Targeting functional RNA structures within the genomes of Chikungunya virus and Zika virus

Oliver Daniel Prosser

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

The University of Leeds

Faculty of Biological Sciences

School of Molecular and Cellular Biology

March 2021

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.

Acknowledgements

I would first like to thank Dr Andy Tuplin. I could not have asked for a better supervisor, who both encouraged me when I was doing well and provided invaluable support and advice during more difficult times.

Thank you to Professor Nicola Stonehouse for all of your help and advice throughout this project, again I could not have asked for a better second supervisor.

To Marietta, thank for providing me with constant support and advice and always being there to help.

I would like to thank Carsten for being a fantastic lab manger and keeping things in order.

Raymond and Christian for being my brothers in arms!

Past and present members of the Tuplin group and Garstang 8.61, thank you for being great colleagues and fellow combatants against the lab gremlins.

Finally, I would like to thank my parents for being great and always believing in me.

Abstract

Chikungunya virus (CHIKV) and Zika virus (ZIKV) share a common feature, in that the genomes of both viruses consists of a single strand of positive sense RNA, which folds into complex secondary structures. These structures are highly conserved and play essential roles in virus replication. The Tuplin group previously characterised RNA structures within the 5' region of the CHIKV genome, demonstrating that such RNA structures function in a structure dependent manner. We hypothesised that these functional RNA structures could be specifically targeted using antisense oligonucleotides, binding in a sequence specific manner and disrupting key RNA interactions required for efficient virus replication.

A panel of locked nucleic acid (LNA) oligonucleotides was shown to bind specific RNA structures in the context of a folded CHIKV RNA molecule. Using a range of methods, LNA oligonucleotides were shown to inhibit virus replication, indicating that RNA elements are accessible within the CHIKV RNA replication complex. Furthermore, LNA oligonucleotides were able to inhibit a CHIKV sub-genomic replicon system, indicating that the oligonucleotides specifically inhibit genome replication.

RNA structural elements within the 5' UTR and adjacent capsid coding region of ZIKV were modelled using a combination of SHAPE chemistry, thermodynamic predictions and phylogenetic analysis. ZIKV RNA structure models were generated using both full-length genomic RNA and truncated RNA molecules at both 37 °C and 28 °C, reflecting both mammalian and insect body temperatures.

The RNA structural models were used as a basis for reverse genetic analysis of ZIKV RNA structures and a novel pseudoknot interaction was shown to be essential for efficient ZIKV replication. RNA structures within the ZIKV genome were specifically targeted using antisense oligonucleotides, causing impairment of ZIKV replication in both infectious ZIKV and sub-genomic replicon systems. Together, these results provide a basis for the development specific antiviral agents targeting CHIKV and ZIKV as no specific agents currently exist.

Table of Contents

Table of Figures

Table of Tables

Abbreviations list

ADP - Adenosine diphosphate

Akt - Protein kinase B

AMP - Ampicillin

APS - Ammonium persulphate

ATP – Adenosine triphosphate

AUD - Alphavirus unique domain

AUG – Methionine codon (start codon)

BCA - Bicinchoninic acid

BHK – Baby hamster kidney

bp – Base pair

C protein – Capsid protein (CHIKV and ZIKV)

cDNA – Complimentary DNA

CD8 – Cluster of differentiation 8

CHIKV - Chikungunya virus

CHIKV ICRES - Chikungunya virus infectious clone

CHIKV 1-337 – Chikungunya nucleotides 1-337

cHP - Capsid-coding region hairpin element

CSFV - Classical swine fever virus

CME - Clathrin mediated endocytosis

CPV - Cytopathic vacuole

CR – Conserved region

CSE – Conserved sequence element

CTP – Cytosine triphosphate

CS - Cyclisation sequences

C6/36 - *Aedes albopictus* larval cells

ddATP – Dideoxyadenosine triphosphate

ddH_2O – Double distilled water

DENV - Dengue virus

DEPC - Diethyl pyrocarbonate

DIS - dimerization initiation sites

DMEM - Dulbecco's modified minimal essential medium

DMSO - Dimethyl sulfoxide

DNA – Deoxyribose nucleic acid

DNase - Deoxyribonuclease

dNTP – Deoxynucleotide triphosphate

DTT - Dithiothreitol

E protein – ZIKV envelope protein

ECSA - East/Central/South African (CHIKV strain)

EDTA - Ethylenediaminetetraacetic acid

EILV - Eliat virus

ELISA - Enzyme-linked immunosorbent assay

EMSA - Electromobility shift assays

ER – Endoplasmic reticulum

E1 – CHIKV envelope protein 1

E2 – CHIKV envelope protein 2

E3 – CHIKV envelope protein 3

Fluc-Rluc-(GDD>GAA)-SGR – Chikungunya virus sub genomic replicon translation reporter

Fluc-SGR - Chikungunya virus firefly luciferase sub-genomic replicon

FMDV - Foot and mouth disease virus

F-SLA - Fluorescein-labelled SLA

FXR - Farnesoid X receptor

GAG – Glycosaminoglycan

GBV-A – G. Barker virus A

GBV-C – G. Barker virus C

GFP - Green fluorescent protein

GTase - Guanylyltransferase

GTP – Guanosine triphosphate

GBS - Guillain-Barré syndrome

G3BP - Ras GTPase-activating protein-binding protein

HCV – Hepatitis C virus

HEK 293T – Human embryonic kidney cells, expressing a mutant version of the simian virus 40 large T antigen

HEPES - (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)

HIV - Human immunodeficiency virus

HIV TAR - Human immunodeficiency virus trans-activation response element

Huh 7 - Human hepatocellular carcinoma

HVR - Hypervariable region

IFIT - Interferon-induced protein with tetratricopeptide repeats 1

IFN – Interferon

IgG – Immunoglobulin G

IRES - internal ribosome entry site

kDa - Kilodalton

ORF - Open reading frame

LB – Lysogeny broth

LB-AMP - Lysogeny broth containing ampicillin

LC3 - Microtubule-associated protein light chain 3

LNA - Locked nucleic acid oligonucleotides

LoVo cells - Human colon adenocarcinoma cells

M - Molar

M protein – ZIKV membrane protein

miRNA – Micro RNA

MPO - Modified phosphorothioate oligonucleotide

MOI – Multiplicity of infection

MOPS - 3-(N-morpholino)propanesulfonic acid

ml – Millilitres

mM – Millimolar

mRNA – Messenger RNA

ms – Milliseconds

MTase - Methyltransferase

mTOR – Mechanistic target of rapamycin

MTT - Thiazolyl Blue Tetrazolium Bromide

Mxra8 - Matrix Remodelling Associated 8

NaOAc – Sodium acetate

ng – Nanogram

nm – Nanometres

nM – Nanomolar

NMIA - N-methyl-nitroisatoic anhydride

NMR - Nuclear magnetic resonance

NP-40 - nonyl phenoxypolyethoxylethanol

NS1 – ZIKV non-structural protein 1

NS2A – ZIKV non-structural protein 2A

NS2B – ZIKV non-structural protein 2B

NS3 – ZIKV non-structural protein 3

NS4A – ZIKV non-structural protein 4A

NS4B – ZIKV non-structural protein 4B

NS5 – ZIKV non-structural protein 5

nsp1 – CHIKV non-structural protein 1

nsp2 – CHIKV non-structural protein 2

nsp3 – CHIKV non-structural protein 3

nsp4 – CHIKV non-structural protein 4

NTP – Nucleoside triphosphate

nt - Nucleotide

Oligo - Oligonucleotide

ONNV - O'nyong-nyong virus

P ≤ - P value less than or equal to

PABP - Poly(A) binding protein

P_{add} - Probability of adduct formation at each nucleotide

PAGE - Polyacrylamide gel electrophoresis

PBS – Phosphate-buffered saline

PCR – Polymerase chain reaction

PE2 - CHIKV envelope protein 2 and CHIKV envelope protein 3 precursor protein

pH – Potential of hydrogen

PMO - Phosphorodiamidate morpholino oligomer

PK – Pseudoknot

PFU - Plaque forming units

PNK - Polynucleotide kinase

pmoles - picomoles

pre-miR-372 - microRNA-372 precursor

prM protein – ZIKV precursor membrane protein

P1234 – CHIKV non-structural polyprotein NSP1, -2, -3 and -4

P123 - CHIKV non-structural polyprotein NSP1, -2 and -3

qRT-PCR – Quantitative real time polymerase chain reaction

RdRp - RNA dependent RNA polymerase

RLU – Relative light units

RNA – Ribonucleic acid

RNase – Ribonuclease

rRNA – Ribosomal RNA

RRE - RNA replication element

rpm – revolutions per minute

Rsample - Restrained sample

RT-PCR – Reverse transcription polymerase chain reaction

SARS-CoV - Severe acute respiratory syndrome coronavirus

SDS - Sodium dodecyl sulphate

SDS-PAGE - Sodium dodecyl sulphate–polyacrylamide gel electrophoresis

SELEX - Systematic evolution of ligands by exponential enrichment

SFV - Semliki Forest virus

sfRNA - Sub-genomic flavivirus RNAs

SHAPE - Selective 2'-Hydroxyl Acylation Analysed by Primer Extension

sHP - Short hairpin

SINV – Sindbis virus

SL – RNA stem loop

SLA - Stem-loop A

SLB - Stem-loop B

SLD – Stem-loop D

SLE – Stem-loop E

SMN - Survival of motor neuron

STAT - Signal transducer and activator of transcription

TAE buffer – Tris base – acetic acid – EDTA buffer

TATase – Adenylytransferase

TBE buffer - Tris/Borate/EDTA buffer

TBEV - Tick-borne encephalitis virus

TE buffer – Tris base - EDTA buffer

TEMED - N,N,N’,N’- tetramethylethylenediamine

Tm – Melting temperature

T7_Ren - Transfection control *Renilla* RNA

UAR - Upstream of AUG region

UFS - 5'-UAR-flanking stem

UGA – Opal stop codon

UTP – Uracil triphosphate

UTR – Untranslated region

V - Volts

VEEV - Venezuelan Equine Encephalitis virus

Vero – African green monkey kidney epithelial cells

WA – West African (CHIKV strain)

WEEV - Western Equine Encephalitis virus

WNV - West Nile virus

WT – Wild-type

x g – Times gravity

XRNs - Exoribonucleases

YFV - Yellow fever virus

ZIKV – Zika virus

ZIKV IC – ZIKV infectious clone

ZIKV N-luc IC – ZIKV infectious clone, expressing nano luciferase

ZIKV N-luc Replicon – ZIKV replicon expressing nano luciferase

ZIKV 1-772 – ZIKV nucleotides 1-772

ZLP1 – ZIKV LNA panel 1

ZLP2 - ZIKV LNA panel 2

α-^{32}P – Alpha 32 phosphorus-32

γ-^{32}P – Gamma phosphorus-32

µCi - Microcurie

µg – Microgram

µl – Microlitre

µM – Micromolar

6-FAM - 6-Carboxyfluorescein

^{32}P - Phosphorus-32

Chapter 1 Introduction

1.1.1 General introduction

Arboviruses such as Chikungunya virus (CHIKV) and Zika virus (ZIKV) virus represent a significant public health concern. In the near future these pathogens are likely to increase in prevalence, as the disease vectors of their urban infectious cycle (*Aedes aegypti* and *Aedes albopictus*) increase in global distribution, driven by a combination of human movement, change of land use and climate change [1]. At present, there are no licenced antiviral therapies or vaccines available for either pathogen and consequently, development of such agents is a research priority. In order to identify novel therapeutic antiviral targets, it is essential to improve our knowledge of the CHIKV and ZIKV lifecycles at a molecular level, including the role played by RNA replication elements (RREs) in virus replication. RNA structures within the 5' UTR and adjacent nsp1 coding region of CHIKV have been previously shown to be critical for virus replication [2] and therefore represent potential targets for antiviral agents. In addition, RNA structures within the 5' genome region of ZIKV remain understudied, but by analogy to related flaviviruses, are likely to play essential roles in virus replication. Characterisation of these structures will increase our knowledge of factors required for ZIKV replication and potentially elucidate novel targets for antiviral agents.

1.1.2 RNA interactions

Viral RNA genomes can have a high degree of structural complexity, due to the ability of RNA molecules to form multiple interactions, and the overall topology of an RNA molecule is composed of base-paired and single-stranded regions. Base pairing of complementary sequences causes the formation of RNA duplexes, which represent a major fraction of the overall RNA structure. Simple RNA elements are composed of a single base paired duplex stem and a single unpaired apical loop as shown by Fig. 1.1A. More complex structures can be composed of multiple base paired and unpaired regions as shown by Fig. 1.1C. Single stranded regions of nucleotides are found between RNA elements, and in terminal loops, as indicated in Fig. 1.1A. Base pairing mismatches may also result in small single stranded bulges within base paired RNA structures.

In addition to these secondary structural elements, RNA molecules can also form tertiary structures, such as pseudoknot interactions (Fig. 1.1B). Pseudoknot interactions occur when an unpaired region of an RNA element, such as an apical loop, forms a bipartite duplex with a region outside of the RNA element [3,4]. RNA pseudoknots have a number of functions, such as -1 ribosomal frameshifting, whereby the ribosome shifts back one nucleotide in order to transcribe from another reading frame [5]. For example, in order to investigate ribosomal frame shifting in severe acute respiratory syndrome coronavirus (SARS-CoV), a reporter system was used, whereby mutant pseudoknot sequences were cloned upstream of a firefly reporter gene, where the reporter is only expressed if frame shifting occurs. Mutagenesis disrupting the pseudoknot interaction caused a reduction in frame shifting, relative to WT [6]. In addition, pseudoknot interactions can promote stop codon readthrough. For example, in

murine leukaemia virus, readthrough of a stop codon by 5-10% of ribosomes is required for expression of the Gag-Pol fusion protein. Mutagenesis to disrupt base pairs in the pseudoknot resulted in a decrease in stop codon readthrough products compared with truncated products; a finding demonstrated using *in vitro* translation assays were products were visualised using sodium dodecyl-sulphate polyacrylamide gel electrophoresis (SDS-PAGE). Compensatory mutations, which restore the pseudoknot structure with alternative base pairing, restored stop codon readthrough to near WT levels [7]. In addition, long range RNA-RNA interactions can occur throughout an RNA molecule, whereby complementary base pairing can occur between spatially separated regions of the RNA molecule (Fig. 1.1D). Long range base pairing occurs within the genomes of flaviviruses and is further discussed in section 1.3.5.2 [8].

RNA replication elements (RREs), are RNA structures that play functional roles in the virus lifecycle. The roles of RREs in CHIKV and ZIKV replication are discussed in sections 1.2.5.1 and 1.3.5.1/1.3.5.2 respectively. Briefly, RREs may function by binding to proteins, for example in ZIKV an RRE termed stem loop A binds to the virus RNA polymerase during RNA replication [9]. RREs may also function via RNA-RNA interactions, for example complimentary sequences in the 5′ and 3′ ends of the flavivirus genome interact, resulting in genome cyclisation [8]. RREs also function in the alphavirus lifecycle, for example CHIKV stem loop 165 has been shown to be essential for efficient virus replication [2]. Identification of RREs within the genomes of CHIKV and ZIKV both provides insight into the biology of these viruses and identifies potential targets for antiviral agents

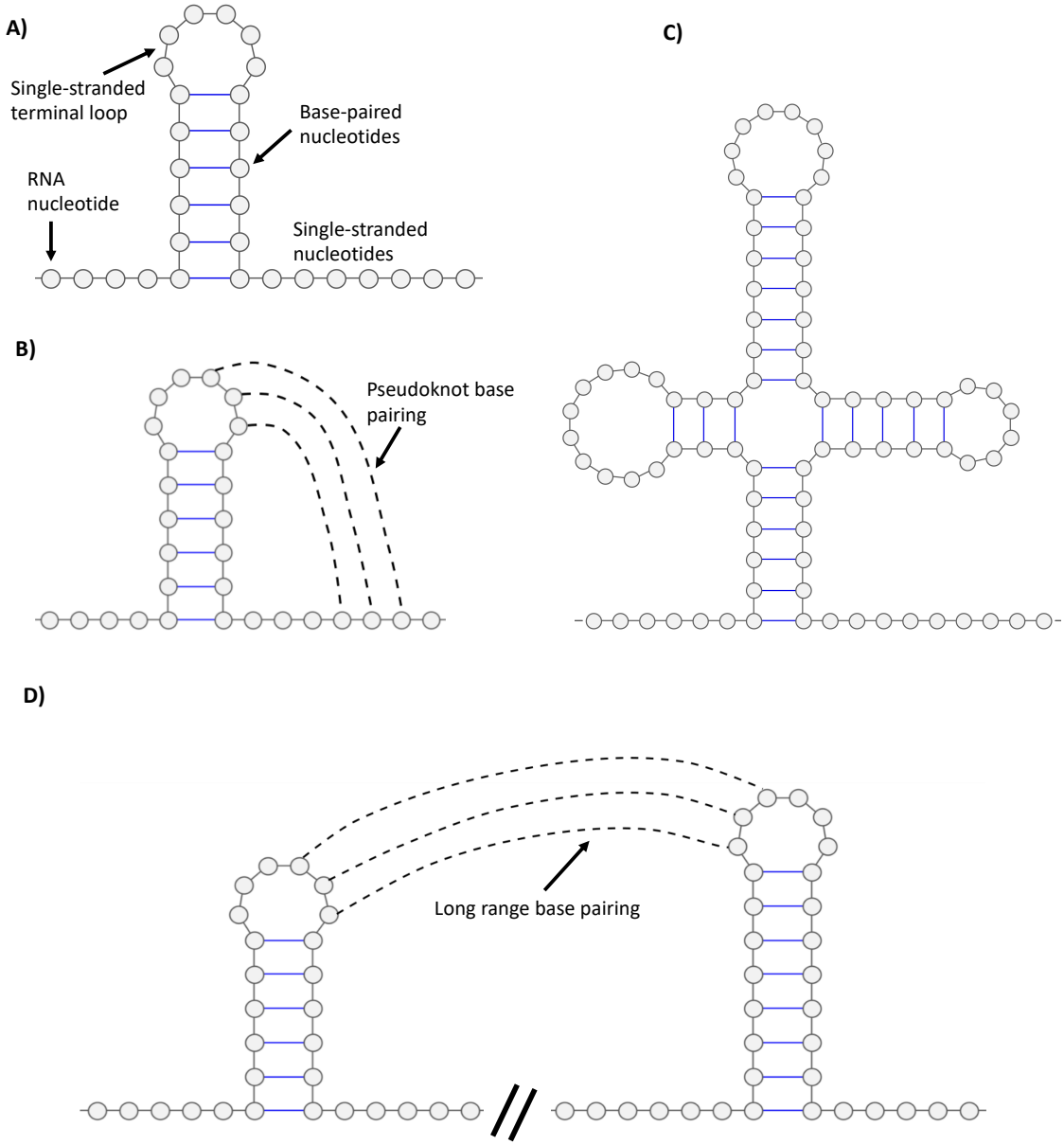


Figure 1.1 Schematic of RNA structural elements.
A) *RNA stem loop. Base pairs form a double-stranded helix stem and non-base-paired terminal nucleotides form a loop.* ***B)*** *RNA pseudoknot base pairing between the terminal loop and downstream nucleotides.* ***C)*** *Complex RNA structure with multiple base paired and non-base paired regions.* ***D)*** *Long range RNA-RNA interactions between terminal loops. Figures created using VARNA.*

RNA structures can be investigated using a variety of approaches. RNA helix structures have been solved using X-ray crystallography. For example, the crystal structure of $[U(UA)_6A]_2$ duplex revealed a double helix structure, with major and minor grooves, forming an A type helix. An A-type helix is a right handed double helix where the major groove is deep and narrow and the minor groove is wide and shallow [10]. Crystal structure analysis has also been performed on RNA molecules containing non-Watson-Crick, G-U base pairs. In a 14-mer sequence RNA, G-U base pairs altered duplex structure compared with an A type helix, with the greatest effect seen where the guanine of a G-U base pair stacks over a purine of the opposite strand, which has the effect of increasing the width of the major groove [11].

