replicase polyprotein. *J Virol*. 2012;86(1):553-65.

156. Dé I, Fata-Hartley C, Sawicki SG, Sawicki DL. Functional Analysis of nsP3 Phosphoprotein Mutants of Sindbis Virus. *J Virol.* 2003;77(24):13106-16.
157. LaStarza MW, Lemm JA, Rice CM. Genetic analysis of the nsP3 region of Sindbis virus: evidence for roles in minus-strand and subgenomic RNA synthesis. *J Virol.* 1994;68(9):5781-91.
158. Wang Y-F, Sawicki SG, Sawicki DL. Alphavirus nsP3 functions to form replication complexes transcribing negative-strand RNA. *J Virol.* 1994;68(10):6466-75.
159. Gao Y, Goonawardane N, Ward J, Tuplin A, Harris M. Multiple roles of the non-structural protein 3 (nsP3) alphavirus unique domain (AUD) during Chikungunya virus genome replication and transcription. *PLoS pathogens.* 2019;15(1):e1007239.
160. Götte B, Liu L, McInerney GM. The enigmatic alphavirus non-structural protein 3 (nsP3) revealing its secrets at last. *Viruses.* 2018;10(3):105.
161. Strauss EG, Levinson R, Rice CM, Dalrymple J, Strauss JH. Nonstructural proteins nsP3 and nsP4 of Ross River and O'Nyong-nyong viruses: sequence and comparison with those of other alphaviruses. *Virology.* 1988;164(1):265-74.
162. Tuittila M, Hinkkanen AE. Amino acid mutations in the replicase protein nsP3 of Semliki Forest virus cumulatively affect neurovirulence. *Journal of General Virology.* 2003;84(6):1525-33.
163. 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 pathogens.* 2011;7(11):e1002383.
164. Panas MD, Ahola T, McInerney GM. The C-terminal repeat domains of nsP3 from the Old World alphaviruses bind directly to G3BP. *J Virol.* 2014;88(10):5888-93.

165. Panas MD, Varjak M, Lulla A, Er Eng K, Merits A, Karlsson Hedestam GB, et al. Sequestration of G3BP coupled with efficient translation inhibits stress granules in Semliki Forest virus infection. *Molecular biology of the cell*. 2012;23(24):4701-12.
166. Vihinen H, Saarinen J. Phosphorylation site analysis of Semliki forest virus nonstructural protein 3. *Journal of Biological Chemistry*. 2000;275(36):27775-83.
167. Peränen J, Takkinen K, Kalkkinen N, Kääriäinen L. Semliki Forest virus-specific non-structural protein nsP3 is a phosphoprotein. *Journal of general virology*. 1988;69(9):2165-78.
168. Vihinen H, Ahola T, Tuittila M, Merits A, Kääriäinen L. Elimination of phosphorylation sites of Semliki Forest virus replicase protein nsP3. *Journal of Biological Chemistry*. 2001;276(8):5745-52.
169. Saksela K, Permi P. SH3 domain ligand binding: What's the consensus and where's the specificity? *FEBS letters*. 2012;586(17):2609-14.
170. Kim DY, Reynaud JM, Rasalousskaya A, Akhrymuk I, Mobley JA, Frolov I, et al. New World and Old World alphaviruses have evolved to exploit different components of stress granules, FXR and G3BP proteins, for assembly of viral replication complexes. *PLoS pathogens*. 2016;12(8):e1005810.
171. Cristea IM, Carroll J-WN, Rout MP, Rice CM, Chait BT, MacDonald MR. Tracking and elucidating alphavirus-host protein interactions. *Journal of Biological Chemistry*. 2006;281(40):30269-78.
172. Frolov I, Kim DY, Akhrymuk M, Mobley JA, Frolova EI. Hypervariable domain of eastern equine encephalitis virus nsP3 redundantly utilizes multiple cellular proteins for replication complex assembly. *J Virol*. 2017;91(14):e00371-17.

173. Frolova E, Gorchakov R, Garmashova N, Atasheva S, Vergara LA, Frolov I. Formation of nsP3-specific protein complexes during Sindbis virus replication. *J Virol.* 2006;80(8):4122-34.
174. Fros JJ, Domeradzka NE, Baggen J, Geertsema C, Flipse J, Vlak JM, et al. Chikungunya virus nsP3 blocks stress granule assembly by recruitment of G3BP into cytoplasmic foci. *J Virol.* 2012;86(19):10873-9.
175. Panas MD, Schulte T, Thaa B, Sandalova T, Kedersha N, Achour A, et al. Viral and cellular proteins containing FGDF motifs bind G3BP to block stress granule formation. *PLoS pathogens.* 2015;11(2):e1004659.
176. Schulte T, Liu L, Panas MD, Thaa B, Dickson N, Götte B, et al. Combined structural, biochemical and cellular evidence demonstrates that both FGDF motifs in alphavirus nsP3 are required for efficient replication. *Open biology.* 2016;6(7):160078.
177. O'Reilly EK, Kao CC. Analysis of RNA-dependent RNA polymerase structure and function as guided by known polymerase structures and computer predictions of secondary structure. *Virology.* 1998;252(2):287-303.
178. Thal MA, Wasik BR, Posto J, Hardy RW. Template requirements for recognition and copying by Sindbis virus RNA-dependent RNA polymerase. *Virology.* 2007;358(1):221-32.
179. Li M-L, Stollar V. Identification of the amino acid sequence in Sindbis virus nsP4 that binds to the promoter for the synthesis of the subgenomic RNA. *Proceedings of the National Academy of Sciences.* 2004;101(25):9429-34.
180. Li M-L, Stollar V. Distinct sites on the Sindbis virus RNA-dependent RNA polymerase for binding to the promoters for the synthesis of genomic and subgenomic RNA. *J Virol.* 2007;81(8):4371-3.

181. Lello LS, Bartholomeeusen K, Wang S, Coppens S, Fragkoudis R, Alphey L, et al. nsP4 is a major determinant of alphavirus replicase activity and template selectivity. *J Virol.* 2021;95(20):e00355-21.
182. Tomar S, Hardy RW, Smith JL, Kuhn RJ. Catalytic core of alphavirus nonstructural protein nsP4 possesses terminal adenylyltransferase activity. *J Virol.* 2006;80(20):9962-9.
183. Rubach JK, Wasik BR, Rupp JC, Kuhn RJ, Hardy RW, Smith JL. Characterization of purified Sindbis virus nsP4 RNA-dependent RNA polymerase activity in vitro. *Virology.* 2009;384(1):201-8.
184. Shirako Y, Strauss JH. Requirement for an aromatic amino acid or histidine at the N terminus of Sindbis virus RNA polymerase. *J Virol.* 1998;72(3):2310-5.
185. Rana J. Molecular interactions of chikungunya virus non-structural proteins. 2014.
186. Sharma R, Kesari P, Kumar P, Tomar S. Structure-function insights into chikungunya virus capsid protein: Small molecules targeting capsid hydrophobic pocket. *Virology.* 2018;515:223-34.
187. Voss JE, Vaney M-C, Duquerroy S, Vornrhein C, Girard-Blanc C, Crublet E, et al. Glycoprotein organization of Chikungunya virus particles revealed by X-ray crystallography. *Nature.* 2010;468(7324):709-12.
188. Kril V, Aïqui-Reboul-Paviet O, Briant L, Amara A. New Insights into Chikungunya Virus Infection and Pathogenesis. *Annual Review of Virology.* 2021;8(1):327-47.
189. Choi H-K, Tong L, Minor W, Dumas P, Boege U, Rossmann MG, et al. Structure of Sindbis virus core protein reveals a chymotrypsin-like serine proteinase and the

organization of the virion. *Nature*. 1991;354(6348):37-43.

190. Garoff H, Frischauf A, Simons K, Lehrach H, Delius H. The capsid protein of Semliki Forest virus has clusters of basic amino acids and prolines in its amino-terminal region. *Proceedings of the National Academy of Sciences*. 1980;77(11):6376-80.

191. Thomas S, Rai J, John L, Schaefer S, Pützer BM, Herchenröder O. Chikungunya virus capsid protein contains nuclear import and export signals. *Viol J*. 2013;10(1):1-13.

192. Perera R, Owen KE, Tellinghuisen TL, Gorbalenya AE, Kuhn RJ. Alphavirus nucleocapsid protein contains a putative coiled coil α -helix important for core assembly. *J Virol*. 2001;75(1):1-10.

193. Hong EM, Perera R, Kuhn RJ. Alphavirus capsid protein helix I controls a checkpoint in nucleocapsid core assembly. *J Virol*. 2006;80(18):8848-55.

194. Aggarwal M, Sharma R, Kumar P, Parida M, Tomar S. Kinetic characterization of trans-proteolytic activity of Chikungunya virus capsid protease and development of a FRET-based HTS assay. *Scientific Reports*. 2015;5(1):1-12.

195. Hahn CS, Strauss JH. Site-directed mutagenesis of the proposed catalytic amino acids of the Sindbis virus capsid protein autoprotease. *J Virol*. 1990;64(6):3069-73.

196. Sariola M, Saraste J, Kuismanen E. Communication of post-Golgi elements with early endocytic pathway: regulation of endoproteolytic cleavage of Semliki Forest virus p62 precursor. *Journal of Cell Science*. 1995;108(6):2465-75.