However, as RNA molecules are often highly structured and dynamic, obtaining RNA crystals can be very challenging, as crystal formation tends to favour relatively ridged structures, such as those formed by short RNA sequences. An alternative method for investigating RNA structure is nuclear magnetic resonance (NMR), which allows more flexible biomolecules to be resolved at atomic resolution. For example, the structure of the 132 nt Moloney Murine Leukemia virus conserved RNA double hairpin element, which is required for genome packaging, has been determined using NMR [12]. NMR also allows the dynamics of biomolecules to be studied. For example, real time NMR has been used to study conformation changes in an adenine sensing riboswitch, where the initial riboswitch conformation contains two protruding stem loops, which fold inwards upon ligand binding. Real time NMR allows the identification of intermediate steps in the conformation transition and can track the conformational transition of individual nucleotides, based on changes in the NMR spectra [13].

Both X-ray crystallography and NMR tend to be performed on RNA molecules which are significantly smaller than virus genomic RNAs. In order to study RNA structures in the context of viral genomes, biochemical probing methods techniques can be used to determine secondary RNA structure and investigate tertiary RNA interactions. Early biochemical methods involved the use of specific nucleases however, more recently RNA structure have been determined using Selective 2' Hydroxyl Acylation analysed by Primer Extension (SHAPE). These techniques allow structures within RNA molecules to be mapped without the need for labour intensive crystal diffraction or NMR spectroscopy and are described in chapter 4.

1.2.1 Chikungunya virus

CHIKV is a member of the *Togaviridae* family. Togaviruses are small, enveloped viruses which have single-stranded, non-segmented, positive sense RNA genomes of 9.7-12 kb. The family consists of a single genus named Alphavirus, which contains > 30 species [14] (Fig. 1.2). The family previously included the Rubivirus genus however, this genus was recently moved to the new *Matonaviridae* family [15,16]. Most Alphaviruses are mosquito transmitted arboviruses that cycle between mosquito and vertebrate hosts, such as humans, non-human primates, birds and pigs. Whilst some alphaviruses, such as CHIKV, are significant pathogens and can replicate in purely human/mosquito infectious cycle, the majority of alphaviruses cannot develop sufficient viremia in humans to infect mosquitoes [14]. Alphaviruses can also be divided according to geographical placement into Old World and New World viruses. Typically, old world alphaviruses, such as CHIKV and O'nyong'nyong virus, cause arthritogenic disease; whereas New world alphaviruses, such as Western Equine Encephalitis virus (WEEV) and Venezuelan Equine Encephalitis virus (VEEV), cause neurological disease [17]. Not all alphaviruses are transmitted to humans; the Eliat virus (EILV) and the related Taï Forest alphavirus are thought to be insect-specific viruses [18,19]. EILV is unable to infect vertebrate cell lines, but can infect insect cells lines, such as Aedes albopictus mosquito larvae (C6/36) cells, indicating that mammalian cells lack the entry receptors for ELIV. In addition, vertebrate cell lines transfected with EILV RNA showed no sign of replication at both 37 °C and 28 °C, as measured by expression of red fluorescent protein under the control of a sub-genomic promoter. Since no ELIV RNA replication was detected, the authors suggest that ELIV is restricted as a result of incompatibility between ELIV RNA/ proteins and vertebrate cell co-factors [18].

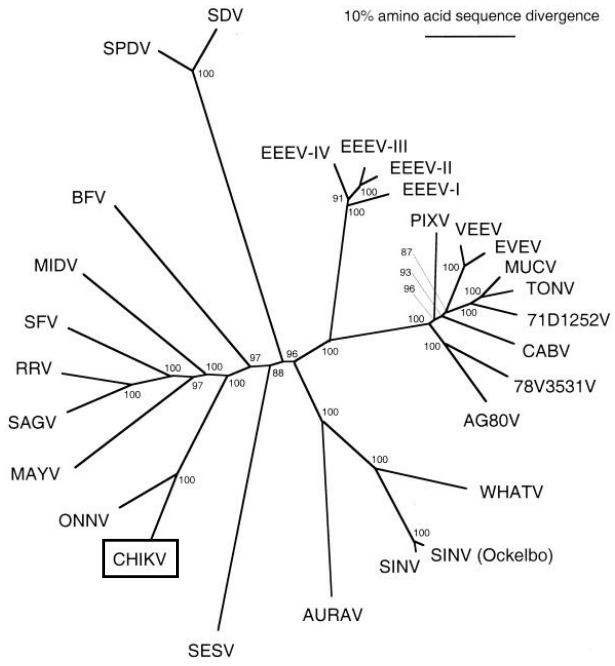


Figure 1.2 Phylogenetic tree of Alphavirus species
Based on complete structural polyprotein amino acid sequences by using a neighbor-joining program. CHIKV has the highest identify with ONNV. CHIKV, one of the subjects of this thesis is highlighted. Adapted from Powers et al. 2001 [20].

CHIKV was first identified from infected patient sera, in 1952 during an outbreak in present day Tanzania [21–23]. The name Chikungunya comes for the Kimakonde language and translates as "that which bends up", which describes the stooped posture and ridged gait which individuals may experience following infection [21]. Phylogenetic analysis of CHIKV sequences suggest that CHIKV originated in Africa over 500 years ago [24] and subsequently diverged from a common ancestor into two separate genotypes, termed West African (WA) and East/Central/South African (ECSA) [24,25]. The WA strain has mainly been responsible for small enzootic outbreaks in West African countries, where the virus is maintained in an animal population and occasionally spreads to humans [26]. In contrast the ECSA genotype has spread and established urban transmission cycles throughout Africa, the Middle East [27,28], the Americas [29] and Asia, leading to selection of a divergent Asian CHIKV genotype [24,26]. More recently, a separate lineage of ECSA diverged during an epidemic in South East Asia and the Indian Ocean basin, which has been classified as the Indian Ocean lineage of the ECSA genotype [24,26].

Since the initial discovery of CHIKV, outbreaks have occurred sporadically in Africa, Asia and the Indian Ocean islands [30–32]. In Kenya, a major epidemic occurred in 2004 [33], which later spread throughout Indian, the Indian Ocean islands, Asia and La Reunion [31,34,35]. CHIKV has become endemic in the Americas, where the *A*. *albopictus* vector has a wide range [36], and CHIKV cases have been reported throughout North and South America [37]. This includes autochthonous spread of an ECSA strain reported in Brazil [38] and outbreaks of the ECSA strain in the French West Indies [39,40] and Puerto Rico [41]. Globalisation, alongside increased deforestation and dense urban living, make it likely that CHIKV will continue to

become established in new geographical regions, where the population is immunologically naïve [26].

1.2.2 CHIKV Transmission

Transmission of CHIKV occurs through the bite of an infected mosquito and occurs in two distinct cycles, the sylvatic cycle and the urban cycle, which vary in prevalence between geographical regions. In general, CHIKV is spread by sylvatic and urban cycles in Africa and in an urban cycle in Asia [26,27]. There are significant differences between the two cycles (Fig. 1.3).

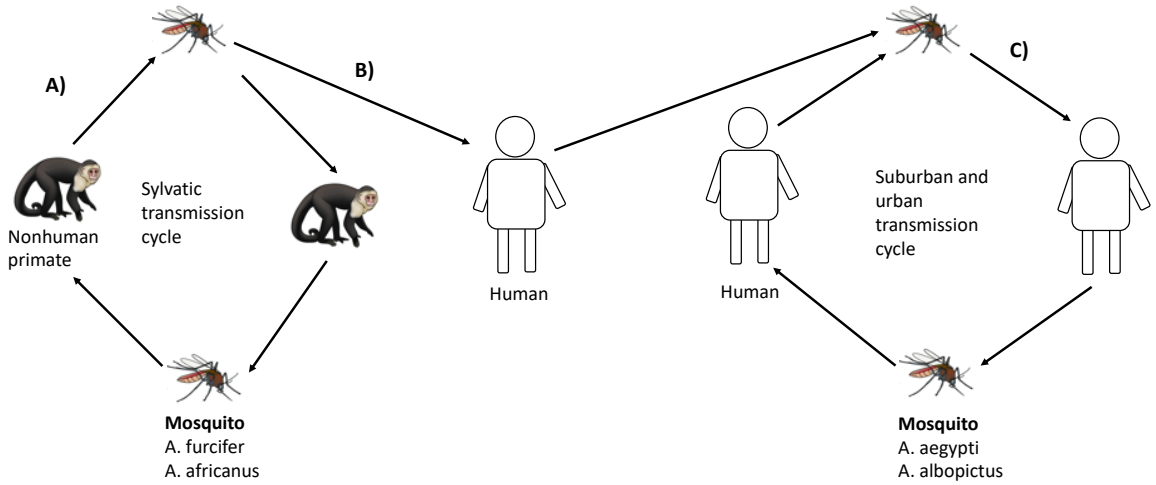


Figure 1.3 CHIKV transmission cycle
A) *In forest environments, CHIKV circulates in a sylvatic transmission cycle between non-human primates and mosquitos (A. furcifer and A africanus).* ***B)*** *Sporadic enzootic transmission can occur between mosquitos and humans.* ***C)*** *In urban and sub-urban environments, CHIKV is mainly transmitted between mosquitos (A. aegypti and A. albopictus) and humans. Adapted from Madariaga et al. 2016* [42].

The sylvatic cycle occurs in rural environments, such as forests or savannahs, where CHIKV is maintained by a cycle of transmission, between infected mosquitos and vertebrates which act as reservoir hosts [43,44]. The sylvatic cycle is prevalent in Africa [26,45], where CHIKV has been isolated from multiple non-human primates [44,46–48] and Aedes species mosquitos [43,49]. Non-human primates are suggested to be the principle animal reservoir host for CHIKV, although small mammals, such as bats and rodents may also act as hosts [44]. Enzootic human infection occurs when forest and savannah environments are encroached upon, (for example through farming, hunting, or deforestation) and infected mosquitoes feed on humans [45]. Although CHIKV has been isolated from multiple Aedes species of mosquito, other factors, such as feeding behaviour and distribution mean that some species are likely not involved in human transmission. For example, *Aedes luteocephalus* and *Aedes taylori* are largely restricted to the forest canopy [43]. Due to its broad land coverage and presence in human settlements, *Aedes furcifer* has been suggested as one species responsible for enzootic spill-over of CHIKV into human populations [43].

Infected humans then travel to urban environments, where the virus is maintained by an urban cycle of transmission. During urban transmission, mosquitoes feed on infected humans and become infected. The infected mosquitos then feed on another (uninfected) human, passing the infection to a new host. Newly infected humans can then transmit the virus to a new mosquito host during a bloodmeal [42,50]. Replication of CHIKV in human tissues can result in high viral serum loads (>10^9 virus particles/ml [26,51]), which makes the virus infection of mosquitos during a bloodmeal extremely efficient. *Aedes aegypti* and *Aedes albopictus* are the main vectors for urban transmission of CHIKV in Asia, Africa and the Indian Ocean [26,42,52]. There is limited evidence for sylvatic transmission in Asia and the prevalence of CHIKV infected non-human primates is low compared with Africa [53,54]. CHIKV may also be maintained in mosquito populations by venereal transmission [55].

1.2.3 CHIKV symptoms

Compared with other alphaviruses, CHIKV causes symptoms in a higher proportion of infected individuals and can be identified by several characteristic factors [29,56]. The disease progression is divided into two stages, an acute stage, lasting approximately a week, and a chronic stage, in which symptoms may persist for several months to years after the initial infection [56].

During the acute stage, an innate immune response is triggered, causing elevated levels of type I interferon and other proinflammatory cytokines in infected patients [57]. Whilst a minority of patients (5-18%) present as asymptomatic [58–60], the remaining patients will develop a high fever (39-40 °C) [21,61] and severe polyarthralgia, which is often bilateral, symmetric and debilitating [56,62]. Between 10% and 40% of patients develop a maculopapular rash [63,64]. In a study of 610 atypical CHIKV cases in the 2005-2006 outbreak in Réunion (where an atypical case was defined as presenting symptoms apart from arthralgia and fever), 222 cases has severe disease and 65 patients died [64]. Atypical symptoms included malaise (45%), diarrhoea/ vomiting (18%), myalgia (13%) and chest pain (4%) [64]. During a 2006 outbreak of CHIKV in India, 16.3% of patients developed neurological symptoms [65] compared with 24% during the Réunion outbreak [64]. In addition development of severe disease in atypical cases was 2.5 times more likely in those aged 40-60 years compared with those aged < 40 years [64].

Acute CHIKV symptoms are attributed to direct cellular damage caused by apoptosis of infected cells, leading to tissue damage, and local inflammation caused by the host immune response [26,66]. In mouse models of CHIKV and in tissue biopsies from fatal neonatal cases, tissue fibroblasts within the joints, skeletal fascia and the dermis were the principal CHIKV cell target [67] and human fibroblasts from primary cultures have been shown to undergo apoptosis when infected with CHIKV in cell culture conditions [66].

The severity of symptoms at the acute stage and increased age are risk factors for the development of chronic disease [68]. Chronic CHIKV symptoms are characterised by relapsing long-term arthralgia, which is usually located in the same joint sites affected during the acute stage [69,70]. In addition, chronic myalgia and cutaneous lesions have been reported [69]. Chronic inflammation is proposed to be maintained by local persistence of viral replication or

failure to clear viral antigens from the infected tissue [69,71]. Histological assessment of synovial joint biopsies identified CHIKV proteins in macrophages using fluorescence microscopy and CHIKV RNA was detected in biopsy cells using RT-PCR. A lack of cytotoxic $CD8^+$ T cells, which specifically target infected cells for apoptosis, was also noted and may contribute to CHIKV persistence join tissue [71]. Patients suffering from chronic CHIKV will often need long term symptomatic treatment and have a reduced quality of life [72]. In a study of over 1,000 patients from La Réunion island, patients who tested positive for CHIKV IgG antibodies were interviewed/ responded to questionnaires concerning their symptoms at around 2 years post infection. 43% of patients reported musculoskeletal pain, 75% reported cerebral disorder and 36% reported skin based symptoms [73]. In a study into CHIKV associated arthralgia, patients underwent clinal investigation at 4-, 6-, 14- and 36-months post infection. After 36 months, 60% of patients experienced arthralgia, occurring in episodic patterns of relapse and recovery periods, 77% of which was described as highly incapacitating [69]. In another study of 147 patients, where data was collected by trained physicians using a standardised questionnaire, by 15-months post infection 43% had no symptoms, 21% had at least one relapse of symptoms and 36% had permanent symptoms [74].

1.2.4 CHIKV genome organisation and replication

The CHIKV positive sense RNA genome consists of two open reading frames (ORFs), separated by a non-coding intergenic region, which are flanked by a 5′ untranslated region (UTR) of 76 nucleotide (nt) and a 3′ UTR of between 450 and 900 nts in length depending on lineage [75–77] (Fig. 1.4). RNA structures within the 5′ UTR play essential roles in CHIKV genome replication and are discussed in detail in section 1.2.5.1. The 5′ UTR is capped with a type-0 cap structure (N7mGppp) for initiation of cap-dependent translation. The 3′ end of the genome has a polyadenylate(A) tail and a 19 nt conserved sequence element is present in the 3′ UTR, adjacent to the poly(A) tail [78]. ORF 1 encodes four non-structural proteins (nsp1-4), and a second ORF, translated from a sub-genomic mRNA, encodes the structural proteins [79]. The majority of CHIKV strains encode an opal stop codon in the nsp3-nsp4 junction however, some strains of CHIKV (and the related O'nyong-nyong virus) encode an arginine codon at the equivalent position. The presence of the arginine codon results in less severe disease phenotypes in mouse models [80].

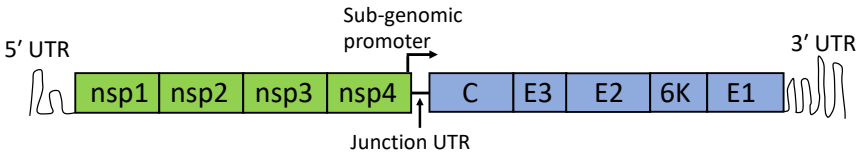


Figure 1.4 Schematic of CHIKV genome
CHIKV genome organization. ORF 1 (Green) encodes the non-structural proteins, which replicate the viral genome and ORF 2 (Blue) encodes the structural proteins, which form the virus particle. ORF 1 and OFR 2 are separated by a junction region. The prompter for sub-genomic RNA synthesis is found partially in the junction region and partially in the nsp4 coding region.

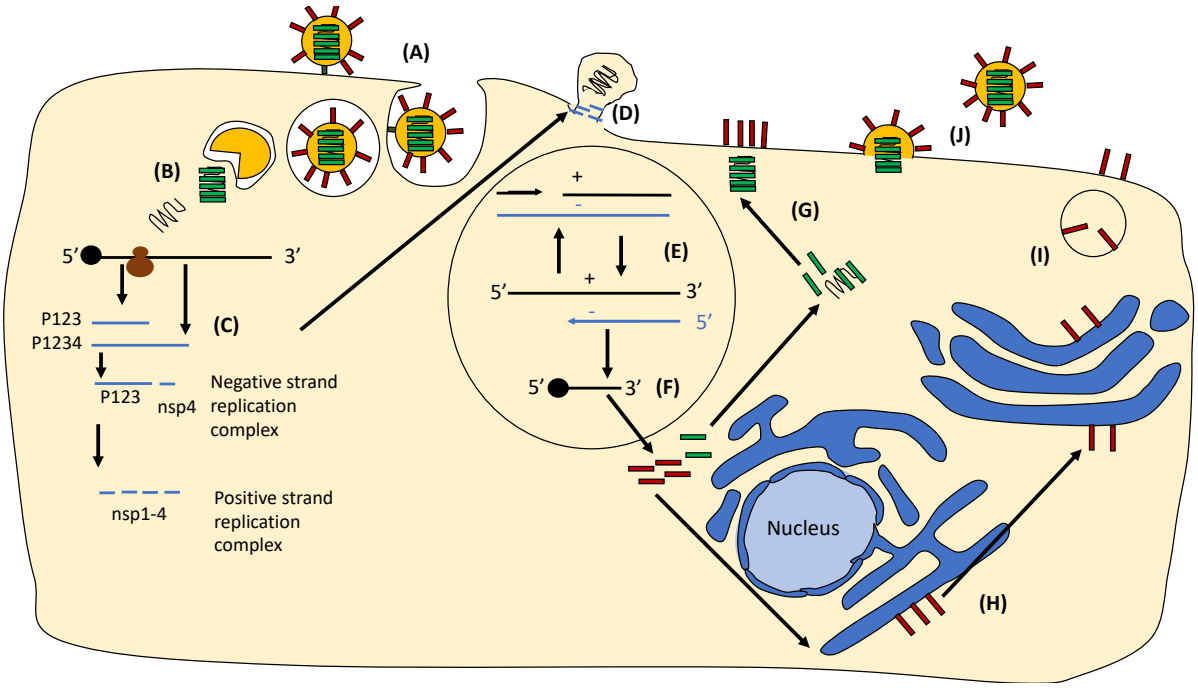


Figure 1.5 CHIKV lifecycle summary

A) *CHIKV enters the host cell via receptor mediated endocytosis.* ***B)*** *pH changes in the endosome result in fusion of the virus envelope and nucleocapsid uncoating.* ***C)*** *Translation of ORF 1 produces the P123 and P1234 polyproteins, which are subsequently processed into individual non-structural proteins. The presence of an opal stop codon at the nsp3/nsp4 junction results in the P123 polyprotein upon translation. Readthrough of the opal stop codon results in the P1234 polyprotein.* ***D)*** *Nonstructural proteins and positive sense genomic RNA assemble at the plasma membrane and form replication compartments (spherules).* ***E)*** *Positive and negative sense RNA are replicated in spherules, producing positive sense, negative sense and sub-genomic RNA.* ***F)*** *Sub-genomic RNA is translated into the capsid and glycoproteins.* ***G)*** *Capsid and genomic RNA form the nucleocapsid complex.* ***H)*** *Glycoproteins insert into the ER.* ***I)*** *Glycoproteins are processed in the Golgi and the mature glycoproteins are trafficked to the plasma membrane.* ***J)*** *Virus particles exit the cells from the plasma membrane.*

In human hosts, CHIKV is transmitted to the bloodstream or skin during the process of mosquito blood meal feeding, where the virus infects fibroblasts [81] and potentially macrophages [67,82]. Fibroblasts were identified as the primary target for CHIKV in both mouse models and human biopsies [67]. Liver derived primary macrophages can be infected *in vitro* but display low levels of replication [67] and macrophages in synovial biopsies contain CHIKV RNA and proteins [71]. The CHIKV life cycle is summarised in Fig. 1.5. CHIKV binds to host mammalian cells via the cell adhesion molecule Mxra8. Preincubation of primary human cells, including dermal fibroblasts, with anti-Mxra8 antibodies inhibited CHIKV infection (as measured by intracellular E2 expression analysed by flow cytometry) and Mxra8 was shown to bind captured CHIKV virus-like-particles using surface plasmon resonance [83]. Mutation of CHIKV E2 inhibited Mxra8 binding in flow cytometry experiments where the E2 mutants were expressed in Human embryonic kidney cells expressing a mutant version of the simian virus 40 large T antigen (HEK 293T), indicating that receptor binding is mediated by E2 [83].

CHIKV is thought to enter cells via clathrin mediated endocytosis (CME) [84], and inhibition of CME using the CME inhibitor Pitstop2 was shown to inhibit CHIKV entry, as measured by flow cytometry using a modified CHIKV which expresses GFP from a sub-genomic promoter [85].

Endosome maturation causes pH reduction, which triggers reorganisation of the E1-E2 heterodimer, as shown by x-ray crystallography [86,87], exposing the fusion peptide in E1. E1 then triggers fusion of the CHIKV and endosomal membranes, releasing the viral RNA into the host cytoplasm [84]. Mutation of conserved residues in E1 reduced fusion activity in a recombinant baculoviral assay, whereby formation and size of syncytia is used as a measure of fusion activity. In this assay, the size of syncytia induced by monomeric E1 was reduced compared with cells where all CHIKV structural proteins had been expressed, indicating that other CHIKV structural proteins may be involved in membrane fusion [88].

Following release of CHIKV particles from the endosome, ORF 1 is translated as a polyprotein directly from the genomic RNA. The majority of translation events produce the viral polyprotein P123, due to the opal stop codon (UGA) at the terminal of the ns3 coding region. Read through of the opal stop produces the P1234 polyprotein. Opal stop codon readthrough has been investigated in Sindbis virus, using cell-free translation assays, where RNA transcripts were added to rabbit reticulocyte lysates. In the translation assay, translational readthrough was determined by size differences in products analysed by SDS-PAGE. These assays determined that mutation of the cysteine residue immediately downstream of the UGA codon resulted in a reduction in translation readthrough, from 10% to 1%, indicating that the context of the UGA is important for readthrough [89]. In addition, mutation of the opal stop codon to an arginine residue did not reduce the infectious titre of CHIKV in Vero cells and did not impair virus replication in mouse models, as determined by plaque assay of harvested tissues. Despite this, mouse models showed a reduction in tendon and joint swelling when infected with the mutant virus, indicating that the opal stop codon could be linked with development of more severe disease phenotypes [90]. RNA structure predictions using the ViennaRNA and pknotsRG web servers/ prediction algorithms indicate that an RNA structure may form downstream of the opal stop codon in VEEV, consisting of a short RNA duplex and a 95-148 nucleotide open loop. Sequences including the opal stop codons and nucleotides predicted to form the RNA structure were cloned between *Renilla* and Firefly luciferase genes, where firefly expression is dependent on readthrough of the opal stop codon, in order to investigate the function of the RNA structure. Mutations disrupting base pairing in the RNA duplex resulted in a reduction in stop codon readthrough, whilst restorative mutations which restored base pairing resulted in stop codon readthrough at near WT levels [91].

The viral non-structural polyprotein is proteolytically cleaved in *cis* by nsp2, releasing the nsp4 RNA dependent RNA polymerase (RdRp) (further discussed in section 1.2.5.3). Polyprotein processing has been investigated in Sindbis virus (SINV). Mutations which block cleavage at the nsp3/nsp4 site resulted in a lack of plaque formation from transfected cells, indicating that nsp4 cleavage is essential for virus replication. Mutations which block cleavage at the nsp1/nsp2 and nsp2/nsp3 sites did not impair synthesis of negative sense RNA, as detected by RT-PCR of SINV infected cells. In contrast, positive sense RNA synthesis was impaired by P123 cleavage mutations, as measured by incorporation of [3H] uridine, when cells were incubated with [3H] uridine during infection. The P123 cleavage mutant incorporated less

[3H] uridine compared to WT, indicating a defect in positive sense RNA synthesis. The authors argue that since 90% of RNA synthesised in SINV infected cells is positive sense, then incorporation of a label is a measure of positive sense RNA synthesis. Mutations which block P123 processing also caused a reduction in viral titre, presumably due to lack of positive sense RNA synthesis [92]. These results indicate that the P123 polyprotein and nsp4 form a replication complex, which uses the positive-sense genomic RNA as a template for synthesis of a negative-sense RNA replication intermediate and that positive sense RNA replication occurs when the P123 polyprotein is fully processed into individual proteins. Viral RNA replication occurs in membrane bound complexes derived from the plasma membrane, termed spherules, which are further discussed in section 1.2.5. CHIKV RNA replication is likely to also involve host proteins, although this is not well characterised [79]. Chloride ion channels inhibitors have been shown to impair virus genome replication, as treatment of CHIKV infected cells with chloride ion channel blockers reduces infectious titre, and in addition, impairs replication of a sub-genomic firefly expressing CHIKV replicon [93].

The role for the 3' poly(A) tail and a 19 nt conserved sequence element, found immediately preceding the poly(A) tail, to act as a promoter for negative-sense RNA synthesis has been investigated in SINV. RNA templates corresponding to the SINV genome were incubated with purified P123 and nsp4 proteins (where nsp2 contained a mutation to prevent polyprotein processing). Templates lacking either the conserved sequence elements alone or conserved sequence element and poly(A) tail were unable to synthesise negative sense RNA, as determined by agarose-phosphate gel electrophoresis where RNA was visualised by [α-^{32}P] CTP incorporation [94].

The positive sense 26S sub-genomic RNA is translated into a polyprotein, which is translated and cleaved into the structural proteins capsid (C), E1, PE2 (which consist of E2 and E3), and the small 6K peptide. C is proposed to self-cleave from the polyprotein, as mutations within the C protein protease domain of Sindbis virus impaired protein cleavage during *in vitro* translation assays [95]. Immunoelectron microscopy experiments using Semliki Forest virus indicates that virus proteins E1 and PE2 pass from the ER and through the Golgi network [96] and treatment of cells with the reagent monensin causes the retention of E1/E2 in the Golgi [97]. Monensin facilitates the exchange of Na^+ ions for protons across membranes, thereby removing Na^+ and H^+ gradients and neutralising the pH of the Golgi network (and other acidic organelles) [98]. This has the effect of causing the Golgi cisternae to swell and separate, blocking transport of proteins between the different components of the Golgi complex [97].

PE2 is proposed to be cleaved by furin in the Golgi network. Use of a furin inhibitor prevented PE2 cleavage, as shown by western blot of CHIKV infected cell lysates, and treatment of CHIKV infected cell with a furin inhibitor reduced viral titre [99]. In the later stages of infection (5 hours post infection), Semliki Forest virus E1 and E2 can be visualised at the plasma membrane (in vesicles known as type II cytopathic vesicles), using immunoelectron microscopy [100]. Genomic RNA molecules associate with capsid proteins forming nucleocapsid complexes. This is dependent on the positively charged N terminus of the C protein, as deletion of more than 19 nucleotides from the C protein N terminus prevents nucleocapsid assembly *in vitro* in the related Sindbis virus [101]. Mutations, replacing the positively charged amino acids in the RNA binding region with neutral amino acids, caused a reduction in virus titre and upon passage, a number adaptive mutations were seen in nsp2,

indicating a potential role for this protein in RNA packaging [102]. The resultant nucleocapsid complexes associate with E1 and E2 at the plasma membrane, from which virus particles are released by budding, as shown by electron microscopy [100,103].

1.2.5 The CHIKV RNA replication complex

Alphavirus RNA replication takes place in modified membrane compartments, which contain the viral RNA and proteins. Translation of the first viral ORF produces the non-structural proteins (nsp1-4), which induce the formation of membrane invaginations termed spherules. Spherules are approximately 50 nm in diameter and are proposed to be connected with the cytoplasm via a narrow neck-like structure, as visualised by electron microscopy [104]. Immunofluorescence imaging indicates that both dsRNA and non-structural protein co-localise in replication complexes [105,106]. In the proposed model of RNA replication, non-structural proteins are positioned at the neck of the spherule, with the dsRNA contained within the spherule membrane. Newly synthesised RNA molecules are then released into the cytoplasm via the neck region pore [104,106]. Formation of replication spherules may both provide protection from immune surveillance, for example by keeping dsRNA within the vesicle, and bring the components of the replication complex into spatial proximity [105].

Treatment of Semliki forest virus (SFV) infected cells with the phosphoinositide 3-kinase inhibitor wortmannin leads to retention of the replication complexes at the plasma membrane, indicating that spherules are first formed at the plasma membrane [106]. In mammalian cells, electron microscopy images of SINV show that spherules are endocytosed and fuse with late endosomes/ lysosomes, forming vacuole structures in which spherules are positioned at the cytoplasmic side of the membrane [103]. Replication complexes are trafficked to the perinuclear space via microtubules and mature into larger cytopathic vacuole structures (CPV-I). Disruption of the microtubule network using nocodazole results in CPV-I vesicles having no localisation within the cell, shown by confocal live-cell imaging. In addition, CPV-I were smaller and not stained with the acidic organelle marker Lysotracker, indicating that microtubules are required for the transport of internalised spherules to acidic organelles [103,106].

Work by Thaa et al. [107] indicates that in CHIKV and SFV differ in phosphoinositide 3-kinase - protein kinase B (Akt) - mechanistic target of rapamycin (mTOR) pathway activation. Western blotting of CHIKV or SFV infected BHK cells indicates that CHIKV Akt phosphorylation is much weaker in CHIKV, as less phosphorylated Akt can be visualised in CHIKV lysates compared with SFV lysates. In addition, immunofluorescence microscopy experiments dsRNA and CHIKV nsp3 were shown to accumulate at the cell periphery at 8 hours post infection, suggested that in CHIKV, replication spherules are not internalised. The authors show that in SFV deletion of a hyperphosphorylated region of the nsp3 hypervariable domain impaired persistent Akt-mTOR activation. In cells infected with this SFV mutant, fluorescence microscopy showed that most of the replication complexes were localised at the cell periphery, resembling cells which had been treated with wortmannin. These results indicate that spherule internalisation in alphaviruses is dependent on activation of the Akt-MTOR pathway. The authors therefore suggest that in replication spherules are not internalised in CHIKV due to weak activation of the Akt-mTOR pathway [107].

In mammalian cells, a second type of cytopathic vacuole (CVP-II) transports the viral envelope proteins to the site of viral budding at the plasma membrane. CVP-II vesicles containing the E1 and E2 proteins have been visualised close to the plasma membrane in SFV infected cells, using immuno-electron microscopy [100,103]. Electron miscopy studies of alphavirus infected

C6/36 cells indicates that the sites of viral replication and assembly are not segregated in mosquito cells. In these cells, virus budding can occur from internal cytopathic vesicles, in addition to the plasma membrane, which may enable persistent infections [103].

1.2.5.1 CHIKV 5′ UTR and adjacent nsp1 coding region RNA structures

RNA structures within the first 303 nts of the CHIKV genome, encompassing the 5′ UTR and adjacent nsp1 coding region, are essential for CHIKV genome replication (Fig. 1.6). Using a combination of selective 2′-hydroxyl acylation analysed by primer extension (SHAPE) and reverse genetic studies, several RNA replication elements (RREs) were identified in the 5′ region of the CHIKV genome, all of which function through structure dependent mechanisms. 2 RREs were identified in the 5′ UTR (SL3, SL47) and 5 RREs in the adjacent nsp1 coding region (SL85, SL102, SL165, SL294 and SL246). The 51 nts conserved sequence element (CSE), which corresponds to SL165 and SL194, is conserved across alphaviruses [108–111]. Comparative SHAPE reactivity was seen in both 28 °C and 37 °C (reflecting both mammalian and *A. albopictus* body temperature), indicating that the same structures are present in both human and mosquito hosts. SHAPE reactivities mapped to SL85 showed that a number of terminal loop nucleotides are base paired, indicating that higher order RNA-RNA interactions may be forming with a downstream sequence which overlaps with the AUG start codon [2].

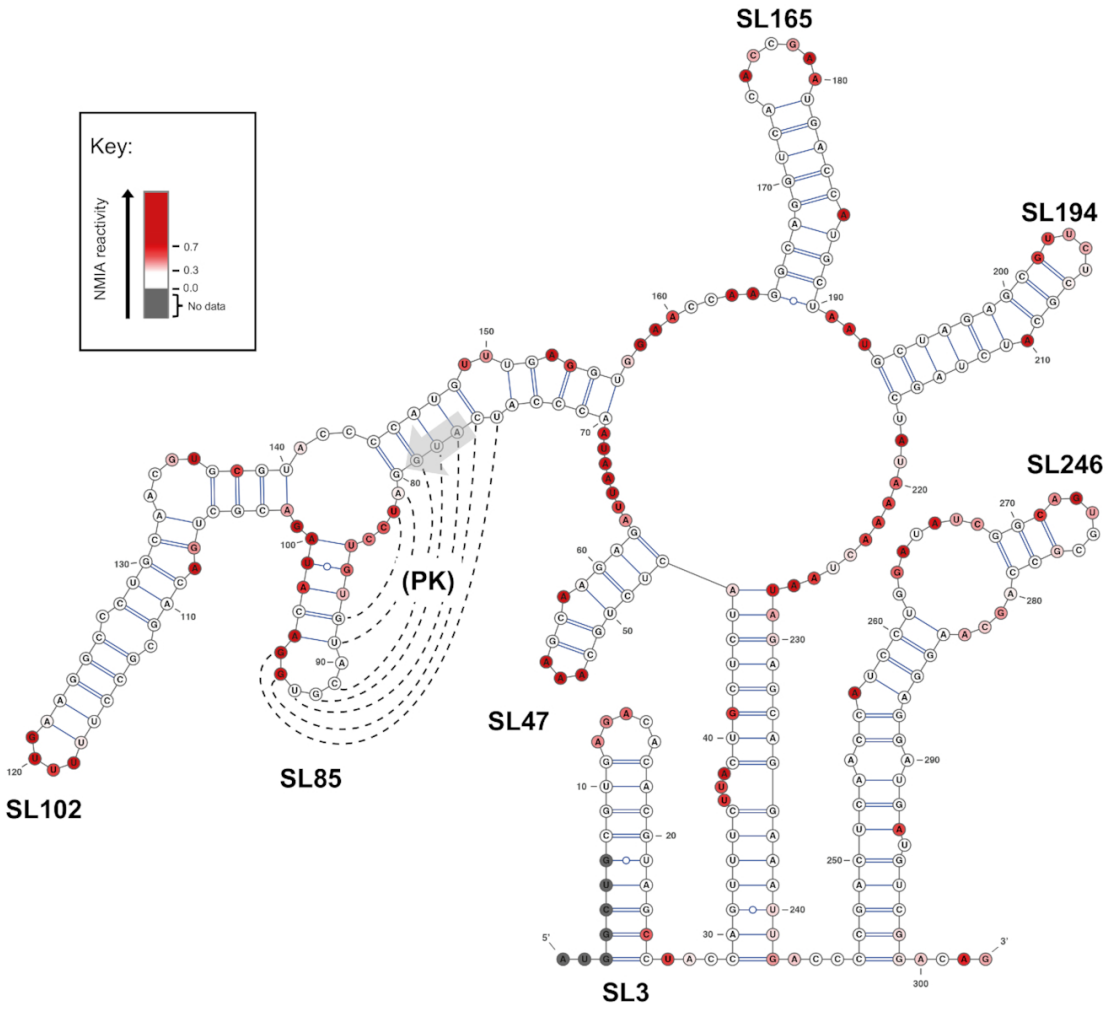


Figure 1.6 RNA structures within the 5' UTR and adjacent coding region of the CHIKV genome.

RREs with this this region are essential for CHIKV genome replication and have been shown to function in a structure dependent manner. SL47 is required for efficient replication in both mosquito and mammalian cell lines. SL86, SL102, SL165 and SL194 are required for efficient replication in mammalian cells, but are not essential for replication in mosquito cells. In contrast, S246 is not required for replication in mammalian cells, but is required for replication in mosquito cells. Nucleotides with a score above 0.7 are highly reactive chemical reagent NMIA, nucleotides with a score between 0.3 and 0.69 are moderately reactive and nucleotides with a reactivity below 0.3 are considered unreactive. The process of SHAPE mapping is explored in detail in chapter 4. Adapted from Kendal et al. 2019 [2].

The RRE SL3 has been characterised in CHIKV and related alphaviruses and plays a role in viral immune evasion. The IFN-induced protein IFIT1 binds to type-0 viral RNA cap structures, preventing binding of eukaryotic translation initiation factors and thereby inhibiting virus replication. Mutations introduced to destabilise SL3 base pairing resulted in inhibition of virus replication in cell lines which overexpress IFIT1. The authors suggest that SL3 provides resistance to IFIT1, possibly by regulating the efficiency of IFIT1 binding to RNA cap structure

[112]. To investigate the remaining RREs, Kendal et al performed mutagenesis to destabilise base-pairing within the duplex stems of SL47, SL85, SL102, SL165, SL194 and SL246, without changing the primary protein coding sequence. The effect of these mutations on infectious CHIKV replication was measured the in both human (Huh7 and Human Caucasian embryo rhabdomyosarcoma) and mosquito (C6/36) cell lines. Disruption of SL47 significantly reduced replication in both human and mosquito cell lines, indicating that the REE is required in both hosts. RREs within the nsp1 coding region showed more host specific phenotypes. Disrupting the structures of SL86, SL102, SL165 and SL194 had no effect on replication in C6/36 cells, but significantly inhibited replication in human cells. Conversely, disruption of SL246 significantly impaired replication in mosquito cells, but had no effect in human cells [2]. The same mutations were also investigated using a sub-genomic replicon system, with the same cell dependent phenotypes observed [2], indicating that the RREs function during CHIKV genome replication. Translation assays confirmed that RRE mutation did not affect translation, further indicating that the REEs function during the genome replication stage of the lifecycle.

To confirm that mutant phenotype was due to disruption of base-pairing, rather than changes to the primary RNA sequence, compensatory synonymous mutations were incorporated to restore base-pairing within the RNA duplex-stem, without reverting to wild-type sequence. Restoration of base pairing in SL85, SL102, SL165 and SL194 significantly rescued replication in Huh7 cells and restoration of SL246 restored replication in C6/36 cells, to near wild-type levels. Restoration of SL47 restored replication in C6/36 cells, but did not rescue replication in Huh7 cells at 37 °C. Replication was rescued when Huh7 cells were incubated at 28 °C, suggesting that the SL47-compensated structure was not stable at 37 °C. Mutations within the single-stranded loop regions of the CHIKV RREs did not affect replication, with the exception of SL194 where a C>U substitution resulted in significant inhibition in Huh7 cells, indicating that for SL194, both the primary loop sequence and the duplex stem are essential for replication. These results indicate that the structure of the RREs is functionally significant and that disruption of the RREs structure inhibits CHIKV genome replication [2].

1.2.5.2 CHIKV Non-structural protein 1 (nsp1)

Nsp1 is a 60kDa protein which has two essential functions during replication. Firstly, nsp1 contains an N-terminal methyltransferase/ GTase domain, responsible for adding the type-0 5’ cap to genomic and sub-genomic viral RNAs [79]. Mutagenesis of Sindbis virus nsp1 resulted in reduced methyltransferase activity, as measured by transfer of radioactivity ([3H-methyl]AdoMet) [113]. In addition, nsp1 anchors the CHIKV replication complex to host membranes within virus replication complexes. Palmitoylated cysteine residues within the C-terminal region of nsp1 have affinity for cholesterol rich lipid raft regions within membranes. Mutation of conserved amino acids led to nsp1 becoming diffuse in the cytoplasm, compared with WT, where nsp1 is associated with membrane structures, and nsp1 colocalises with cholesterol in fluorescent microscopy studies [114].

The structure of the nsp1 complex has been solved using single-particle cryo-electron-microscopy, using recombinant nsp1 expressed in *E. coli* and in a Baculovirus-insect cell expression system [115]. Nsp1 was shown to assemble into dodecameric pores with a central aperture that is 7 nm wide. The 12 subunits complex contains three regions: the crown (formed by the nsp1 capping domains), the waist (which defines the 7-nm-wide, 2nm deep

pore) and the membrane-binding skirt region. Nsp1 pores are proposed to form at the neck of the replication spherule, thereby acting as a pore complex which controls the movement of molecules which enter and exit the replication spherule [115].