197. Simizu B, Yamamoto K, Hashimoto K, Ogata T. Structural proteins of Chikungunya virus. *J Virol*. 1984;51(1):254-8.

198. Metz SW, Geertsema C, Martina BE, Andrade P, Heldens JG, van Oers MM, et al. Functional processing and secretion of Chikungunya virus E1 and E2 glycoproteins in insect cells. *Virology*. 2011;8(1):1-12.
199. Uchime O, Fields W, Kielian M. The role of E3 in pH protection during alphavirus assembly and exit. *J Virol*. 2013;87(18):10255-62.
200. Schmidt M, Bracha M, Schlesinger MJ. Evidence for covalent attachment of fatty acids to Sindbis virus glycoproteins. *Proceedings of the National Academy of Sciences*. 1979;76(4):1687-91.
201. Snyder AJ, Sokoloski KJ, Mukhopadhyay S. Mutating conserved cysteines in the alphavirus e2 glycoprotein causes virus-specific assembly defects. *J Virol*. 2012;86(6):3100-11.
202. Kail M, Hollinshead M, Ansorge W, Pepperkok R, Frank R, Griffiths G, et al. The cytoplasmic domain of alphavirus E2 glycoprotein contains a short linear recognition signal required for viral budding. *The EMBO Journal*. 1991;10(9):2343-51.
203. Lescar J, Roussel A, Wien MW, Navaza J, Fuller SD, Wengler G, et al. The fusion glycoprotein shell of Semliki Forest virus: an icosahedral assembly primed for fusogenic activation at endosomal pH. *Cell*. 2001;105(1):137-48.
204. Pletnev SV, Zhang W, Mukhopadhyay S, Fisher BR, Hernandez R, Brown DT, et al. Locations of carbohydrate sites on alphavirus glycoproteins show that E1 forms an icosahedral scaffold. *Cell*. 2001;105(1):127-36.
205. Calvert AE, Bennett SL, Hunt AR, Fong RH, Doranz BJ, Roehrig JT, et al. Exposing cryptic epitopes on the Venezuelan equine encephalitis virus E1 glycoprotein prior to treatment with alphavirus cross-reactive monoclonal antibody allows blockage of replication early in infection. *Virology*. 2022;565:13-21.

206. Williamson LE, Reeder KM, Bailey K, Tran MH, Roy V, Fouch ME, et al. Therapeutic alphavirus cross-reactive E1 human antibodies inhibit viral egress. *Cell*. 2021;184(17):4430-46. e22.
207. Selvarajah S, Sexton NR, Kahle KM, Fong RH, Mattia K-A, Gardner J, et al. A neutralizing monoclonal antibody targeting the acid-sensitive region in chikungunya virus E2 protects from disease. *PLoS neglected tropical diseases*. 2013;7(9):e2423.
208. Okeoma CM. *Chikungunya Virus*. Cham: Springer International Publishing. 2016.
209. Melton JV, Ewart GD, Weir RC, Board PG, Lee E, Gage PW. Alphavirus 6K proteins form ion channels. *Journal of Biological Chemistry*. 2002;277(49):46923-31.
210. Gaedigk-Nitschko K, Ding M, Levy MA, Schlesinger MJ. Site-directed mutations in the Sindbis virus 6K protein reveal sites for fatty acylation and the underacylated protein affects virus release and virion structure. *Virology*. 1990;175(1):282-91.
211. McInerney GM, Smit JM, Liljeström P, Wilschut J. Semliki Forest virus produced in the absence of the 6K protein has an altered spike structure as revealed by decreased membrane fusion capacity. *Virology*. 2004;325(2):200-6.
212. Firth AE, Chung BY, Fleeton MN, Atkins JF. Discovery of frameshifting in Alphavirus 6K resolves a 20-year enigma. *Virol J*. 2008;5(1):1-19.
213. Chung BY-W, Firth AE, Atkins JF. Frameshifting in alphaviruses: a diversity of 3' stimulatory structures. *Journal of molecular biology*. 2010;397(2):448-56.
214. Kam Y-W, Ong EK, Rénia L, Tong J-C, Ng LF. Immuno-biology of Chikungunya and implications for disease intervention. *Microbes and infection*. 2009;11(14-15):1186-96.
215. Sourisseau M, Schilte C, Casartelli N, Trouillet C, Guivel-Benhassine F, Rudnicka

D, et al. Characterization of reemerging chikungunya virus. *PLoS pathogens*. 2007;3(6):e89.

216. Klimstra WB, Nangle EM, Smith MS, Yurochko AD, Ryman KD. DC-SIGN and L-SIGN can act as attachment receptors for alphaviruses and distinguish between mosquito cell- and mammalian cell-derived viruses. *J Virol*. 2003;77(22):12022-32.

217. Rose PP, Hanna SL, Spiridigliozzi A, Wannissorn N, Beiting DP, Ross SR, et al. Natural resistance-associated macrophage protein is a cellular receptor for Sindbis virus in both insect and mammalian hosts. *Cell host & microbe*. 2011;10(2):97-104.

218. Wang K-S, Kuhn RJ, Strauss EG, Ou S, Strauss JH. High-affinity laminin receptor is a receptor for Sindbis virus in mammalian cells. *J Virol*. 1992;66(8):4992-5001.

219. Khan AH, Morita K, del Carmen Parquet M, Hasebe F, Mathenge EG, Igarashi A. Complete nucleotide sequence of chikungunya virus and evidence for an internal polyadenylation site. *Journal of General Virology*. 2002;83(12):3075-84.

220. van Duijl-Richter MK, Hoornweg TE, Rodenhuis-Zybert IA, Smit JM. Early events in chikungunya virus infection—from virus cell binding to membrane fusion. *Viruses*. 2015;7(7):3647-74.

221. Hoornweg TE, van Duijl-Richter MK, Ayala Nuñez NV, Albulescu IC, van Hemert MJ, Smit JM. Dynamics of chikungunya virus cell entry unraveled by single-virus tracking in living cells. *J Virol*. 2016;90(9):4745-56.