1.2.5.3 CHIKV Non-structural protein 2 (nsp2)

The 90kDa nsp2 carries out several functions in CHIKV replication. The nsp2 N-terminal helicase domain has NTPase activity, which is required for unwinding RNA secondary structures during RNA replication. The crystal structure of nsp2 helicase domain bound to a 14-mer RNA (containing the 3' genome end sequence) and a transition-state analogue of ATP hydrolysis has been reported and the structure of the helicase is highly similar to other RNA helicases [116]. Additionally, this domain has nucleoside triphosphatase activity, as incubation of nps2 with 5' [γ-^{32}P] GTP labelled RNA resulted in removal of the label. Additionally, nsp2 can remove γ-^{32}P, but not α-^{32}P (α-^{32}P is used for labelling during RNA synthesis, whereas γ-^{32}P is specifically used for 5' end labelling of RNA). The authors suggest that nsp2 removes the γ-phosphate of nascent positive-sense viral RNAs, providing a substrate for nsp1-mediated capping reactions [117]. Nsp2 also has a C-terminal protease domain, which carries out cleavage of the non-structural polyprotein. The crystal structure of the protease domain has been solved and the active site identified [118]. Mutations within the nsp2 active site impaired protease activity at all of the polyprotein cleavage sites, both using purified enzyme and substrate, and in a cell-free in vitro transcription/ translation assay. In contrast, incubation with WT nsp2 resulted in a number of bands corresponding to polyprotein digestion products when analysed by SDS-PAGE [119]. In cell based assays using a construct where the non-structural polyprotein was expressed from a T7 promoter, P1234 was the only products detected by western blot when key amino acids in nsp2 were mutated [119].

1.2.5.4 CHIKV Non-structural protein 3 (nsp3)

The least well understood of the CHIKV non-structural proteins is the 60kDa nsp3, which has three domains: The macrodomain, the alphavirus unique domain (AUD) and the hypervariable region. [120,121]. The crystal structure of the N-terminal nsp3 macrodomain has been reported and ADP binding activity was determined by isothermal titration calorimetry [120]. CHIKV nsp3 ADP-ribosylhydrolase activity was demonstrated by incubating an nsp3 macrodomain fusion protein/ full length nsp3 with ^{32}P labelled-mono-ADP-ribosylated proteins. Hydrolase activity was determined by the amount of detectable ^{32}P post-incubation, whereby incubation with nsp3 caused a reduction in detectable ^{32}P [122]. Another study reported that mutant viruses lacking activity ADP-ribosylhydrolase are unable to replicate in BHK or *A. albopictus* cells and underwent reversion to remove catalytically inactivating mutations [123]. The AUD domain is proposed to contain a zinc binding domain consisting of four cysteines, and mutation of two of these cysteine residues in a CHIKV sub-genomic replicon system resulted in a lack of replication and translation in both mammalian and mosquito cell lines, as measured by luciferase expression. Mutation of surface exposed basic residues impaired replication of infectious CHIKV in mammalian cell lines, but not in mosquito cells where the mutations rapidly reverted to WT. Further analysis by confocal microscopy of mutations within the AUD resulted in loss of co-localisation with dsRNA and

loss of nsp3 accumulation at the plasma membrane [121] indicating the role of nsp3 with the replication complex.
The C-terminal hypervariable region (HVR) is divergent between alphaviruses in terms of length and sequence and may play a role in host specificity [79], as nsp3 has been shown to interact with different host proteins in different alphavirus species. In old world alphaviruses, such as CHIKV, the HVR interacts with G3BP proteins, whereas in new world alphaviruses, such as VEEV, the HVR interacts with Farnesoid X receptor (FXR) proteins. The use of different host factors may partially explain the difference in disease phenotype seen between new world and old-world alphaviruses. Use of Ras GTPase-activating protein-binding protein (G3BP) knockout cell lines causes a reduction in CHIKV virus titre and genomic RNA copy numbers. In addition, deletion of nsp3 G3BP binding elements also impairs CHIKV replication. G3BP is proposed to play a role in CHIKV replication complex formation, by binding to P123 and P1234 complexes at the plasma membrane and thereby promoting interactions between the replication proteins and genomic RNA [17].

1.2.5.5 CHIKV Non-structural protein 4 (nsp4)

Due to polyprotein expression and processing, as described previously, significantly less nsp4 is produced compared with other non-structural proteins. The N-terminus of nsp4 is predicted by modelling to be disordered. In Sindbis virus, mutations within the N-terminal region resulted in a 10-fold reduction in virus titre. The effect of N-terminal mutations on RNA synthesis was investigated by [^{3}H] uridine/ actinomycin D treatment of infected cells. Mutations within the N-terminus of nsp4 resulted in a significant decrease in detectable RNA, when visualised using fluorography [124].

Nsp4 is required for synthesis of alphavirus RNA. In an assay where nsp4 was combined with BHK cell membrane fractions containing P123, positive sense Sindbis virus RNA, NTPs and [α-^{32}P] CTP (as a marker for visualisation of products on agarose-phosphate gels), P123 was unable to perform RNA synthesis in the absence of nsp4. Reactions containing nsp4 but lacking P123 could not perform negative sense RNA synthesis, indicating that both nsp4 and the P123 precursor polyprotein are required for negative sense RNA synthesis. Mutation of the GDD active site to GAA resulted in a lack of RNA synthesis. Deletion of the first 97 N terminal amino acids (Δ97nsP4), leaving the predicted catalytic core domain intact, did not result in RNA synthesis, indicating that the N-terminal region is required for RNA synthesis [125]. In addition, nsp4 has been purified from BHK cells, using recombinant vaccinia viruses where the Sindbis P123 polyprotein and nsp4 were expressed, by fractionation/ centrifugation. Fractions containing P123 and nsp4 (pellet) were able to synthesise negative sense RNA, but not positive sense RNA (as measured by [α-^{32}P] CTP/ phosphoimaging) [126].

In contrast, synthesis of positive sense RNA requires cleavage of the P123 polyprotein. Mutations which block cleavage at the nsp2/nsp3 cleavage site did not impair negative sense RNA synthesis (as detected by PCR), however positive sense RNA replication (measured by [3H] uridine incorporation) was impaired [92]. These results indicate that processing of the P123 precursor is required for positive strand synthesis. The authors propose a model whereby early in the infection nsp4 and the P123 precursor form a transient complex, which uses positive sense RNA as a template for synthesis of negative sense RNA. Later in the infection, P123 is cleaved into individual proteins. The P123 cleavage products and nsp4 then

form stable replication complexes, which use negative sense RNA as a template for synthesis of positive sense RNA [92].

The catalytic core domain Δ97nsP4 was used to demonstrate adenylytransferase (TATase) activity. 45 nucleotide RNA templates, corresponding to the 3' region of sindbis virus, were incubated with the purified catalytic domain and either [α-^{32}P] labelled ATP, UTP, CTP or GTP, then separated on a PAGE gel. Only incubation with [α-^{32}P] ATP nucleotides produced products with a higher molecular weight (47, 48 and 49 nucleotides), indicating that Δ97nsP4 is able to specifically add ATP onto RNA transcripts. Δ97nsP4 is able to add ATP to multiple RNA transcripts, indicating a lack of specificity and suggesting that either additional viral or cellular proteins or residues 1-97 of nsp4 are required for template specificity [127].

1.3.1 Flaviviridae and Zika virus

Members of the *Flaviviridae* family are small enveloped viruses with single-stranded non-segmented positive sense RNA genomes of between 9-12 kb. There are four genera in the *Flaviviridae* family. The Flavivirus genus contains over 50 species, including important human pathogens, some spread by mosquitoes, such as Zika virus (ZIKV), dengue virus (DENV), West Nile virus (WNV), yellow fever virus (YFV) and some spread by ticks, such as tick-borne encephalitis virus (TBEV) (Fig. 1.7). Flaviviruses can be further subdivided into four groups based on the vector organism: Mosquito borne (e.g. ZIKV, DENV), Tick Borne (e.g. TBEV), Insect specific (e.g. *Aedes* flavivirus) and no known vector (e.g. Rio Bravo virus). Members of the flavivirus genus cause a range of disease from haemorrhagic fever to neurological complications [128] and affect a high percentage of the global population. For example, DENV causes an estimated 400 million infections per year [129] and ZIKV has caused several significant outbreaks, most recently in Brazil [130]. Comparison of amino acid sequences indicate that on average ZIKV has 55.6% sequence identity with DENV, 57% with WNV and 46% with YFV [131]. Members of the *Pestivirus* genus are economically important viruses which infect pigs and ruminants, such as cattle and sheep, and are transmitted by infected secretions such as urine. *Pestivirus* disease symptoms include enteric and wasting diseases, and in addition, pestiviruses can cross the placenta and infect the foetus, causing malformations and stillbirth [132]. The *Hepacivirus* genus includes the important human pathogen hepatitis C virus (HCV), which can cause chronic liver disease in humans [133]. Hepaciviruses also infect a number of animal hosts, such as horses, rodents and bats [134]. The forth genus are the Pegiviruses, which can infect a range of mammalian species including humans, chimpanzees and bats, but have no identified pathogenesis [135].

ZIKV was first isolated in 1947, from the serum of a rhesus monkey caged in the canopy of the Zika forest in Uganda. ZIKV was then isolated the following year from *A. africanus* mosquitoes, taken from the same forest, and the virus was named after the forest where it was first isolated [136]. Whilst there was no indication of disease, antibodies were detected in 8-20% of Ugandan residents tested [137] and in the next half century, ZIKV antibodies were detected in a range of countries across a narrow equatorial belt, running from Africa to Asia [138]. ZIKV infection was first reported to cause human disease in 1954 in Nigeria, where the virus caused an outbreak of jaundice [139], after which few cases of ZIKV associated disease were reported until 2007, when an outbreak occurred on Yap Island in the Federated States of Micronesia. In the outbreak an estimated 73% of the islands population (aged 3 years or older) were

infected with ZIKV [140] and the strain was identified as originating from South East Asia [141].

The next major epidemic occurred in 2013, on the islands of French Polynesia in the pacific ocean [142]. The most common symptom of ZIKV disease was an influenza-like illness however, during this outbreak, ZIKV infection was associated with higher prevalence Guillain-Barré syndrome in the population [143]. Due to the similarity of ZIKV disease symptoms to other endemic arboviruses (such as DENV and CHIKV) and the high level of antibody cross reactivity with other flaviviruses in diagnostic tests, it is difficult to determine the global prevalence of ZIKV [143]. The most recent ZIKV outbreak occurred in Brazil, where the first case of autochthonous transmission of the Asian clade was recorded in 2015 [144]. During this outbreak ZIKV infection was linked with higher rates of microcephaly seen in affected areas and viral RNA was isolated from the amniotic fluid of two women [145]. A retrospective study also found that the microcephaly was higher than average incidence in the population during French Polynesian outbreak [146].

There are two major lineages of ZIKV: the African lineage and the Asian lineage. Sequence analysis of viral isolates has revealed that the virus originated in Africa and evolved in two distinct groups in Uganda and Nigeria, with the Uganda strain considered to be the classical ZIKV strain. Despite having origins in Africa, approximately 97% of the published ZIKV strains belong to the Asian lineage [147]. The Asian lineage is thought to have originated in Malaysia [148] and has high sequence homology with the strains of ZIKV circulating in the Americas [149]. Amino acid differences between the two strains are present in the non-structural proteins NS1, NS2B and NS5, thought large phenotypic differences between the two lineages has not been observed [150]. Asian strains infect and replicate less efficiently in human neuroprogenitor cells compared with African strains, which may indicates links to the prolonged infection in brain and placental tissues seen in more recent outbreaks [150,151]. Interestingly, of the published ZIKV genomes, African strains have only been isolated from mosquitoes and non-human primates, whilst Asian strains have isolated from humans and mosquitoes [147].

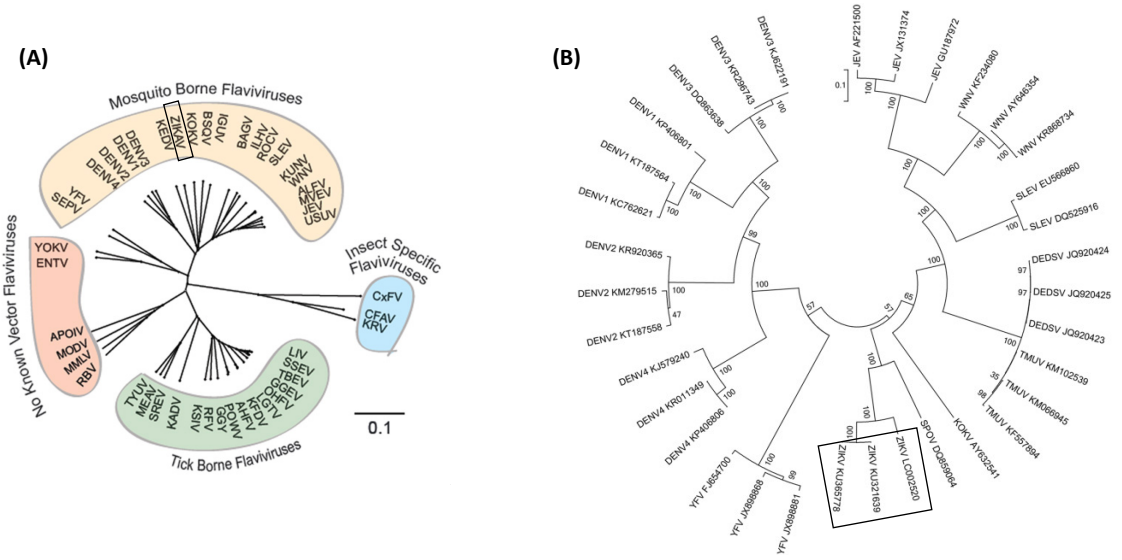


Figure 1.7 Phylogenetic trees of flavivirus species.

A) *The flavivirus family is divided into four sub-groups based on the disease vector. Adapted from Gamarnik et al. 2016* [152]. ***B)*** *Analysis of 14 flavivirus complete genome sequences by MegAlign. ZIKV shows the highest identity with Spondweni virus (SPOV), with 68.6% identity in nucleotide sequence and 75% identity in amino acid sequence. JEV, WNV, DENV and St. Louis encephalitis virus had 58%-60% nucleotide sequence identity with ZIKV and 55%-58% amino acid sequence identity. Adapted from Qing et al. 2016* [153].

1.3.2 ZIKV transmission

Similar to CHIKV, ZIKV is maintained in both sylvatic and urban cycles of transmission (Fig. 1.8). In Africa, the virus is likely maintained in non-human primates, with a number of non-human primates species showing serological evidence of ZIKV exposure [154]. The virus has also been detected in a number of arboreal *Aedes* species, including *A. furcifer* and *A. vittatus*. ZIKV positive *A. furcifer* and *A. vittatus* mosquitoes have been found within human settlements, suggesting that these species are involved in enzootic transmission to human hosts [155]. Urban transmission of ZIKV has been linked with *A. aegypti* and *A. albopictus* [156,157], although other species, such as *A. polynesiensis* and *A. Hensilli* were thought to involved in the French Polynesian [158] and Yap island [159] outbreaks respectively. Arbovirus disease transmission may also be linked to lifestyle and level of economic development. For example, despite the presence of *A. aegypti* in the south-eastern United States, there have been relatively few cases of autochthonous transmission of DENV (a closely related flavivirus). This has been attributed to factors such as abundance of air conditioning in buildings (leading to less open windows), fewer people per house and the increased average distance between houses, all of contribute to lower levels of contact between humans and infected mosquitoes [160].

The main method of ZIKV transmission is *via* mosquito vectors, however significant ZIKV transmission has also been noted via other routes. Sexual transmission was documented in 2008 when an infected scientist returned from Senegal and transmitted the virus to his wife [161]. ZIKV RNA has been detected in patient semen up to 188 days after the onset of symptoms [162], though in 95% of patients viral RNA was cleared after 3 months [163]. Guidelines state that infected individuals should use protection or refrain from sexual contact for 6 months after ZIKV exposure [162,163]. In addition, sexual transmission of ZIKV can occur from asymptomatic patients [164] and the level of sexual ZIKV transmission in endemic areas is difficult to assess, as the population is also exposed to mosquitoes, making it challenging to differentiate between mosquito acquired and sexually acquired cases. ZIKV also has the potential to be transmitted by blood transfusion, as 3% of 1,505 asymptomatic blood donors tested positive for ZIKV by RT-PCR during the French Polynesia outbreak, though again the prevalence of this is difficult to determine in endemic areas [165,166]. Materno-foetal ZIKV transmission was confirmed by detection of ZIKV RNA and anti-ZIKV IgM antibodies in the amniotic fluid of pregnant women, who had microcephalic foetuses during the Brazilian outbreak [167]. ZIKV RNA has also been recovered from foetal brain tissue and virus particles have been visualised within infected foetal brain tissue by electron microscopy [168], indicating that the virus had passed from infected mothers into the womb.

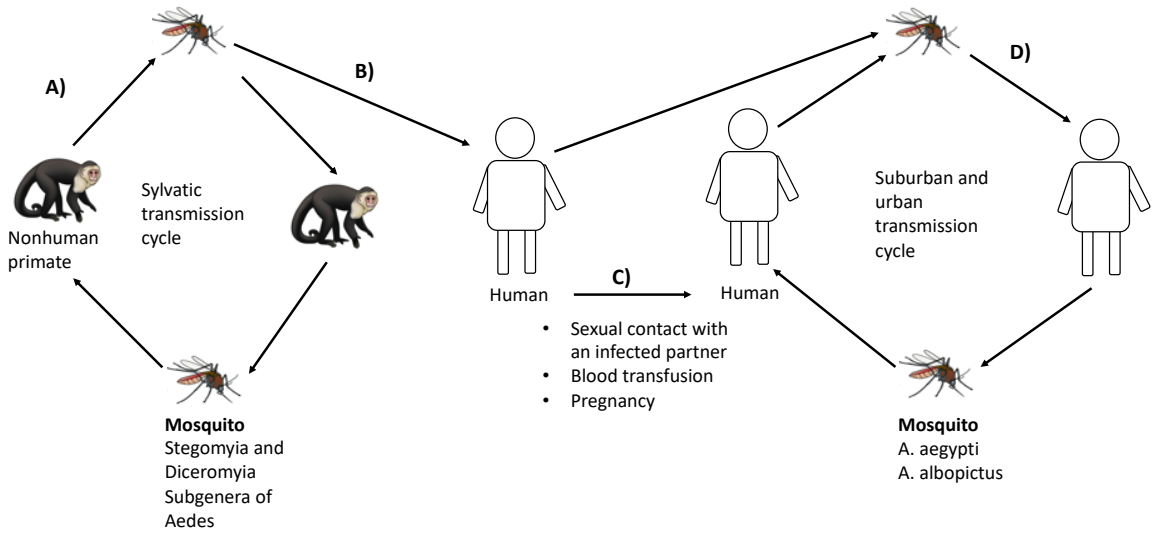


Figure 1.8 ZIKV transmission cycle.
A) *In forest environments, ZIKV circulates in a sylvatic transmission cycle between non-human primates and Stegomyia and Diceromyia Aedes mosquitos.* ***B)*** *Sporadic transmission can occur between mosquitos and humans.* ***C)*** *Transmission can occur between humans, via sexual contact, blood transfusion and pregnancy.* ***D)*** *In urban and sub-urban environments, ZIKV is mainly transmitted between mosquitos and humans. Adapted from Kurscheidt et al. 2019* [169].

1.3.3 ZIKV symptoms

ZIKV infection can cause a variety of symptoms in humans, which range from mild to severe and an estimated 80% of infected individuals will be asymptomatic. The remaining 20% will experience a febrile illness, where patients can experience fever, headache malaise, macular or papular rash, arthritis, arthraligia or nonpurulent conjunctivitis [140,170]. Hematospermia may occur in males [161]. In symptomatic patients, an estimated 50% of individuals will experience symptoms within 1 week and 99% will experience symptoms within 2 weeks [171]. ZIKV was initially considered to be a mild illness. However, in recent outbreaks more severe symptoms have been documented, conceivably because the higher number of infections has led to more cases of rare disease outcomes [172].

During the French Polynesian outbreak, ZIKV was linked with neurological symptoms, due to increased incidence of Guillain-Barré syndrome (GBS) [143]. GBS is an autoimmune disease which affects the peripheral nerves and causes progressive paralysis [173]. GBS carries a death rate of 5% and causes significant disability in up to 20% of patients [174]. The estimated prevalence of ZIKV-associated GBS is 1.23% [173], compared with the average incidence of 0.4 to 1.7 per 100,000 population [175]. There are multiple potential mechanisms for ZIKV-associated GBS. Viral antigens may mimic host proteins, leading to an immune response against the host [172]. Alternatively, prior or concurrent arbovirus infections, which exacerbate the immune response, may cause damage to peripheral nerves by the immunopathogenic process [174].

ZIKV infection is also associated with microcephaly, a neurological condition where the brain is underdeveloped, resulting in reduced brain volume and motor and/or intellectual disabilities. Microcephaly is defined as head circumference of more than two standard deviations below the mean for sex, age and ethnicity, with severe microcephaly defined as more than three standard deviations below the mean. Microcephaly may be diagnosed at birth (primary microcephaly) or postnatally (secondary microcephaly) where the child has a normal head circumference at birth, but then develops microcephaly at a later point [176]. Primary microcephaly is caused by a decrease in the number of neurones generated by neurogenesis and secondary microcephaly is thought to be caused by a decrease in synaptic connections and dendritic processes [177]. ZIKV was linked to microcephaly during the 2015 outbreak in Brazil, when the prevalence of recorded microcephaly cases increase from 1-2 cases per 10,000 live births to approximately 20 cases per 10,000 live births [145]. Ultrasound images of foetuses within infected mothers show calcification within the brain, which is indicative of intrauterine infection. Other presentations included severe damage to the cerebellum, brain stem and thalami [178]. In foetal brain tissue, ZIKV particles were visualised by negative staining electron microscopy and viral RNA was detected using RT-PCR, indicating that the virus actively replicates in the brain. No virus or pathological changes were observed in other foetal organs outside of the brain, indicating a strong neurotropism in foetuses [168].

1.3.4 ZIKV genome organisation and replication

ZIKV has a positive sense single stranded RNA genome consisting of a single open reading frame (ORF) encoding for 10 proteins, with highly structured non-coding UTRs at the 5′ and 3′ ends [152] (Fig. 1.9). The ORF encodes an approximately 3400 amino acid polyprotein precursor, containing three structural proteins (C, prM and E), which form the virus particle; and 7 non-structural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5) which are involved in the formation of replication complexes, particle assembly and immune evasion responses [179]. The ZIKV 5′ UTR is around 100 nts in length and has a type-1 5′ cap (m7GpppAmpN1) structure, which is essential for translation of viral genes by the host machinery and increases endonuclease resistance [180]. The 5′ UTR and adjacent capsid coding region of the ZIKV genome contains RNA structures. Despite previous characterisation of some RNA structures within the first 200 nts of the ZIKV genome, such as stem loop A, this region remains understudied and may contain additional RREs which are important for ZIKV replication [152]. The ZIKV 3′ UTR is around 430 nts in length and the 3′ end of the genome lacks a poly(A) tail. During initiation of translation cellular mRNAs circularise via interaction of the poly(A) binding protein (PABP) to the 5′ and 3′ ends of the mRNA. ZIKV genomes lack a poly(A) tail and evidence suggests that PABP binds RNA elements within the 3′ UTR, in order to enhance translation efficiency. The interaction of purified PABP with [α-^{32}P] labelled DENV 3′ UTR RNA was investigated using electromobility shift assays by native PAGE, where a greater shift was observed at higher concentrations of 3′ UTR RNA. Incubation of PABP with extreme 3′ end of the UTR alone showed no binding, indicating that PABP interactions occur in the upstream region of the 3′ UTR [181,182]. During RNA replication, complimentary conserved sequence elements at the 5′ and 3′ ends of the genome stabilise genome cyclisation [183], which is discussed in greater detail in section 1.3.5.2.

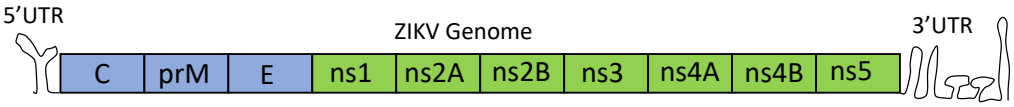


Figure 1.9 ZIKV genome organization.
The coding region expresses a single polyprotein which is processed into individual proteins. The structural proteins (marked in blue) form the virus particle, whilst the non-structural proteins (marked in green) are responsible for replicating the viral genome. The coding region is flanked by 5' and 3' untranslated region, which are highly structured.

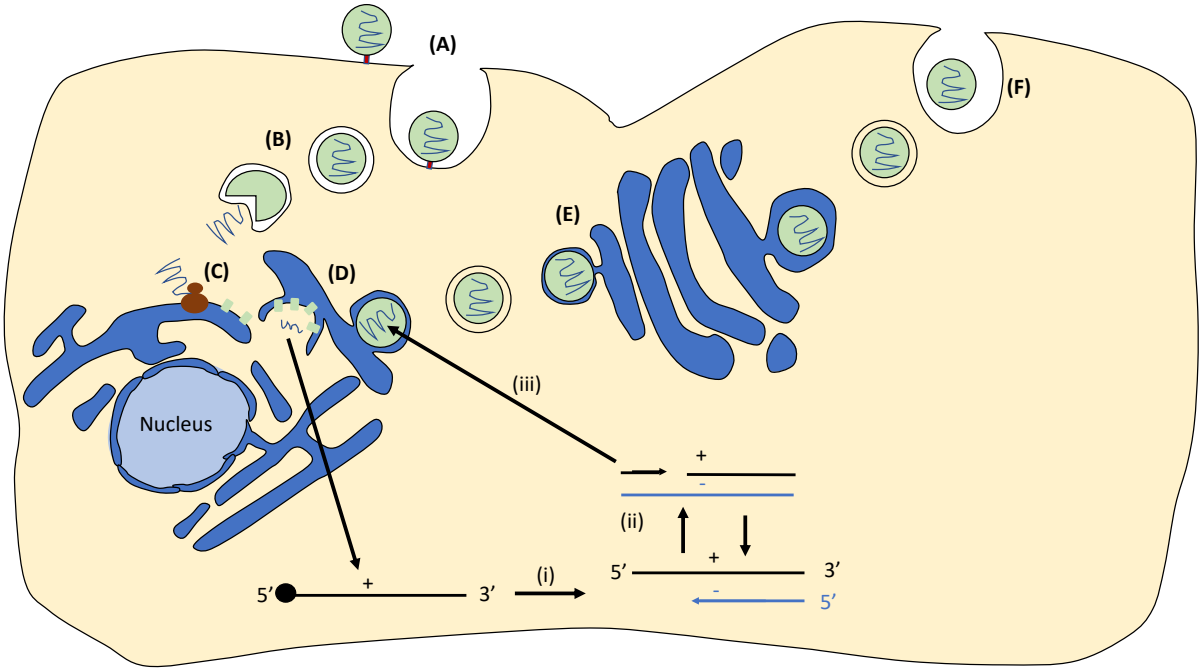


Figure 1.10 ZIKV lifecycle summary.
A) *ZIKV attaches to the cell surface via surface receptors and is internalized by endocytosis.* ***B)*** *Endosome maturation causes a more acidic pH, resulting in virion uncoating and genome release.* ***C)*** *Viral proteins are translated at the ER surface.* ***D)*** *Viral non-structural proteins replicate the viral genome in membrane bound complexes. (i) negative sense RNA is synthesised. (ii) Positive sense RNA is sythesised from negative RNA templates. Some new positive sense templates are used as templates for more negative sense templates. (iii) Positive sense RNA is packaged into immature virus particles.* ***E)*** *Furin cleavage of prM in the Golgi results in virus particle maturation.* ***F)*** *Mature virus particles are released. + denotes positive sense genomes, - denotes negative sense genomes.*

Following infection of the human body, ZIKV targets a variety of tissues and cells. Studies of pregnant mice, infected subcutaneously with ZIKV, identified high levels of infectious virus (detected by plaque assay) and high levels of ZIKV RNA by RT-PCR within the placenta and foetus head. High levels of ZIKV were also detected in the blood spleen and brain. Levels of

ZIKV RNA were 1,000-fold greater in the placenta than in the maternal serum, indicating that ZIKV preferentially replicates in the placenta. RNA fluorescence *in situ* hybridization in histopathological samples of mice placentas identified that ZIKV infects trophoblast cells [184]. Cell culture-based studies have shown that a number of human trophoblast derived cells lines are permissive to infection using immunofluorescence. Human brain microvascular endothelial cells were shown to be permissive to infection by RT-PCR however, primary human trophoblasts were resistant to infection, with only low levels of viral RNA detected. The authors attribute this resistance to the constitutive release of type III interferon by these cells [185]. ZIKV RNA has been detected by RT-PCR in the brain, cerebrospinal fluid, urine, and saliva of rhesus macaques [186] and in pigtail macaque eyes [187]. In addition, ZIKV infection reduces the size of the testes in mice and high levels of ZIKV RNA were detected in the male reproductive tract by RT-PCR [188].

The ZIKV lifecycle is summarised in Fig. 1.10. The ZIKV E protein is present on the virus particle surface [189] and attachment of ZIKV to the surface of host cells has been linked with binding to glycosaminoglycans (GAGs). ZIKV E protein has been shown to bind to placental GAGs using surface plasmon resonance [190]. A range of host receptors may contribute to virus entry, for example removal of sialic acid from the surface of Vero cells by incubation with neuraminidase caused a reduction in viral titre [191]. In addition, pre-incubation with anti-Axl antibodies caused impaired viral replication in male reproductive Sertoli cells, as determined by plaque assay [192] and human umbilical vein endothelial cells, as analysed by flow cell cytometry [193]. ZIKV enters host cells via clatherin mediated endocytosis. Treatment of host cells with either chlorpromazine (which prevents assembly of clatherin associated pits) or pitstop2 impaired virus replicon, as determined by RT-PCR [194].

Receptor bound virus particles are internalised in endosomes which mature into late endosomes. Changes in pH are proposed to trigger rearrangement of the E protein, exposing fusion peptides which carries out fusion of the viral and endosomal membranes. The viral RNA is then released into the cytoplasm via a fusion pore [195]. ZIKV infection requires acidic endosomal pH, as treatment of human glioblastoma cells with NH_4Cl (which raises pH of intracellular acidic vesicles) or chroloquine (an inhibitor of endosomal acidification) resulted in reduction in viral titre/ viral RNA detected by RT-PCR [194]. DENV E protein trimer insertion into liposomal membranes has been visualised using electron microscopy and the structure of the E protein trimer has been solved by x-ray crystallography. DENV E protein was expressed in *Drosophila melanogaster* Schneider 2 cells and purified from cell supernatant using immunoaffinity chromatography. Purified E protein was then combined with liposomes, which resulted in the formation of E protein trimers. Trimers were eluted and further purified by gel filtration before, being used to generate crystals. The trimer structure contains fusion loops, which are thought to be inserted into the endosomal membrane [195]. WNV has been shown to fuse with liposomes at pH 6.3, but not at pH 7.4, indicating that pH is essential for membrane fusion. Membrane fusion was measured using a pyrene-labelled WNV, whereby membrane fusion results in dilution of pyrene-labelled phospholipids into the liposome, resulting in a decrease in fluorescent signal [196].

Following release into the cytoplasm the positive-sense genomic RNA is translated as a polyprotein precursor, which is subsequently cleaved into the viral proteins. Polyprotein processing involves both host and viral proteins. For example, NS1/NS2A cleavage occurs

when NS1/NS2A is *in vitro* translated in the presence of dog pancreas microsomes (which contain membrane organelles similar to the ER and signal peptidases) and visualised by SDS-PAGE. The authors therefore propose that NS1/NS2A cleavage is mediated by a signal peptidase contained in the ER [197]. Cleavage of the NS4B/NS5 precursor has been demonstrated by the NS2A-NS3 complex *in vitro* (discussed further in section 1.3.5.6) [198].

Viral RNA replication occurs in membrane compartments, formed by NS4A integration into the ER membrane and remodelling the ER into virus replication vesicles (further discussed in section 1.3.5.7) [199]. Immunofluorescence microscopy of DENV infected cells indicates that dsRNA co-localises with viral proteins such as NS2B in the perinuclear space, indicating the site of viral RNA replication. Further analysis using electron microscopy revealed a complex set of membrane structure in DENV infected cells, which is derived from the ER, as determined by immuno-gold EM using the ER marker calnexin and immunofluorescence microscopy, using the smooth ER maker Syntaxin17. In addition, immuno-gold EM identified both viral protein and RNA in the ER derived membranes. Electron tomography analysis indicates that the induced membranes structures may form a large interconnected network of membranes [200]. RNA replication inside membrane structures likely functions as a means of bringing the components of the replication complex into close proximity, in addition to preventing recognition by the immune system. Positive-sense genomic RNA acts as a template for synthesis of intermediate negative-sense RNA, which in turn acts as a template for positive-sense RNA synthesis. New positive-sense strands are either packaged, enter a new cycle of RNA replication or act as mRNA for further translation of viral proteins. RNA replication is dependent on a number of viral RNA and protein factors, which are discussed in section 1.3.5.

Immature, non-infectious virus particles are proposed to form in the ER. Immunofluorescence analysis of TBEV infected cells confirmed that both E and prM localise to the ER, whilst electron microscopy experiments revealed that virus like-particles were present in the rough and smooth ER [201]. In order for virus particles to mature and become infectious, prM must be cleaved into M. Cleavage of prM has been linked with the cellular protein furin. *In vitro* reactions, where immature TBEV was incubated with recombinant bovine furin indicate that furin is able to cleave prM, as the cleavage products were visualised by immunoblotting. Furin cleavage was shown to be pH dependent, as cleavage was inefficient at higher pH values. In furin deficient human colon adenocarcinoma derived epithelial cells (LoVo), prM was not cleaved when cell lysates were analysed by immunoblotting, indicating that furin is the cellular protein responsible for cleavage. In addition, immature TBEV particle can be generated by treatment of primary chicken embryo cells with 20 mM ammonium chloride, which raises the pH of intracellular compartments, thereby preventing prM cleavage. Since furin is highly expressed in the Golgi, flaviviruses are proposed to be trafficked through the Golgi [202]. After assembly, flavivirus particles exit the cell via the plasma membrane. Virus particles have been observed to accumulate at the plasma membrane using electron microscopy and were shown to be released from the plasma membrane by small-vesicle exocytosis. The authors suggest that individual virus particles are released in small exocytosis vesicles [203].

1.3.5 Flavivirus RNA replication

1.3.5.1 Flavivirus RNA structures

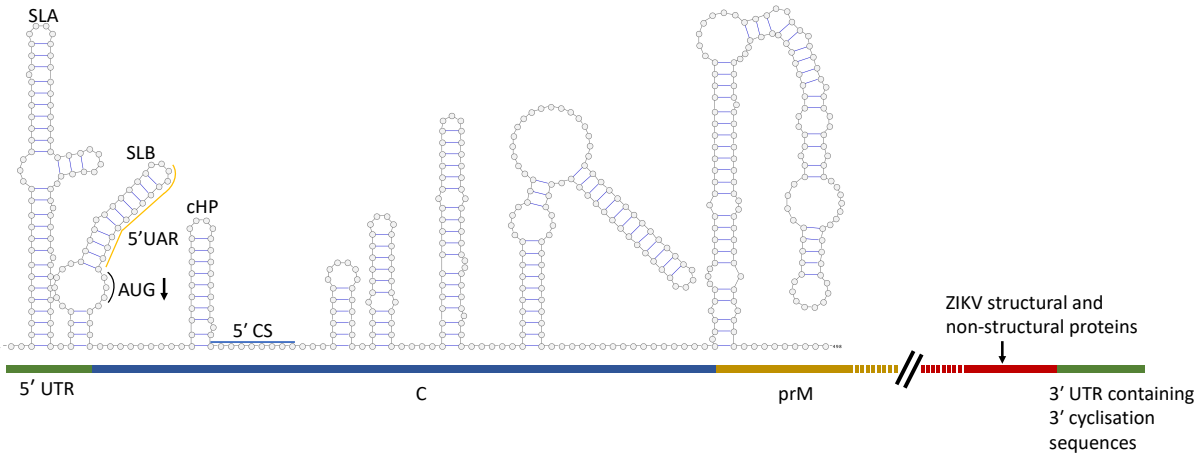


Figure 1.11 Schematic of ZIKV RNA structures in the 5′ UTR and adjacent coding region
RREs SLA, SLB and cHP have been characterized in related flaviviruses and are known to function in virus replication. Downstream RNA structures have not been previously characterized in ZIKV and were mapped in the current study. The 5′ UAR and 5′ CS bind to complementary sequences in the 3′ UTR during genome cyclisation. (Schematic is based on RNA SHAPE models shown in chapter 4).

RNA structures are present across the flaviviruses, indicating a high degree of selection pressure for them to be maintained [152,204]. Given the high level of conservation for several RNA structures amongst the mosquito born flaviviruses (e.g. DENV, WNV), conclusions about the function of RREs are often broadly applicable to closely related viruses. For example, conclusions drawn from studies into RREs in DENV have been subsequently shown to also apply for ZIKV and WNV [205–207]. At the current time DENV has been more widely studied than ZIKV, and given the high level of sequence conservation between the two viruses [131], much of the background for RNA structures in flaviviruses presented in this thesis has been derived from studies investigating DENV.

The 5′ UTR and adjacent coding region of flavivirus genome contains several highly conserved RNA structures which are essential for genome replication; consequently, deletions within this region result in impairment of virus replication [208,209] (Fig. 1.11). The first conserved structure is stem-loop A (SLA). Formed from approximately 70 nucleotides at the upstream end of the 5′ UTR, SLA acts as a promoter for viral RdRp binding and activation. SLA is discussed further in section (i) [152,210]. A second conserved structure named stem loop B (SLB), is found immediately downstream of SLA. Consisting of 30-50 nucleotides, SLB contains the 5′ upstream AUG region (UAR) sequence that is essential in forming long range RNA-RNA interactions between both ends of the virus genome during replication [211]. SLB is discussed further in section (ii). The third conserved flavivirus RNA structure is the capsid-coding region hairpin element (cHP), which is involved in start codon recognition and is required for efficient replication [205]. cHP is discussed further in section (iii).

The flavivirus 3′ UTR ranges between 400 and 700 nucleotides in length and consists of three domains, which have been confirmed by intracellular SHAPE mapping [152,212]. RNA structures present in domain I are involved in the production of sub-genomic flavivirus RNAs (sfRNA). During flavivirus infections, the viral genomic RNA may be digested by the host 5′ → 3′ exoribonucleases (XRNs). In ZIKV, two tight pseudoknot RNA structures (xrRNA1 and 2) in domain I cause the XRNs to stall, leaving the remaining 3′ RNA undigested [213]. When the 5′ UTR or viral proteins were deleted in a WNV replicon system (in addition to replication deficient replicons), sfRNA production was not affected, indicating that sfRNA generation is dependent on cellular factors rather than viral factors. In addition, siRNA depletion of XRN-1 resulted in reduced sfRNA production in WNV infected A549 cells, as detected by northern blot [214]. Northern blot analysis of total RNA from WNV infected cells confirmed the presence of two sfRNAs of different length, reflecting the positions of the two RNA structures. Mutation of xrRNA1 results in loss of sfRNA 1 detection and mutation of xrRNA2 results in loss of sfRNA 2, indicating that sfRNA production is dependent on these structures [215]. Northern blotting of ZIKV infected mammalian and insect cells lines confirmed sfRNA production and the crystal structure of xrRNA1 has been solved [213]. Mutations which prevent sfRNA production resulted in reduced viral titre in both Vero and C6/36 cells [214] and production of flavivirus sfRNAs has been linked to evasion of IFN responses. An sfRNA deficient WNV replicated more efficiently in mice deficient in major factors associated with the IFN response compared with WT mice as determined by RT-PCR of serum samples [216]. In *Aedes aegypti* mosquitoes, ZIKV deficient in sfRNA production was shown to have a lower viral titre in mosquito saliva than WT ZIKV [217].