222. Vancini R, Wang G, Ferreira D, Hernandez R, Brown DT. Alphavirus genome delivery occurs directly at the plasma membrane in a time- and temperature-dependent process. *J Virol*. 2013;87(8):4352-9.

223. Singh I, Helenius A. Role of ribosomes in Semliki Forest virus nucleocapsid

uncoating. *J Virol.* 1992;66(12):7049-58.

224. Wengler G, Wengler G. Identification of a transfer of viral core protein to cellular ribosomes during the early stages of alphavirus infection. *Virology.* 1984;134(2):435-42.

225. Sokoloski KJ, Nease LM, May NA, Gebhart NN, Jones CE, Morrison TE, et al. Identification of interactions between Sindbis virus capsid protein and cytoplasmic vRNA as novel virulence determinants. *PLoS pathogens.* 2017;13(6):e1006473.

226. Li G, Rice CM. The signal for translational readthrough of a UGA codon in Sindbis virus RNA involves a single cytidine residue immediately downstream of the termination codon. *J Virol.* 1993;67(8):5062-7.

227. Abraham R, Hauer D, McPherson RL, Utt A, Kirby IT, Cohen MS, et al. ADP-ribosyl-binding and hydrolase activities of the alphavirus nsP3 macrodomain are critical for initiation of virus replication. *Proceedings of the National Academy of Sciences.* 2018;115(44):E10457-E66.

228. Peränen J, Laakkonen P, Hyvönen M, Kääriäinen L. The alphavirus replicase protein nsP1 is membrane-associated and has affinity to endocytic organelles. *Virology.* 1995;208(2):610-20.

229. Salonen A, Vasiljeva L, Merits A, Magden J, Jokitalo E, Kääriäinen L. Properly folded nonstructural polyprotein directs the Semliki Forest virus replication complex to the endosomal compartment. *J Virol.* 2003;77(3):1691-702.

230. Kallio K, Hellström K, Jokitalo E, Ahola T. RNA replication and membrane modification require the same functions of alphavirus nonstructural proteins. *J Virol.* 2016;90(3):1687-92.

231. Jose J, Taylor AB, Kuhn RJ. Spatial and temporal analysis of alphavirus replication and assembly in mammalian and mosquito cells. *MBio*. 2017;8(1):e02294-16.
232. Hellström K, Kallio K, Utt A, Quirin T, Jokitalo E, Merits A, et al. Partially uncleaved alphavirus replicase forms spherule structures in the presence and absence of RNA template. *J Virol*. 2017;91(18):e00787-17.
233. Frolova EI, Gorchakov R, Pereboeva L, Atasheva S, Frolov I. Functional Sindbis virus replicative complexes are formed at the plasma membrane. *J Virol*. 2010;84(22):11679-95.
234. Spuul P, Balistreri G, Kääriäinen L, Ahola T. Phosphatidylinositol 3-kinase-, actin-, and microtubule-dependent transport of Semliki Forest Virus replication complexes from the plasma membrane to modified lysosomes. *J Virol*. 2010;84(15):7543-57.
235. Carrasco L, Sanz MA, Gonzalez-Almela E. The Regulation of Translation in Alphavirus-Infected Cells. *Viruses-Basel*. 2018;10(2).
236. Shirako Y, Strauss JH. 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(3):1874-85.
237. Shirako Y, Strauss JH. Cleavage between nsP1 and nsP2 initiates the processing pathway of Sindbis virus nonstructural polyprotein P123. *Virology*. 1990;177(1):54-64.
238. Sawicki DL, Sawicki SG. Short-lived minus-strand polymerase for Semliki Forest virus. *J Virol*. 1980;34(1):108-18.
239. Sawicki DL, Sawicki SG. A second nonstructural protein functions in the

- regulation of alphavirus negative-strand RNA synthesis. *J Virol.* 1993;67(6):3605-10.
240. Suopanki J, Sawicki DL, Sawicki SG, Kääriäinen L. Regulation of alphavirus 26S mRNA transcription by replicase component nsP2. *Journal of general virology.* 1998;79(2):309-19.
241. Scholte FE, Tas A, Albulescu IC, Žusinaite E, Merits A, Snijder EJ, et al. Stress granule components G3BP1 and G3BP2 play a proviral role early in Chikungunya virus replication. *J Virol.* 2015;89(8):4457-69.
242. Sanz M, Madan V, Nieva J, Carrasco L. The alphavirus 6K protein. *Viral membrane proteins: structure, function, and drug design: Springer; 2005.* p. 233-44.
243. Sanz MA, Carrasco L. Sindbis virus variant with a deletion in the 6K gene shows defects in glycoprotein processing and trafficking: lack of complementation by a wild-type 6K gene in trans. *J Virol.* 2001;75(16):7778-84.
244. Snyder JE, Kulcsar KA, Schultz KL, Riley CP, Neary JT, Marr S, et al. Functional characterization of the alphavirus TF protein. *J Virol.* 2013;87(15):8511-23.
245. Jose J, Snyder JE, Kuhn RJ. A structural and functional perspective of alphavirus replication and assembly. *Future microbiology.* 2009;4(7):837-56.
246. Lee CY, Kam Y-W, Fric J, Malleret B, Koh EG, Prakash C, et al. Chikungunya virus neutralization antigens and direct cell-to-cell transmission are revealed by human antibody-escape mutants. *PLoS pathogens.* 2011;7(12):e1002390.
247. Lum F-M, Ng LF. Cellular and molecular mechanisms of chikungunya pathogenesis. *Antiviral research.* 2015;120:165-74.
248. Barton DJ, O'Donnell BJ, Flanagan JB. 5' cloverleaf in poliovirus RNA is a cis-

acting replication element required for negative-strand synthesis. *The EMBO journal*. 2001;20(6):1439-48.

249. Li X-F, Jiang T, Yu X-D, Deng Y-Q, Zhao H, Zhu Q-Y, et al. RNA elements within the 5' untranslated region of the West Nile virus genome are critical for RNA synthesis and virus replication. *Journal of General Virology*. 2010;91(5):1218-23.

250. Filomatori CV, Lodeiro MF, Alvarez DE, Samsa MM, Pietrasanta L, Gamarnik AV. A 5' RNA element promotes dengue virus RNA synthesis on a circular genome. *Genes & development*. 2006;20(16):2238-49.

251. Herold J, Andino R. Poliovirus RNA replication requires genome circularization through a protein-protein bridge. *Molecular cell*. 2001;7(3):581-91.

252. Friebe P, Harris E. Interplay of RNA elements in the dengue virus 5' and 3' ends required for viral RNA replication. *J Virol*. 2010;84(12):6103-18.

253. Friebe P, Shi P-Y, Harris E. The 5' and 3' downstream AUG region elements are required for mosquito-borne flavivirus RNA replication. *J Virol*. 2011;85(4):1900-5.

254. Sakakibara S, Nakamura Y, Satoh H, Okano H. Rna-binding protein Musashi2: developmentally regulated expression in neural precursor cells and subpopulations of neurons in mammalian CNS. *J Neurosci*. 2001;21(20):8091-107.

255. Sakakibara S-i, Nakamura Y, Yoshida T, Shibata S, Koike M, Takano H, et al. RNA-binding protein Musashi family: roles for CNS stem cells and a subpopulation of ependymal cells revealed by targeted disruption and antisense ablation. *Proceedings of the National Academy of Sciences*. 2002;99(23):15194-9.

256. Caldas-Garcia GB, Schüler-Faccini L, Pissinatti A, Paixão-Côrtés VR, Bortolini MC. Evolutionary analysis of the Musashi family: What can it tell us about Zika? *Infection*,

Genetics and Evolution. 2020;84:104364.

257. Zearfoss NR, Deveau LM, Clingman CC, Schmidt E, Johnson ES, Massi F, et al. A conserved three-nucleotide core motif defines Musashi RNA binding specificity. *J Biol Chem*. 2014;289(51):35530-41.

258. Duggimpudi S, Kloetgen A, Maney SK, Münch PC, Hezaveh K, Shaykhalishahi H, et al. Transcriptome-wide analysis uncovers the targets of the RNA-binding protein MSI2 and effects of MSI2's RNA-binding activity on IL-6 signaling. *Journal of Biological Chemistry*. 2018;293(40):15359-69.

259. Gerstberger S, Hafner M, Tuschl T. A census of human RNA-binding proteins. *Nature Reviews Genetics*. 2014;15(12):829-45.

260. Choi Y, Kim K, Lee J, Chun Y, An I, An S, et al. DBC2/RhoBTB2 functions as a tumor suppressor protein via Musashi-2 ubiquitination in breast cancer. *Oncogene*. 2017;36(20):2802-12.

261. Kudinov AE, Deneka A, Nikonova AS, Beck TN, Ahn Y-H, Liu X, et al. Musashi-2 (MSI2) supports TGF- β signaling and inhibits claudins to promote non-small cell lung cancer (NSCLC) metastasis. *Proceedings of the National Academy of Sciences*. 2016;113(25):6955-60.

262. UniProt: the universal protein knowledgebase in 2021. *Nucleic acids research*. 2021;49(D1):D480-D9.

263. Chavali PL, Stojic L, Meredith LW, Joseph N, Nahorski MS, Sanford TJ, et al. Neurodevelopmental protein Musashi-1 interacts with the Zika genome and promotes viral replication. *Science*. 2017;357(6346):83-8.

264. Schneider AB, Wolfinger MT. Musashi binding elements in Zika and related

Flavivirus 3'UTRs: A comparative study in silico. *Sci Rep.* 2019;9(1):6911.

265. Pohjala L, Utt A, Varjak M, Lulla A, Merits A, Ahola T, et al. Inhibitors of alphavirus entry and replication identified with a stable Chikungunya replicon cell line and virus-based assays. *PloS one.* 2011;6(12):e28923.

266. Clingman CC, Deveau LM, Hay SA, Genga RM, Shandilya SM, Massi F, et al. Allosteric inhibition of a stem cell RNA-binding protein by an intermediary metabolite. *Elife.* 2014;3:e02848.

267. Plaskon NE, Adelman ZN, Myles KM. Accurate strand-specific quantification of viral RNA. *PloS one.* 2009;4(10):e7468.