Domain II in ZIKV contains two dumbell secondary structures which have been confirmed by SHAPE mapping and are predicted to form pseudoknot interactions [212,218]. SHAPE analysis of domain II showed that potential pseudoknot interactions are disrupted by mutagenesis, as N-methyl-nitroisatoic anhydride (NMIA) reactivity for nucleotides involved in pseudoknot formation was increased in the mutant RNA [218]. In DENV, mutation of the dumbbell structures led to reduced replication of a replicon, measured by RT-PCR, but did not impair translation (measured by luciferase expression), indicating a role in RNA replication [219]. Domain III contains two RNA structures, a functional short hairpin (sHP) and a terminal 3′ stem loop (3′ SL), as indicated by intracellular SHAPE reactivities [212]. sHP consists of a 5-nucleotide stem with a highly conserved 6 nucleotide apical loop. Mutations which disrupt sHP base pairing were shown using immunofluorescence to impair DENV replication in BHK cells. DENV isolates with reversion mutations, which are predicted to restore sHP structure, emerged after 12 days post infection [220]. Mutations which disrupt base pairing within the upper portion of the 3′ SL have been shown to be lethal to replication in DENV, as no infectious virus was detected in Rhesus monkey kidney epithelial cells at up to 20 days post electroporation. Mutations affecting the sequence of the upper portion of the 3′ SL impaired virus replication, but were not lethal for replication, indicating that the structure of the 3′ SL is more critical than the sequence [221]. In WNV, mutations within the base region of the 3′ SL were lethal for replication, as no virus was detectable by plaque assay after transfection of viral RNA [222]. In addition, the host protein translation elongation factor-1 alpha has been shown using electromobility shift assays to bind the 3′ SL, and the authors propose that this interaction may play a role in targeting viral RNA into intracellular membranes [223].

(i) SLA

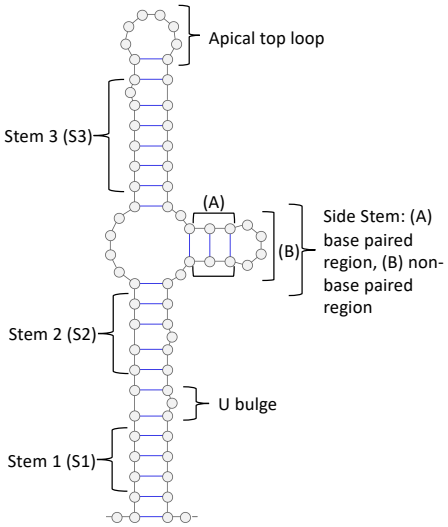


Figure 1.12 Labelling of SLA sub-regions
SLA binds the viral polymerase NS5, allowing RNA replication to initiate from the 3' end of the viral genome. All sub-regions have been shown to function in flavivirus replication, for example mutations which disrupt base pairing in the S1 and S2 regions impair virus replication.

The flavivirus RRE SLA, found at the 5' end of the 5' UTR, is described as having a Y-shape structure, with a heteroduplex base paired stem with an unpaired apical top loop and a short side stem [224]. SLA is highly structurally conserved between mosquito borne flaviviruses, such as DENV and WNV. The structure of SLA is inverted in tick-borne encephalitis virus, with the side stem present on the opposite side of the stem, and in the mosquito specific Aedes flavivirus the side stem is truncated [152]. SLA functions as a promoter of viral RNA synthesis, by acting as a binding site for the viral RdRp prior to initiation of RNA replication at the 3' end of the genome (discussed further in section 1.3.5.8) [207,224]. Much of the molecular detail for the function of SLA has been obtained from studies using DENV, which is closely related to ZIKV [131].

Several mutations within the unpaired apical top loop, base paired stem and short side stem regions of SLA were shown to inhibit RNA replication in a sub-genomic DENV replication, without affecting translation. Virus replication was measured using immunofluorescence and additionally for ZIKV, using a luciferase expressing replicon and qRT-PCR. Translation was measured using an NS5 knockout DENV replicon expressing firefly luciferase, which has a translation dependent luciferase signal at 10-20 hours post transfection but does not show a second, replication dependent, luciferase signal, which is seen in the WT replicon after 40 hours [207]. The base paired stem region can be further divided into three regions as shown in Fig. 1.12. Mutations which disrupt base pairing in DENV regions S1 and S2 were shown using immunofluorescence, over a period of 1-8 days post infection, to impair virus replication and compensatory mutations (which restore base pairing using alternative nucleotides) restored replication [207,224]. In ZIKV, disruption of the S2 region caused a significant reduction in virus replication, whilst complementary mutations restored replication, although

not to a WT level [206]. By contrast, mutations disrupting the structure of the S3 regions showed replication similar to WT DENV, although deletion of the S3 region abolishes viral replication. Deletion of the S3 region impaired viral polymerase activity, as measured by [α-^{32}P] GMP incorporation into RNA templates, indicating that the promoter activity of SLA was lost by this mutation, i.e. SLA can no longer bind to the polymerase [207]. In addition, mutation of the U bulge which separates the S1 and S2 regions showed impaired replication, when U nucleotides were changed to A, although the presence of only one U nucleotide was sufficient for replication [207].

Likewise, the need for the apical top loop and short side stem regions have been investigated. Mutations disrupting the structured region of the short side stem in DENV impaired virus replication, whilst compensatory mutations restored replication. In contrast, mutations changing the sequence of the non-base paired nucleotides of the side stem had no effect on replication. Deletion of the entire side stem abolishes viral replication and impaired SLA polymerase activity *in vitro*, as measured by [α-^{32}P] GMP incorporation into RNA templates [207,224]. Mutations which change the sequence of the 5 nucleotide SLA apical top loop impair virus replication [207,224]. Gamarnik et al. introduced a mutation which reduced the size of the apical top loop to 3 nucleotides, which impaired virus replication. However, by six days post transfection, a spontaneous insertion mutation occurred which restored the apical top loop from 3 nucleotides to 5 nucleotides, but with a non-WT sequence (CAGAU→AAGAC) and this had the effect of restoring replication [207]. Similarly in ZIKV, mutation of unpaired nucleotides in the apical top loop impaired virus replication [206]. Analysis of ZIKV long range RNA interactions identified that much of the SLA has covariance with a region of the ZIKV genome around 1100 nucleotides downstream, indicating that this interaction may play an as yet unknown role in the virus lifecycle [212].

Downstream of SLA is a region of nucleotides termed the oligo(U) track, which also plays a role in RNA replication. In DENV, deletion of the 3, 6 or 10 nucleotides from this sequence impaired replication, with a greater effect seen as more nucleotides are deleted. Insertion of addition nucleotides (6 or 10 nucleotides) resulted in increased levels of RNA synthesis. The authors suggest that a minimum of 10 residues must be present in this region for efficient RNA synthesis to occur [207].

(ii) SLB

Stem loop B contains the 5’ UAR sequence. During genome cyclisation, the 5’ UAR sequence hybridises with a complementary sequence in the 3’ genome region (3’ UAR) [211]. In addition, the proposed 5’-UAR-flanking stem (UFS) is formed by the nucleotides found within the base region of SLB (genome cyclisation and the role of the UFS switch are discussed in detail in section 1.3.5.2) [225]. Studies investigating the requirement of the 5’ RREs SLA and SLB for RNA replication revealed that templates containing folded SLB structures blocked RNA synthesis, regardless of the presence of an active SLA promoter. Templates lacking SLB or containing sequences in which SLB base pairing was disrupted were competent for RNA replication [226], indicating that SLB plays a role in regulating RNA replication by blocking RNA replication initiation.

(iii) cHP

The third 5′ flavivirus REE, cHP, is a short hairpin which has been shown to play a role in regulation start codon selection [227]. cHP is found 14-15 nucleotides downstream of the first AUG start codon at the terminal of the capsid coding region, a distance which corresponds to the footprint of a ribosome paused over a start codon. During translation initiation, the 43S ribosomal complex scans from the 5′ end of an RNA molecule until a start codon is encountered [228]. The ribosome complex is hypothesised to stall at cHP, thereby positioning the AUG-recognition site directly over the start codon, which has the effect of enhancing translation initiation from the first AUG site [229–231].

In order to investigate the role of cHP in translation initiation, Clyde and Harris performed mutagenesis to disrupt DENV cHP base pairing. This resulted in reduced translation from the first start codon (compared with a second downstream start codon), indicating that the unstable cHP structure was unable to stall the ribosome complex. Translation start codon selection was measured using a T7 promoter construct containing, the 5′ and 3′ UTRs and a C-FLAG fusion protein transfected into mammalian and insect cells. Immunoblotting was used to detect 1st and 2nd start codon translation initiation products, which have a detectable size difference, from cell lysates [227]. Stabilisation of the cHP structure by increasing the GC content resulted in greater levels of translation initiation from the first start codon, as presumably the more stable cHP is more effective at stalling the ribosome complex. The function of cHP is dependent on its position in the genome, as moving cHP upstream or downstream of the WT position results in inhibition of translation initiation from the first AUG. In addition, replacement of the WT cHP sequence with a heterologous sequence, which forms a similar hairpin, leads to translation initiation which is similar to WT, indicating that cHP ribosome stalling is sequence independent [227]. Mutations destabilising the cHP structure cause a reduction in viral titre and passaging of destabilised cHP viruses in BHK cells results in spontaneous compensatory mutations which restore cHP base pairing and restore WT levels of replication [205].

When cHP is disrupted, first start codon selection still occurs at around 40% of WT [227]. It is important to note that whilst disruption of cHP reduces translation from the first AUG, the overall level of translation is unaffected, due to the presence of a second downstream AUG. If the translation machinery bypasses the first AUG, translation is initiated from the second AUG. Mutation of either the first or second AUG sequence causes a reduction in viral titre, indicating that both AUGs are required for viral replication [227].

In DENV, translation of input RNA occurs within the first few hours of transfection, after which viral translation is reduced. A second peak of translation occurs by 48 hours, which is dependent on synthesis of new RNA transcripts. A lack of translation in the second peak indicates that RNA synthesis is defective. Using a DENV replicon, Harris et al. [205] demonstrated that disruption of cHP has no effect on input translation (0-8 hours post-transfection), but significantly reduces translation at later time points (48 and 72 hours). These results indicate that the cHP disrupted mutant initially translates at a WT level, but does not progress to the second peak of translation due to defects in RNA synthesis. In addition, in DENV and WNV replicons, less viral RNA was detected in cells transfected with cHP mutant RNAs and NS1 expression was reduced at both 18 and 22 hours post infection in

infectious clones [205]. The authors propose a model suggesting that during translation, cHP acts to direct start codon selection by enhancing initiation from the first AUG and that cHP also functions in RNA replication, forming part of the 5′ RNA complex which binds to the replication complex or functioning by stabilising the circular genome conformation [205].

1.3.5.2 Flavivirus genome cyclisation

Flavivirus genomes exist in both a linear conformation (in which the 5′ RRE SLB is present) and a circularised conformation (in which complementary sequences in the 5′ and 3′ of the genome hybridise and SLB is no longer formed) (Fig. 1.13). Mutant genomes which shift the equilibrium in favour of either conformation undergo reversion mutations, which restore the equilibrium of both linear and circularised genome conformation, indicating that both forms are required for efficient replication [220]. Genome cyclisation is proposed to have a critical role in viral RNA replication, at the stage of initiation of negative-sense RNA synthesis (Fig. 1.14). In the proposed model, the viral RdRp binds to the 5′ Flavivirus RRE SLA. In order to bring the polymerase into proximity with the 3′ initiation site, at the opposite end of the viral genome to SLA, the genome undergoes cyclisation via long-range RNA-RNA interactions between complementary sequence elements. The cyclised conformation allows the polymerase to access the 3′ end of the genome and therefore initiate minus strand synthesis [8,232]. The requirement for SLA binding and genome cyclisation has been proposed to have several advantages. These include improved RNA stability, and allowing discrimination between viral templates and host templates which populate the ER. Genome cyclisation may also allow the RNA replication machinery to sample the ends of the genome, meaning that only full length genomes act as templates [232,233]. It has been suggested that the linear to circular genome switch acts as a mechanism for controlling the levels of negative sense RNA produced, as significantly more positive sense genomes are produced during an infection. In addition, the two genome conformations may regulate the switch between virus replication and translation, with circular genomes acting as templates for replication and linear templates act as templates for translation [220].

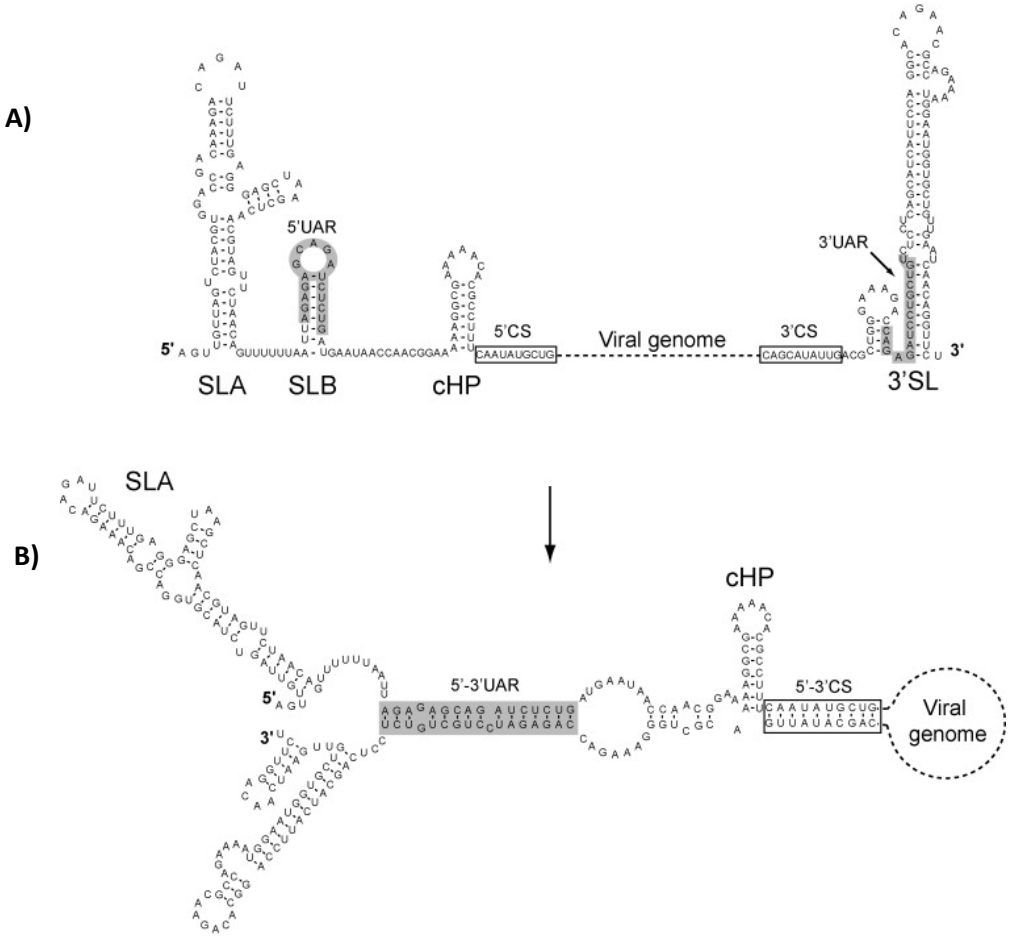


Figure 1.13 Model of flavivirus genome cyclisation.
A) *In the linear genome conformation, the RNA structure SLB is present and the 5' and 3' complementary sequences are not base paired.* ***B)*** *Genome cyclisation is mediated by complementary sequences found within the 5' and 3' regions of the genome, termed the upstream of AUG region (UAR) and cyclisation sequences (CS). In the circularized genome conformation, the 5' and 3' complementary sequences hybridise, resulting in the loss of SLB and reconfiguring the structures at the 3' end of the genome. Adapted from Villordo et al* [183].

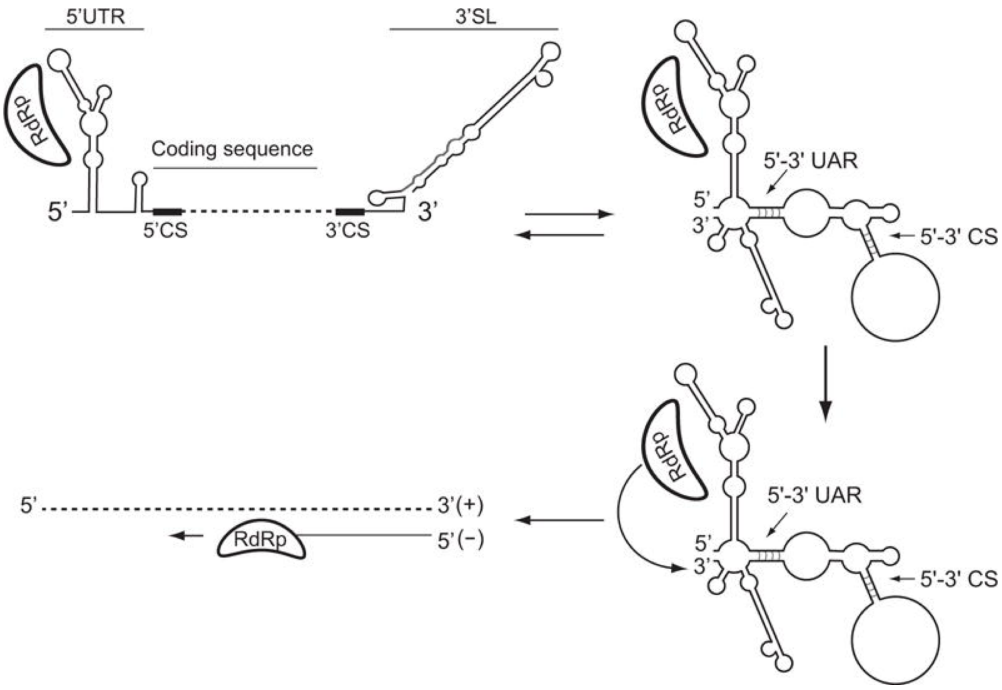


Figure 1.14 Proposed model for the initiation of negative sense viral RNA synthesis.
The viral polymerase (RdRp) binds to SLA at the 5′ end of the genome and the genome is circularised by hybridisation of the upstream of AUG region (UAR) and cyclisation sequences (CS). Genome cyclisation brings the bound polymerase into the proximity of the 3′ end of the genome, initiating RNA synthesis. Adapted from Villordo et al [183].

Complementary sequences which mediate genome cyclisation were first identified in the 5′ and 3′ ends of mosquito-borne flavivirus genomes in the 1980s [234] and these sequences were later shown to be essential for virus replication, without affecting translation [235,236]. These sequences, termed the 5′ and 3′ cyclisation sequences (5′ and 3′ CS) were demonstrated to be necessary, but not sufficient for genome cyclisation, which led to the identification of further cyclisation sequences in a region upstream of the AUG start codon [211,237]. Gamarnik et al. performed RNA binding assays using RNA molecules which represent the 5′ and 3′ ends of the DENV genome. Sequences within the 5′-3′ CS and the upstream of AUG region (UAR) were mutated, which prevented the formation of RNA complexes, indicating that both regions are required for genome cyclisation [211]. Additionally, disruption of UAR base pairing was demonstrated to impair RNA synthesis (but not translation) in a DENV replicon system and in the context of an infectious virus system, disruption of the 5′ or 3′ UAR resulted in spontaneous mutations which restored base pairing and rescued virus replication [238]. In ZIKV, mutation of the 5′ CS to prevent hybridisation with the 3′ CS caused inhibition of virus replication. Compensatory mutations, which reconstituted the 5′-3′ CS interaction with alternative sequences, restored replication but to a lower level than WT [206]. Sweeney et al. proposed that genome cyclisation in ZIKV and DENV blocks initiation of viral translation, as the structures formed by hybridisation of the cyclisation sequences blocks ribosome scanning of the RNA [239]. This therefore indicates that the two genome conformations play distinct roles in the virus lifecycle: with the circular conformation essential for RNA replication and the linear form essential for translation.

An additional factor proposed to play a role in the cyclisation process is the 5'-UAR-flanking stem (UFS). The UFS encompasses the base of the flavivirus RRE SLB and was shown to be base paired by *in vitro* SHAPE mapping [225]. Disruption of base pairing within the UFS by mutagenesis showed impaired replication in a DENV replicon and an infectious ZIKV system. Reconstitution of the UFS structure using alternate base pairing rescued replication, indicating that the UFS functions in a structure dependent, sequence independent manner. The UFS has also been proposed to play a role in RdRp recruitment to viral RNA, as RNA molecules in which UFS base pairing had been disrupted had reduced affinity for the NS5 protein in EMSA experiments. The authors propose a model in which the base paired UFS promotes binding of NS5 onto genomes which have a linear conformation. Once the RdRp binds, the genome folds into a circularised conformation, during which the UFS structure unfolds, decreasing the affinity of NS5 to the 5' end of the genome and therefore facilitating transfer of NS5 to the 3' end of the genome, where negative sense RNA replication is initiated [225].