268. Scutigliani EM, Kikkert M. Interaction of the innate immune system with positive-strand RNA virus replication organelles. *Cytokine & growth factor reviews.* 2017;37:17-27.

269. Utt A, Quirin T, Saul S, Hellström K, Ahola T, Merits A. Versatile trans-replication systems for chikungunya virus allow functional analysis and tagging of every replicase protein. *PLoS One.* 2016;11(3):e0151616.

270. Torres M, Becquet D, Guillen S, Boyer B, Moreno M, Blanchard M-P, et al. RNA pull-down procedure to identify RNA targets of a long non-coding RNA. *JoVE (Journal of Visualized Experiments).* 2018(134):e57379.

271. Arkin MR, Wells JA. Small-molecule inhibitors of protein–protein interactions: progressing towards the dream. *Nature reviews Drug discovery.* 2004;3(4):301-17.

272. Niederhauser O, Mangold M, Schubanel R, Kuszniir E, Schmidt D, Hertel C. NGF ligand alters NGF signaling via p75NTR and TrkA. *Journal of neuroscience research.* 2000;61(3):263-72.

273. Minuesa G, Albanese SK, Xie W, Kazansky Y, Worroll D, Chow A, et al. Small-molecule targeting of MUSASHI RNA-binding activity in acute myeloid leukemia. *Nature communications*. 2019;10(1):1-15.
274. Roberts GC, Zothner C, Remenyi R, Merits A, Stonehouse NJ, Harris M. Evaluation of a range of mammalian and mosquito cell lines for use in Chikungunya virus research. *Scientific reports*. 2017;7(1):1-13.
275. Das I, Basantray I, Mamidi P, Nayak TK, BM P, Chattopadhyay S, et al. Heat shock protein 90 positively regulates Chikungunya virus replication by stabilizing viral non-structural protein nsP2 during infection. *PloS one*. 2014;9(6):e100531.
276. Minuesa G, Antczak C, Shum D, Radu C, Bhinder B, Li Y, et al. A 1536-well fluorescence polarization assay to screen for modulators of the MUSASHI family of RNA-binding proteins. *Combinatorial chemistry & high throughput screening*. 2014;17(7):596-609.
277. Lin A, Giuliano CJ, Palladino A, John KM, Abramowicz C, Yuan ML, et al. Off-target toxicity is a common mechanism of action of cancer drugs undergoing clinical trials. *Science translational medicine*. 2019;11(509):eaaw8412.
278. Steegmann JL, Cervantes F, le Coutre P, Porkka K, Saglio G. Off-target effects of BCR–ABL1 inhibitors and their potential long-term implications in patients with chronic myeloid leukemia. *Leukemia & lymphoma*. 2012;53(12):2351-61.
279. Stepanenko A, Dmitrenko V. Pitfalls of the MTT assay: Direct and off-target effects of inhibitors can result in over/underestimation of cell viability. *Gene*. 2015;574(2):193-203.
280. Zheng X, Cho S, Moon H, Loh TJ, Jang HN, Shen H. Detecting RNA–protein interaction using end-labeled biotinylated RNA oligonucleotides and immunoblotting.

RNA-Protein Complexes and Interactions: Springer; 2016. p. 35-44.

281. Qoronfleh MW, Ren L, Emery D, Perr M, Kaboord B. Use of immunomatrix methods to improve protein-protein interaction detection. *Journal of Biomedicine and Biotechnology*. 2003;2003:291-8.

282. Windisch MP, Frese M, Kaul A, Trippler M, Lohmann V, Bartenschlager R. Dissecting the interferon-induced inhibition of hepatitis C virus replication by using a novel host cell line. *J Virol*. 2005;79(21):13778-93.

283. Horscroft N, Bellows D, Ansari I, Lai VC, Dempsey S, Liang D, et al. Establishment of a subgenomic replicon for bovine viral diarrhoea virus in Huh-7 cells and modulation of interferon-regulated factor 3-mediated antiviral response. *J Virol*. 2005;79(5):2788-96.

284. Stuyver LJ, McBrayer TR, Tharnish PM, Hassan AE, Chu CK, Pankiewicz KW, et al. Dynamics of subgenomic hepatitis C virus replicon RNA levels in Huh-7 cells after exposure to nucleoside antimetabolites. *J Virol*. 2003;77(19):10689-94.

285. Zhu H, Dong H, Eksioglu E, Hemming A, Cao M, Crawford JM, et al. Hepatitis C virus triggers apoptosis of a newly developed hepatoma cell line through antiviral defense system. *Gastroenterology*. 2007;133(5):1649-59.

286. Zamore PD, Haley B. Ribo-gnome: the big world of small RNAs. *science*. 2005;309(5740):1519-24.

287. Lam JK, Chow MY, Zhang Y, Leung SW. siRNA versus miRNA as therapeutics for gene silencing. *Molecular Therapy-Nucleic Acids*. 2015;4:e252.

288. Zhang H, Kolb FA, Jaskiewicz L, Westhof E, Filipowicz W. Single processing center models for human Dicer and bacterial RNase III. *Cell*. 2004;118(1):57-68.

289. Lee Y, Hur I, Park SY, Kim YK, Suh MR, Kim VN. The role of PACT in the RNA silencing pathway. *The EMBO journal*. 2006;25(3):522-32.
290. Tang G. siRNA and miRNA: an insight into RISCs. *Trends in biochemical sciences*. 2005;30(2):106-14.
291. Czauderna F, Fechtner M, Dames S, AyguÈn H, Klippel A, Pronk GJ, et al. Structural variations and stabilising modifications of synthetic siRNAs in mammalian cells. *Nucleic acids research*. 2003;31(11):2705-16.
292. Wang T, Larcher LM, Ma L, Veedu RN. Systematic screening of commonly used commercial transfection reagents towards efficient transfection of single-stranded oligonucleotides. *Molecules*. 2018;23(10):2564.
293. Rao DD, Vorhies JS, Senzer N, Nemunaitis J. siRNA vs. shRNA: similarities and differences. *Advanced drug delivery reviews*. 2009;61(9):746-59.
294. Taxman DJ, Moore CB, Guthrie EH, Huang MT-H. Short hairpin RNA (shRNA): design, delivery, and assessment of gene knockdown. *RNA therapeutics*: Springer; 2010. p. 139-56.
295. Imai T, Tokunaga A, Yoshida T, Hashimoto M, Mikoshiba K, Weinmaster G, et al. The neural RNA-binding protein Musashi1 translationally regulates mammalian numb gene expression by interacting with its mRNA. *Molecular and cellular biology*. 2001;21(12):3888-900.
296. Chua H, Abdul Rashid K, Law C, Hamizah A, Khairul A, Chua K. A fatal case of chikungunya virus infection with liver involvement. *Medical Journal of Malaysia*. 2010;65(1):83-4.
297. Connolly J. Lentiviruses in gene therapy clinical research. *Gene therapy*.

2002;9(24):1730-4.

298. Devhare PB, Desai S, Lole KS. Innate immune responses in human hepatocyte-derived cell lines alter genotype 1 hepatitis E virus replication efficiencies. *Scientific reports*. 2016;6(1):1-10.

299. Van Dijk AA, Makeyev EV, Bamford DH. Initiation of viral RNA-dependent RNA polymerization. *Journal of general virology*. 2004;85(5):1077-93.

300. Khromykh AA, Meka H, Guyatt KJ, Westaway EG. Essential role of cyclization sequences in flavivirus RNA replication. *J Virol*. 2001;75(14):6719-28.

301. You S, Falgout B, Markoff L, Padmanabhan R. In vitro RNA synthesis from exogenous dengue viral RNA templates requires long range interactions between 5'- and 3'-terminal regions that influence RNA structure. *Journal of Biological Chemistry*. 2001;276(19):15581-91.

302. Liu Z-Y, Li X-F, Jiang T, Deng Y-Q, Ye Q, Zhao H, et al. Viral RNA switch mediates the dynamic control of flavivirus replicase recruitment by genome cyclization. *Elife*. 2016;5:e17636.

303. Sanford TJ. Mechanistic analysis of the Zika virus translation-replication switch: University of Cambridge; 2020.

304. Gui H, Lu C-W, Adams S, Stollar V, Li M-L. hnRNP A1 interacts with the genomic and subgenomic RNA promoters of Sindbis virus and is required for the synthesis of G and SG RNA. *Journal of biomedical science*. 2010;17(1):1-7.

305. Kumar R, Chander Y, Khandelwal N, Nagori H, Verma A, Pal Y, et al. hnRNPA1 regulates early translation to replication switch in SARS-CoV-2 life cycle. *bioRxiv*. 2021.

306. Bhattacharya T, Yan L, Crawford JM, Zaher H, Newton IL, Hardy RW. Differential viral RNA methylation contributes to pathogen blocking in Wolbachia-colonized arthropods. *PLoS pathogens*. 2022;18(3):e1010393.
307. Hirota Y, Okabe M, Imai T, Kurusu M, Yamamoto A, Miyao S, et al. Musashi and seven in absentia downregulate Tramtrack through distinct mechanisms in *Drosophila* eye development. *Mechanisms of development*. 1999;87(1-2):93-101.
308. Sharma A, Akagi K, Pattavina B, Wilson KA, Nelson C, Watson M, et al. Musashi expression in intestinal stem cells attenuates radiation-induced decline in intestinal permeability and survival in *Drosophila*. *Scientific reports*. 2020;10(1):1-16.
309. Carey MF, Peterson CL, Smale ST. Experimental strategies for the identification of DNA-binding proteins. *Cold Spring Harbor Protocols*. 2012;2012(1):pdb.top067470.
310. Song C, Zhang S, Huang H. Choosing a suitable method for the identification of replication origins in microbial genomes. *Frontiers in microbiology*. 2015;6:1049.
311. Shirako Y, Strauss EG, Strauss JH. Modification of the 5' terminus of Sindbis virus genomic RNA allows nsP4 RNA polymerases with nonaromatic amino acids at the N terminus to function in RNA replication. *J Virol*. 2003;77(4):2301-9.
312. Gorchakov R, Hardy R, Rice CM, Frolov I. Selection of functional 5' cis-acting elements promoting efficient Sindbis virus genome replication. *J Virol*. 2004;78(1):61-75.
313. Hussian CHAC, Rahman RNZRA, Chor ALT, Salleh AB, Ali MSM. Enhancement of a protocol purifying T1 lipase through molecular approach. *PeerJ*. 2018;6:e5833.
314. BERBEN-BLOEMHEUVEL G, KASPERAITIS MA, van HEUGTEN H, THOMAS AA, van STEEG H, VOORMA HO. Interaction of initiation factors with the cap structure of

chimaeric mRNA containing the 5'-untranslated regions of Semliki Forest virus RNA is related to translational efficiency. *European journal of biochemistry*. 1992;208(3):581-7.