1.3.5.3 RNA structures in other members of the Flaviviridae family

RNA structures at the 5' end of the flavivirus genera genome differ from those found in other viruses in the Flaviviridae family. The genomes of Hepaciviruses, such as HCV, contain a highly structure region within the 5' UTR of the genome (and around 30 nts into the coding region), known as an internal ribosome entry site (IRES) [240]. Using cell free translation assays using a dicistronic RNA, containing a capped methylated region expressing a chloramphenicol acetyltransferase reporter protein at the 5' region and a second HCV RNA region. Both the capped reporter region and the HCV region were translated when the translation experiment was performed using rabbit reticulocyte lysate. In contrast, using lysates from human cervical cancer cells which had been infected with coxsackievirus B1 (which suppresses cap-dependent translation), only the HCV region was translated, indicating that HCV contains an IRES [241].

Chemical and enzymatic probing of the HCV IRES sequence revealed the structure of the IRES, which consists of 2 major domains. Domain II contains a single large RNA stem loop structure, with two subdomains, IIa representing the base of the stem and IIb representing the top of the stem. Domain III contains branching hairpin stem loops, with the base of domain III containing a 4-way junction, which contains a pseudoknot interaction (IIIf) and a small stem loop (IIIe). The middle part of domain III consists of a 3-way junction with a conserved stem loop (IIId). The upper part of domain III contains a 4-way junction (IIIabc) [242]. In addition, the structure of the HCV IRES has been confirmed using both x-ray crystallography [243] and NMR [244].

Members of Pestivirus genera, such as classical swine fever virus (CSFV), also contain an IRES in the 5' region of the genome, which is structurally similar to the HCV IRES [245,246]. Using HCV and CSFV, IRES mediated initiation of translation was shown to occur in the absence of eukaryotic initiation factors. HCV and CSFV IRES RNA sequences were shown to bind 40S ribosomal subunits by incubating short RNAs containing the IRES sequence in rabbit reticulocyte lysate in the presence of a nonhydrolyzable GTP analogue that causes 48S complex accumulation. The ribosome/ viral RNA complexes were then resolved using sucrose density gradient centrifugation. Both the HCV and CSFV RNA were able to form ribosome

complexes when eukaryotic initiation factors were omitted from the experiment, therefore indicating that IRES binding to the ribosome is not dependent on these factors [247].

In addition, the genomes of Pegiviruses, such as G. Barker viruses A and C (GBV-A, GBV-C) also contain IRES RNA element in the 5' region of the genome. The IRES sequences in GBV-A and GBV-C were investigated using bicistronic constructs, where the IRES sequences from GBV-A and GBV-C were inserted between a chloramphenicol acetyltransferase reporter and a luciferase gene. *In vitro* transcription and translation assays were used to determine translation of both products, whereby translation of the luciferase gene was used as an indication of IRES dependent translation. Luciferase activity was measurable when the bicistronic constructs contained the GBV-A and GBV-C IRES sequence in the sense orientation, but not when the IRES sequence was in the antisense orientation [248]. The authors also present secondary structure predictions for the RNA structures in 5' region of the genomes of Pegiviruses, based on a combination of thermodynamic models and phylogenetic analysis. The secondary model indicates the 5' regions of GBV-A and GBV-C contain 5 structural domains. Domain I contains a stem loop, consisting of a top loop region and a base region which contains a side stem. Domain II contains a single large stem loop structure. Domain III contains two small stem loop structures, the first of which is predicted to form a pseudoknot interaction with the nucleotides immediately upstream of itself. Domain IV contains a large RNA stem loop with a base region, containing a short side stem, and a top region containing a non-base paired bulge region. Domain V contains a short RNA stem loop with top and side stem of roughly equal length [248].

In addition to the 5' region RNA structures, an RNA element within the RNA dependent RNA polymerase coding region of HCV has been shown to be required for RNA replication. Termed 5BSL3 the structure is formed by 171 nt of the C-terminus of the RNA dependent RNA polymerase coding region. Structure probing analysis using RNase digestion revealed that 5BSL3 has a cruciform structure. Mutations were introduced to disrupt folding of one of the cruciform arms, termed 5BLS3.2, which consists of an upper and lower helix, a base terminal loop and a side loop. Mutations within 5BLS3.2 resulted in a lack of HCV growth when 5BLS3.2 mutant RNA was transfected into Huh7.5 cells and inhibited replication in a HCV sub-genomic replicon system, indicating that 5BLS3.2 plays a role in HCV genome replication [249].

The 3' region of HCV contains an approximately 40 nt variable region, a poly (U/UC) tract and a 98 nt X tail. The variable region contains two RNA stem loops, and deletion of either one or both of these stem loops was shown to impair replication in a HCV replicon assay [250]. Using replicon assays, the length of the poly (U/UC) tract required for efficient replication was investigated. Replicons containing polyuridine tracts of 46 or 26 nt replicated as efficiently as those containing the parental 83 nt a poly (U/UC) tract [250]. The structure of the 98 nt X tail was determined using RNase digestion experiments. The X tail consists of three RNA stem loops, with the 3' most RNA stem loop being elongated compared with stem loops 1 and 2 [251]. Deletion of the entire X tail region or any of the three stem loops individual resulted in a lack of replication in luciferase replicon assays [250]. Taken together, these results indicate that RNA elements within the 3' region of HCV are important for efficient RNA genome replication.

1.3.5.4 Non-structural protein 1 (NS1)

The Flavivirus protein NS1 contains three domains and is found as both a dimer and as a secreted hexamer [252]. Using an antigen capture enzyme-linked immunosorbent assay, DENV NS1 was shown to be present at high levels (up to 15 µg/ml) in patients sera [253]. The crystal structure for both dimeric and hexameric NS1 has been solved, with the dimeric form containing a hydrophobic protrusion which may mediate interactions with membranes. Cryo-electron microscopy analysis of secreted NS1 obtained from DENV infected Vero cell extracellular media confirmed the hexameric structure and the presence of triglyceride lipids was confirmed by treatment of NS1 with organic solvents, followed by thin layer chromatography and NMR [254].

Incubation of WNV NS1 with large liposomes resulted in formation of small, NS1 coated lipid-protein nano-particles when visualised using negative-stain electron microscopy, indicating that NS1 has the ability to remodel membranes. This suggests that NS1 plays a role in remodelling host membranes to form virus replication complexes [255]. Mutations within YFV NS1 resulted in a decrease in released viral titre, with mutations in the N-terminal region having a greater effect, indicating that NS1 plays a role in virus replication [256]. In DENV, NS1 has been shown using confocal and cryo-immuno-gold electron microscopy to co-localise with dsRNA [257] and in Kunjin virus, immunofluorescence indicates that NS1 localises with dsRNA, NS2B and NS3 in the perinuclear space, indicating that NS1 forms part of the replication complex [258].

NS1 has been shown by co-immunoprecipitation to interact with the complement regulatory protein factor H, using supernatants from BHK cells which stably expresses the WNV NS1-NS5. Binding of NS1 to factor H in solution results in degradation of the complement protein C3b, as shown by western blot. In addition, cell surface associated NS1 was shown using flow cryometry to inhibit complement deposition at the cell surface. The authors therefore suggest that NS1 prevents membrane attack complex formation at the cell surface [259].

1.3.5.5 Non-structural protein 2A (NS2A)

NS2A is a small hydrophobic protein which associates with the ER and has multiple roles in viral replication. Bioinformatics analysis suggest that ZIKV NS2A has seven transmembrane segments and using immunofluorescence microscopy has been shown to co-localise with the ER membrane protein calnexin in the perinuclear region [260]. Kunjin virus NS2A has been shown to co-localise with dsRNA, NS1, NS3 and NS5 in the perinuclear region by immunofluorescent microscopy [261]. Mutagenesis of ZIKV NS2A resulted in a 10-fold decrease in viral titre and caused inhibition of RNA synthesis in a replicon system. In a sperate set of mutations, the sub genomic replicon was unaffected however, no virus was detectable by plaque assay, indicating that NS2A may play a role in virus particle assembly [260]. In DENV, mutagenesis blocking the N-terminal cleavage of NS2A impaired RNA synthesis in a GFP expressing replicon [262].

The NS2A C-terminus is thought to be involved in RNA replication, as C-terminal alanine substitution mutations impaired DENV replication when measured at 7 days post infection

using immunofluorescence. Another set of mutations resulted in WT replicon replication, but with low viral titre by transient *trans*-packaging experiments, where the DENV structural proteins were supplied in trans, indicating a potential role in virus particle assembly [263]. In addition, NS2A may play a role in inhibition of IFN signalling, as transfection of an NS2A expressing plasmid into cells where an IFN response had been triggered enhanced replication of Green fluorescent protein (GFP)-tagged Newcastle disease virus [264].

1.3.5.6 Non-structural protein 2B (NS2B)

NS2B acts as the co-factor for NS3 to form the NS2B-NS3 protease complex, which cleaves the viral polyprotein. The crystal structure for the protease complex has been solved [265]. Protease activity has been demonstrated *in vitro* using a purified [^{35}S] Methionine-labelled NS4B-NS5 substrate, where reactions were separated using SDS-PAGE. NS2B-NS3 demonstrated protease activity, whereas NS3 alone and NS2B/NS3 mutant precursor were unable to cleave the substrate [198]. Crystallography based studies of the NS2B-NS3 in complex with an aldehyde peptide indicate that NS2B may form part of the protease active site [266].

1.3.5.7 Non-structural protein 3 (NS3)

The NS3 protein carries out enzymatic activities which are essential for flavivirus replication. The N-terminal domain is a serine protease which, when acting with the co-factor NS2B, form a protease which processes the viral polyprotein and the C-terminus contains an ATPase/ helicase domain which unwinds the viral RNA during RNA synthesis [267,268]. The crystal structure for both the helicase and protease domain of DENV has been published [267]. ZIKV helicase activity is dependent on ATP hydrolysis and ATPase activity was determined using purified full length NS3 via a Malachite green colourimeter assay, which can detect free organic phosphates released by ATP hydrolysis. Helicase activity was measured *in vitro* using a FRET based assay, where fluorescence is quenched when nucleic acid strands are annealed and fluorescence increases when strands are separated. The rate of nucleic acid unwinding was increased in the presence of NS5. Additionally, mutations within the ATP binding region impaired ZIKV replication in a luciferase expressing replicon and using mutant infectious virus when quantified using RT-PCR [268].

NS3 also plays a role in RNA capping, measured using transfer of [α-^{32}P] GMP to an RNA substrate, visualised using PAGE. Radiolabelled GMP was not transferred to the RNA substrate in the absence of either NS3 or NS5, indicating that both proteins are required for de novo cap synthesis. The authors propose the following model: During viral genome synthesis the RNA triphosphatase activity of NS3 removes the γ-phosphate from the RNA and the modified RNA then interacts with the NS5 methyltransferase-RNA guanylyltransferase (MTase-GTase). Association of NS3 with NS5 increases GTase activity, resulting in transfer of [α-^{32}P] GMP to the modified RNA and creating a capped transcript which is methylated by the NS5 MTase [269,270].

1.3.5.8 Non-structural protein 4A/4B (NS4A/NS4B)

The C-terminal domain of NS4A is hydrophobic and is referred to as the 2K fragment. Using a construct consisting of 2K-NS4B, cell free translation assays showed that cleavage of 2K is dependent on the presence of microsomal membranes. Processing of NS4A-2K-NS4B was not observed in cell free translation assays, but was observed when NS4A-2K-NS4B was expressed in addition to NS2B/NS3 using in vivo transient expression assays. The authors suggest that NS4A-2K-NS4B is first cleaved at the NS4A-2K junction, by the NS2B/NS3 protease, and then the 2K fragment is cleaved from NS4B by signal peptidases [271]. The 2K fragment acts as a signal for translocation of NSB4 to the ER. Immunofluorescence microscopy of HEK 293T cells expressing 2K-NS4B shows co-localisation of 2K-NS4B and the ER marker Calreticulin [272].

NS4A is a transmembrane ER associated protein and is thought to induce membrane modifications which fold the ER membrane into replication vesicles [199,273]. NS4A has been shown using immunofluorescence to co-localise with dsRNA and other viral proteins in membrane bound complexes found in close proximity to the ER. Electron microscopy analysis of NS4A expression in Huh7/T7 cells revealed that NS4A induces cytoplasmic membrane rearrangement. The membrane remodelling activity may be regulated by the 2K fragment, as NS4A/2K expression did not induce membrane remodelling and was localised to the ER [199]. NS4A has been reported as having an N-terminal alpha helix in DENV, which has been shown using surface plasmon resonance (Biacore L1 chip) to bind membranes immobilised on a chip surface, whilst mutations within the alpha helix reduced membrane binding. NS4A alpha helix mutations were also shown to impair replication of a luciferase expressing DENV replicon [273]. NS4B is also a membrane associated protein which associates with ER derived membranes and has been shown to co-localise with dsRNA and NS3 using immunofluorescence microscopy of DENV infected cells [274].

NS4A/NS4B has been shown to reduce protein Akt phosphorylation, which supresses Akt-mTOR signalling, an essential pathway in neurogenesis from neural stem cells, determined by immunoblotting of ZIKV infected human foetal neuronal stem cell lysate. Expression of either NS4A or NS4B from a stably expressing vector resulted in reduced Akt phosphorylation at positions Thr308 and Ser473, whereas co-expression of NS4A/NS4B reduced mTOR phosphorylation at position Ser2448. In addition, expression of NS4A and NS4B in human foetal neuronal stem cells, increases microtubule-associated protein light chain 3 (LC3) I $\rightarrow$ II conversion, measured by immunoblotting (LC3-II migrates further than LC3-I when separated by PAGE). Expression of NS4A/NS4B had a greater effect on LC3 conversion than either protein in isolation. LC3 conversion is a widely used marker of autophagy, indicating that NS4A-NS4B induce autophagy [275,276].

1.3.5.9 Non-structural protein 5 (NS5)

NS5 is the most conserved and largest of the ZIKV proteins [277]. The crystal structure of NS5 has been reported, consisting of an RNA-dependent RNA polymerase domain at the C-terminus and a methyltransferase domain at the N-terminus [278]. The methyltransferase activity of the N-terminal domain has been demonstrated in DENV by measuring Ado[*methyl*-^{3}H] Met addition to RNA substrates, using a filter binding assay. Transfer of the methyl group from AdoMet occurred in the presence of capped, but not uncapped RNA transcripts [279].

Additionally, NS5 plays a role in synthesis of the RNA cap structure. RNA GTases enzymes catalyse a two-step reaction, first forming a covalent enzyme-GMP intermediate. DENV, WNV and YFV NS5 were shown to form GMP-enzyme complexes when incubated with [α-^{32}P] GTP, by measuring levels of radiolabelled GMP when reactions were separated using SDS-PAGE. In the second step of the capping reaction, GMP is transferred to an acceptor RNA molecule. The authors incubated NS5 with 5'-diphosphate acceptor RNAs of 81 nts to determine if NS5 was capable of GMP transfer. In RNA acceptor molecules containing a 5' Guanosine diphosphate nucleotide, GMP transfer was inefficient. In contrast, RNA acceptor molecules containing a 5' Adenosine diphosphate showed efficient [α-^{32}P] GMP transfer, indicating that NS5 has a preference for RNA molecules containing a 5' Adenosine diphosphate nucleotide [269]. The authors therefore suggest that NS5 is a true GTase, capable of catalysing both steps of the RNA capping reaction.

NS5 is also responsible for viral RNA synthesis, carried out by the RNA dependent RNA polymerase domain. ZIKV NS5 RNA synthesis, either *de novo* or from a primed template, has been demonstrated using purified NS5, measured using [α-^{32}P] CTP incorporation after reactions were separated by PAGE. Mutations which cause the deletion of the MTase domain resulted in a decrease in the amount of synthesised RNA product, indicating that the MTase also plays a role in RNA synthesis [278]. Mutations within the thumb domain of the RdRP reduced binding to fluorescein-labelled SLA (F-SLA) RNA molecules, where changes in florescence intensity were used as a measure of binding of ZIKV NS5 in a fluorescence titration assay. Using fluorescence anisotropy, (where binding reduces fluorophore tumbling speed, resulting in detectable polarised emission), mutations within the MTase domain were shown to reduce F-SLA binding affinity. Circular dichroism analysis of MTase mutants vs. WT showed identical curves, indicating that the reduction in binding affinity was not due to structural rearrangements [9]. ZIKV NS5 is also involved in immune regulation and can block interferon signalling by targeting signal transducer and activator of transcription 2 (STAT2). Transfection of increasing amounts of NS5 expressing constructs into HEK 293T cells resulted in decreasing STAT2 expression, as measured by immunoblotting. STAT proteins are key signal transducers of interferon-stimulated genes, therefore degradation of these proteins allows ZIKV to impair interferon signalling [277].

1.4 Locked nucleic acids

Nucleic acid oligonucleotides have been investigated as potential antiviral agents since the 1970s, when Zamecnik and Stevenson used an antisense single-stranded synthetic DNA oligonucleotide to target cyclisation sequences in Rous Sarcoma virus [280]. Since then, advances in chemistry have enabled the synthesis of modified nucleic acids, which have improved properties compared with natural nucleotides. Incorporation of modified bases into oligonucleotide sequences can improve binding affinity, specificity and nuclease resistance; improving the antiviral effectiveness of these agents within the cellular environment. Amongst the most widely used modified nucleotides are locked nucleic acids (LNA). In LNA nucleotides, the ribose moiety is locked in a single conformational state. This effect is achieved by the addition of an extra methylene bridge, which links the 2'-O- and 4'-C-atoms and reduces the number of conformational degrees of freedom (Fig. 1.15) [281]. Depending on the form of ribose, L- or D-ribose, LNA nucleotides can have either a C3'-endo (β-D-LNA) conformation (similar to a ribose sugar), or a C2'-endo (α-L-LNA) conformation (similar to a

deoxyribose sugar) [282,283]. In this thesis, the term 'LNA' is used synonymously to describe β-D-LNA.

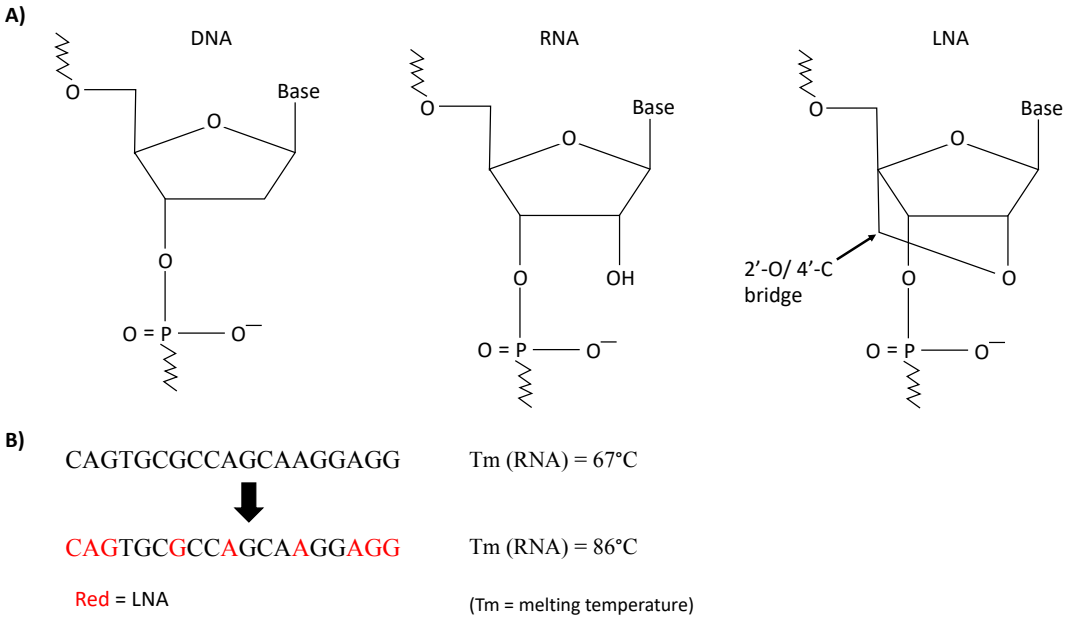


Figure 1.15 LNA nucleotides
A) *Schematic of DNA, RNA and LNA bases. LNA bases have an additional 2'O/ 4'-C bridge which stabilises the ribose ring in an ideal conformational for Watson crick base pairing.* ***B)*** *Incorporation of LNA bases into an oligonucleotide sequence increases the RNA binding Tm and nuclease resistance of the oligo, in addition to preventing RNaseH activation.*

LNA oligonucleotide designs are typically divided into two categories: mixmers and gapmers. Gapmers consist of a central DNA segment, which is generally longer than 7-8 nucleotides, flanked by two LNA segments at the 5' and 3' ends of the oligonucleotide. Gapmers are used to inhibit mRNA expression. When an antisense gapmer binds to an mRNA target, the central segment forms an DNA/ RNA hybrid, which is recognised by RNase H and enzymatically cleaved [284–286]. In a mixmer, LNA and DNA nucleotides are interspersed within the oligonucleotide sequence, avoiding RNase H activation and instead functioning via steric hindrance [287,288].