315. Castelló A, Sanz MÁ, Molina S, Carrasco L. Translation of Sindbis Virus 26 S mRNA Does Not Require Intact Eukariotic Initiation Factor 4G. *Journal of molecular biology*. 2006;355(5):942-56.

316. Hyde JL, Gardner CL, Kimura T, White JP, Liu G, Trobaugh DW, et al. A viral RNA structural element alters host recognition of nonself RNA. *Science*. 2014;343(6172):783-7.

317. White LJ, Wang J-G, Davis NL, Johnston RE. Role of alpha/beta interferon in Venezuelan equine encephalitis virus pathogenesis: effect of an attenuating mutation in the 5' untranslated region. *J Virol*. 2001;75(8):3706-18.

318. Bouraï M, Lucas-Hourani M, Gad HH, Drosten C, Jacob Y, Tafforeau L, et al. Mapping of Chikungunya virus interactions with host proteins identified nsP2 as a highly connected viral component. *J Virol*. 2012;86(6):3121-34.

319. Liang Y, Hong Y, Parslow TG. cis-Acting packaging signals in the influenza virus PB1, PB2, and PA genomic RNA segments. *J Virol*. 2005;79(16):10348-55.

320. Chaitanya K. Structure and organization of virus genomes. *Genome and Genomics: Springer*; 2019. p. 1-30.

321. Fried MG. Measurement of protein-DNA interaction parameters by electrophoresis mobility shift assay. *Electrophoresis*. 1989;10(5-6):366-76.

322. Tuplin A, Struthers M, Cook J, Bentley K, Evans DJ. Inhibition of HCV translation by disrupting the structure and interactions of the viral CRE and 3' X-tail. *Nucleic acids*

research. 2015;43(5):2914-26.

323. Villordo SM, Alvarez DE, Gamarnik AV. A balance between circular and linear forms of the dengue virus genome is crucial for viral replication. *Rna*. 2010;16(12):2325-35.

324. Kulasegaran-Shylini R, Thiviyanathan V, Gorenstein DG, Frolov I. The 5' UTR-specific mutation in VEEV TC-83 genome has a strong effect on RNA replication and subgenomic RNA synthesis, but not on translation of the encoded proteins. *Virology*. 2009;387(1):211-21.

325. Diviney S, Tuplin A, Struthers M, Armstrong V, Elliott RM, Simmonds P, et al. A hepatitis C virus cis-acting replication element forms a long-range RNA-RNA interaction with upstream RNA sequences in NS5B. *J Virol*. 2008;82(18):9008-22.

326. Manzano M, Reichert ED, Polo S, Falgout B, Kasprzak W, Shapiro BA, et al. Identification of cis-acting elements in the 3'-untranslated region of the dengue virus type 2 RNA that modulate translation and replication. *Journal of Biological Chemistry*. 2011;286(25):22521-34.

327. Li X-D, Deng C-L, Yuan Z-M, Ye H-Q, Zhang B. Different degrees of 5'-to-3'DAR interactions modulate Zika virus genome cyclization and host-specific replication. *J Virol*. 2020;94(5):e01602-19.

328. Janecki DM, Swiatkowska A, Szpotkowska J, Urbanowicz A, Kabacińska M, Szpotkowski K, et al. Poly (C)-binding protein 2 regulates the p53 expression via interactions with the 5'-Terminal region of p53 mRNA. *International Journal of Molecular Sciences*. 2021;22(24):13306.

329. Boersma S, Rabouw HH, Bruurs LJ, Pavlovič T, van Vliet AL, Beumer J, et al. Translation and replication dynamics of single RNA viruses. *Cell*. 2020;183(7):1930-45.

e23.

330. Madden EA, Plante KS, Morrison CR, Kutchko KM, Sanders W, Long KM, et al. Using SHAPE-MaP to model RNA secondary structure and identify 3' UTR variation in chikungunya virus. *J Virol.* 2020;94(24):e00701-20.

331. Su H, Lei C-T, Zhang C. Interleukin-6 signaling pathway and its role in kidney disease: an update. *Frontiers in immunology.* 2017;8:405.

332. Hyde JL, Chen R, Trobaugh DW, Diamond MS, Weaver SC, Klimstra WB, et al. The 5' and 3' ends of alphavirus RNAs—Non-coding is not non-functional. *Virus research.* 2015;206:99-107.

333. Kuhn RJ, Griffin DE, Zhang H, Niesters H, Strauss JH. Attenuation of Sindbis virus neurovirulence by using defined mutations in nontranslated regions of the genome RNA. *J Virol.* 1992;66(12):7121-7.