1.5 Project aims and objectives

Functional RREs in both CHIKV and ZIKV are fundamental for efficient replication of the viral genomes and represent a novel target for the develop of specific antiviral agents. Since no specific treatments or vaccines are available for either virus, there is an urgent need to investigate potential strategies for antiviral development. The aims of the thesis are to investigate the potential of RNA structures within the 5' regions of the CHIKV and ZIKV genomes to act as targets for antiviral therapies.
Based on previously acquired knowledge of the CHIKV 5' RREs, we designed a small panel of antisense locked nucleic acid oligonucleotides to target specific RREs, with the aim of disrupting essential RNA-RNA/ *trans* activating protein interactions which are required for

RRE functions. Since we aimed to block RRE formation by steric hinderance, we chose to use mixmer LNA oligonucleotides to target functional RREs in CHIKV. We studied the effect of these antisense LNA oligonucleotides on CHIKV replication in both infectious virus and sub-genomic replicon systems in order to determine the lifecycle stage at which antisense LNA oligonucleotides act.

RNA structures within the 5′ region of the ZIKV genome were then characterised using a combination of biochemical SHAPE mapping and thermodynamic algorithms. Based on the RNA structural map generated by this process, we analysed understudied RNA structures using reverse genetics, in order to determine the role of these structures in virus replication. Antisense-LNA were also designed to target RNA structures within the ZIKV 5′ UTR and adjacent capsid coding region.

Objectives:

1. Target RREs in the CHIKV 5′ UTR and adjacent nsp1 coding region in order to inhibit virus replication.
2. Characterise RNA structures in the 5′ UTR and adjacent capsid coding region of the ZIKV genome, using a combination of SHAPE mapping and reverse genetic approaches.
3. Using data from objective 2, target specific RNA structures in ZIKV using antisense LNA oligonucleotides.

Chapter 2 Methods

2.1 Materials

2.1.1 Plasmids

CHIKV:

The CHIKV infectious clone (CHIKV ICRES), CHIKV-Firefly luciferase sub-genomic replicon (Fluc-SGR) and CHIKV-translation reporter (Fluc-Rluc-(GDD>GAA)-SGR) cDNA clones are based on the isolate LR2006 OPY1, representing the East Central South African genotype and have been previously described [289]. Briefly, in Fluc-SGR, the second ORF (coding for the structural proteins) is replaced by a firefly luciferase gene. The Fluc-Rluc-SGR (GDD>GAA) additionally has a *Renilla* luciferase encoding gene fused within nsp3 and a GDD>GAA mutation in nsp4 which removes polymerase activity. Transfection control *Renilla* RNA (T7_Ren) contains a *Renilla* luciferase gene under the control of a T7 promoter.

ZIKV:

pCCI-SP6-ZIKV infectious clone DNA (ZIKV IC) is based on a consensus sequence of Brazilian Isolates of ZIKV. pCCI-SP6-ZIKV-N-luciferase (ZIKV N-luc IC) contains a truncated capsid coding sequence fused to a nano-luciferase reporter, with an upstream FMDV 2A sequence. After the FMDV sequence a second, full length codon optimised copy of capsid is present. pCCI-SP6-ZIKV-N-luciferase sub-genomic replicon (ZIKV N-luc Replicon) is based on ZIKV N-luc IC, but lacks the codon optimised structural proteins. ZIKV constructs were kindly provided by Andres Merits (University of Tartu) [290]. pUC57-M-SLE, which contains the first 1500 bp of the ZIKV IC sequence with M-SLE mutations, was purchased from Genewiz. pUC57-M-PK and pUC57-M-SLD, which contains the first 1500 bp of the ZIKV N-luc IC sequence, in addition to M-PK and M-SLD mutations respectively, was purchased from Genewiz. (The sequences changes in the M-SLE, M-PK and M-SLD constructs are described in section 5.2).

2.1.2 Primers

Oligonucleotide primers (IDT). The oligonucleotide sequences are provided in 5'-3' orientation.

Table 1 Primers

	Forward	Reverse
ZIKV Full genome	CGATTAAGTTGGGTAACGCCAGGGT	TAGACCCATGGATTTCCCCACACC
ZIKV 1-772	ATTTAGGTGACACTATAGAGTTGTTG	GCTCTTCTAGATCTCCGTGC
CHIKV 1-337	TAATACGACTCACTATAGGGATGGC TGCGTGAGACACACG	CGCACTGCGCATCGGGCAGA
CHIKV E1 region	GCATCAGCTAAGCTCCGCGTC	GGTGTCCAGGCTGAAGACATTG
ZIKV Outer amplicon	GTTGATCTGTGTGAATCAGACTG	GCTTCGGCTCTTGGTGAATTG
ZIKV Inner amplicon	CAGACTGCGACAGTTCGAG	CAAGGTAGGCTTCACCTTGTG
ZIKV mut sequence 5'	ACGGCCAGTGAATTCATTTAGG	
ZIKV Overlap extension	TCGGAATTGTTGGCCTCCTGC	CAGCAGGAGGCCAACAATTCC
RP_ZIKV Upstream of AvrII		AAGTCAAGGCCTGTCCTCGG
SHAPE 575		ACCATCCTGATGGGCCCATGAC
SHAPE 583		TCGCCAAGACCATCCTGATG
SHAPE 590		(FAM/HEX)-GCTAGAATCGCCAAGACCATC
SHAPE 802		CTGCCATAGCTGTGGTCAGCAG
SHAPE 824		CTCCCACGTCTAGTGACCTC
SHAPE 826		(FAM/HEX)-CACTCCCACGTCTAGTGACCTC

2.1.3 Bacterial strains

CHIKV DNA constructs were amplified following transformation of high efficiency DH5α *E. coli* cells (NEB). ZIKV DNA constructs were amplified following transformation of TransforMax EPI 300 Chemically Competent *E. coli* cells (Lucigen) or Turbo Competent *E. coli* cells (NEB).

2.1.4 Cell culture

Huh7 cells (hepatocytes derived from human hepatocellular carcinoma), BHK-21 cells (fibroblasts derived from Syrian golden hamster kidney) and Vero cells (African green monkey kidney) were a gift from M. Harris (University of Leeds, UK). Mammalian cells were cultured in Dulbecco's modified minimal essential medium (DMEM, Sigma) supplemented with 10% (v/v) foetal bovine serum (Thermo Fisher Scientific), 100 units/ml penicillin, 100 µg/ml streptomycin, 100 µM non-essential amino acids (Lonza) and incubated at 37 °C in 5% CO_2. To maintain mammalian cell lines, cells were harvested using trypsin/EDTA and seeded at dilutions of 1:3 to 1:20 and incubated at 37 °C in 5% CO_2.

C6/36 cells (*Aedes albopictus* larva) were a gift from S. Jacobs (The Pirbright Institute, UK). C6/36 cells were cultured in Leibovitz's L-15 media supplemented with 10% (v/v) foetal bovine serum, 10% tryptose phosphate broth and 100 U penicillin/100 µg streptomycin/ml (Leibovitz's L-15/PS). C6/36 cells were harvesting by scraping, passaged at dilutions of 1:3 to 1:8 and maintained at 28 °C.

All cell lines tested negative for mycoplasma.

2.1.5 Buffer composition

Buffer compositions are listed in appendix 1.

2.1.6 Antisense DNA oligonucleotides and LNA oligonucleotides

"+" indicates the positions of LNA nucleotides. The oligonucleotide sequences are provided in 5'-3' orientation.

Table 2 CHIKV LNA oligonucleotides

CHIKV LNA-162	+A+T+TCGG+TGTG+ACC+TGCC+T+T+G
CHIKV LNA-202	+T+A+TAGC+TAGA+TGCG+AG+A+A+C
CHIKV LNA-225	+G+T+CAA+TTT+CCTG+CTCTA+T+T+A
CHIKV LNA-760	+G+T+TTTA+GCTT+TTTC+CCT+C+T+C
Scrambled LNA (also used against ZIKV)	+G+C+ACAG+CGC+AAGT+ATG+T+T+A

Table 3 ZIKV antisense DNA oligonucleotides

ZIKV Oligo-15	GAACTGTCGCAGTCTGATTCAC
ZIKV Oligo-25	GCTTCAAACTCGAACTGTCGCAG
ZIKV Oligo-44	TGTTGCTAGCTTTCGCTT
ZIKV Oligo-56	TAAAACCTGTTGATACTGTTGC
ZIKV Oligo-80	ACTCTCGTTTCCAAATCCAA
ZIKV Oligo-91	TTCATGACCAGAAACTCTCGTT
ZIKV Oligo-96	GGTTTTTCATGACCAGAAACTC
ZIKV Oligo-123	GAATCCTCCGGATTTCTT
ZIKV Oligo-147	CTCCGCGTTTTAGCATATTGAC
ZIKV Oligo-163	GCTCACACGGGCTACTCCGC
ZIKV Oligo-204	CAGAAGTCCGGCTGGCAG
ZIKV Oligo-229	GACCATCCTGATGGGCCCAT
ZIKV Oligo-236	CGCCAAGACCATCCTGATG
ZIKV Oligo-242	GCTAGAATCGCCAAGACCATC
ZIKV Oligo connector 1	AGCTTTCGCTCATACTAACTGT
ZIKV Oligo Connector 2	CTCTCGTTCTTAGCTTTCGCT
ZIKV Oligo Connector 3	AATCCTTCATTCTCTCG
ZIKV Oligo connector 4	CCATGATAACAAAGG
ZIKV Oligo connector 5	CCATGTAATTAAGCCC
Scrambled Oligo	GCACAGCGCAAGTATGTTA

Table 4 ZIKV LNA oligonucleotides

ZIKV LNA set 1	
ZIKV LNA-29	+G+C+TTCA+AACT+CG+AACT+G+T+C
ZIKV LNA-142	+T+T+AG+CATA+TTG+ACAA+T+C+C
ZIKV LNA-158	+G+G+CT+AC+TC+CGCG+T+T+T+T
ZIKV LNA-191	+G+C+AGCC+TCT+TCA+AG+C+C+C
ZIKV LNA-213	+C+C+ATG+ACCC+AGC+AG+AAG+T+C+C
ZIKV LNA set 2	
ZIKV LNA-15	+G+A+ACTGT+CGCAG+TCTG+A+T+T
ZIKV LNA-25	+C+T+TCA+AACT+CGAA+CTGTC+G+C+A
ZIKV LNA-56	+T+A+AAA+CCTG+TTGA+TAC+TGT+T+G+C
ZIKV LNA-96	+T+T+TTCA+TGACC+AGAAA+C+T+C
ZIKV LNA-204	+A+G+A+AGTCC+GGCTG+G+C+A

2.1.7 Statistical tests

Statistical analysis of data were performed using Graphpad Prism 7.0c

2.2 Molecular Biology Methods

2.2.1 Native agarose gel electrophoresis

The integrity, purity and size of plasmid DNA, PCR products and restriction digests were determined by native agarose gel electrophoresis using 0.4 g (1%) or 0.8 g (2%) of analytical grade agarose (Sigma Aldrich) in 30 ml 1 x TAE with SYBR Safe DNA stain (Life Technologies at 1:10,000 dilution). Samples were combined with Gel loading Dye, purple (6X) (NEB), to a final concentration of 1 x, and were loaded alongside 3 μg 1 kb plus DNA ladder (NEB). DNA was resolved at 90 V for 45-60 minutes in 1 x TAE and visualised using a UV transilluminator.

2.2.2 Denaturing MOPS gel electrophoresis

Products of *in vitro* transcription were checked for integrity, purity and size using denaturing formaldehyde agarose gel electrophoresis in 3-(N-morpholino)propanesulfonic acid (MOPS) buffer. Samples were analysed using 0.3 g (1%) or 0.6 g (2%) agarose gels, made with analytical grade agarose (Sigma Aldrich) in 28.6 ml 1 x MOPS buffer and 1.4 ml of 37% formaldehyde with SYBR Safe DNA stain (Life Technologies at 1:10,000 dilution). Samples were combined with RNA loading dye (2 x) (NEB), to a final concentration of 1 x, and heated to 85 °C for 5 minutes to denature the RNA. Samples were loaded alongside either Millennium (for 1% gels) or Century RNA ladder (for 2% gels) (Ambion) and electrophoresis was performed at 80 V for 45 minutes in 1 x MOPS buffer and visualised using a UV transilluminator.

2.2.3 SDS-polyacrylamide gel electrophoresis

5 ml SDS-PAGE gels were made with a 10% resolving gel (10 ml) (4 ml 30% bis-acrylamide, 2.5 ml 1.5 M Tris-HCl pH 8.8, 4 ml ddH_20, 100 µl 10% SDS, 100 µl ammonium persulphate (APS), 10 µl N,N,N',N'- tetramethylethylenediamine (TEMED)) with a 5% stacking gel (0.83 ml 30% bis- acrylamide, 0.63 ml 1 M Tris-HCl pH 6.8, 3.4 ml ddH_20, 50 µl 10% SDS, 50 µl 10% APS, 5 µl TEMED). Samples were combined with 2 x denaturing sample buffer, to a final concentration of 1 x, and denatured by heating to 95 °C for 5 minutes. Samples were loaded alongside 3 µl colour pre-stained protein standard, broad range marker (11-245 kDa) (NEB). Electrophoresis was performed in 1 x SDS buffer for 90 minutes at 180V.

2.2.4 Denaturing polyacrylamide Urea gel electrophoresis

A gel stock solution was prepared for 7.5 % urea denaturing polyacrylamide gels by combining 80 ml 10 x TBE, 325 g urea, 150 ml 40% Acrylamide and made up to 800 ml with DEPC H_2O. Prior to electrophoresis, 400 µl APS and 40 µl TEMED were added to 30 ml of gel stock solution. Samples were combined with 1 µl of 4M NaOH and 32 µl acid stop mix (4:25 (v/v) mixture of 1 M unbuffered Tris-HCl and stop dye (85% formamide, ½ x TBE, 50 mM EDTA, pH 8.0, containing bromophenol blue and xylene cyanol tracking dyes and then incubated at 95 °C for 5 minutes. Electrophoresis was performed in 1 x TBE (Thermo Fisher) 180 V for 1 hour 30 minutes.

2.2.5 Native polyacrylamide gel electrophoresis

Native polyacrylamide gels (8%) were prepared by combining 6 ml Gel mix (Accugel 19:1, 40% (w/v) 19:1 Acrylamide: bis acrylamide solution, gas stabilized, (National diagnostics)), 21 ml DEPC H_2O, 400 μl APS, 40 μl TEMED and 3 ml 10 x TBE buffer (Thermo Fisher). Samples were combined with 3 μl native RNA loading dye and electrophoresis was performed in 1 x TBE (Thermo Fisher) at 180V for 1 hour 40 minutes.

2.2.3 Bacterial Transformation

For routine amplification of plasmids, DH5α, Turbo competent or EPI 300 *E. coli* cells were transformed according to the manufactures protocol. Briefly, approximately 50 ng of plasmid DNA was mixed with the *E. coli* cells and incubated for 15 minutes on ice, followed by 30 second heat shock at 42 °C and recovery on ice for 2 minutes. 250 μl of SOC outgrowth medium (NEB) was added to the transformation mix and incubated in a shaking incubator for 1 hour. 200 μl of the transformation mix was spread onto LB-ampicillin (AMP) agar plates (CHIKV) or soy agar plates (ZIKV), containing an appropriate concentration of antibiotic, and incubated for 16 hours at 37 °C

2.2.4 CHIKV Plasmid DNA amplification

For CHIKV DNA purification, 5 ml cultures containing LB-AMP and an appropriate concentration of ampicillin were inoculated with a single transformed bacterial colony and incubated for 5-7 hours. Starter cultures were then used to inoculate 250 ml of LB-AMP which was subsequently incubated for 16 hours at 37 °C. *E. coli* cells were the pelleted by centrifugation at 4000 x g for 10 minutes. Plasmid DNA was isolated and purified using a GeneJET Plasmid Maxiprep Kit (Thermo Scientific) following the manufacturer's instructions. DNA purity and quantity were assessed using a Nanodrop 1000 (Thermo Scientific) spectrophotometer, which measures the ratio of absorbance at different wavelengths. Nucleic acids absorb wavelengths of 260 nm, proteins absorb wavelengths of 280 nm and organic compounds such as phenol absorb wavelengths of 230 nm. By comparing the 260/280 and 260/ 230 ratio of absorbance, the purity of nucleic acid in a sample can be assessed.

2.2.5 ZIKV Plasmid DNA amplification

ZIKV DNA was purified using a method adapted from personal communication with Prof A. Merits. 20 ml of soy broth supplemented with 12.5 μg/ml chloramphenicol was inoculated with a transformed EPI 300 cell colony and incubated overnight. The following morning, 80 ml of soy broth was added to the culture, along with 0.1% Arabinose to activate the second origin of replication, and chloramphenicol to a final concentration of 12.5 μg/ml. Cultures were then incubated for a further 5 hours before cells were pelleted by centrifugation at 10,000 x g for 10 minutes. Pellets were resuspended, lysed and neutralised using Qiagen midi prep kit buffers, according to the manufacturer's instructions. Neutralised lysate was then centrifuged at 4000 x g for 5 minutes and filtered through gauze. Next, 0.7 volumes of isopropanol were added to the lysate and DNA was pelleted by centrifugation at 15,000 x g for 20 minutes at 4 °C. Pellets were then washed with 70% ethanol and resuspended in 750

µl DNA mini prep resuspension buffer, lysed in 750 µl lysis buffer and neutralised in 1,050 µl neutralisation buffer (Thermo). Lysate was then separated and purified through three DNA purification columns according to the manufacturer's instructions. DNA purity and quantity were assessed using a Nanodrop 1000 (Thermo Scientific) spectrophotometer (as described in 2.2.4).

2.2.6 PCR amplification of full-length ZIKV DNA

PCR products containing the full-length ZIKV IC, ZIKV IC mutant, ZIKV N-luc IC and ZIKV N-Luc replicon DNA sequence were PCR-amplified for use as templates for *in vitro* transcription reactions. For PCR reactions, 100 ng of ZIKV/Nano-luc ZIKV DNA was combined with 0.5 nM ZIKV full genome forward primer, 0.5 nM ZIKV full genome reverse primer, 1 µM dNTPs, 1 x GC buffer (NEB), 4 units of Phusion DNA polymerase (NEB) and nuclease-free water was added to a final volume of 100 µl. Reactions were performed in a Proflex PCR system (Life Technologies) with conditions set at 98 °C for 2 minutes, followed by 35 cycles of 98 °C for 15 seconds, 69.4 °C for 30 seconds, 72 °C for 6 minutes and 1 cycle 72 °C for 10 minutes. DNA was then precipitated by addition of 0.1 volume 3M NaOAc (pH 5.2) and 3 volumes of ice-cold 100% ethanol, followed by incubation at -80 °C for 60 minutes. DNA was pelleted by centrifugation at 16,000 x g at 4 °C for 30 minutes. The DNA pellet was then washed with ice cold 70% ethanol followed by centrifugation at 16,000 x g at 4 °C for 10 minutes. Remaining ethanol was aspirated and the pellet was resuspended in 20 µl of nuclease free water. DNA was analysed using a 1% agarose gel to check the size and purity of the DNA product and quantified using a NanoDrop 1000 spectrophotometer (Thermo scientific) (as described in 2.2.4).

2.2.7 PCR amplification of ZIKV 1-772

PCR reactions were performed to generate DNA fragments encompassing the ZIKV 5' genome region, base pairs 1-772. Reactions contained 100 ng of template DNA, 0.5 nM primers (ZIKV 1-772 forward primer:) (ZIKV 1-772 reverse primer: 1 µM dNTPs, 1 x GC buffer (NEB), 2 units of Phusion DNA polymerase (NEB) and nuclease free water was added to a final volume of 50 µl. Reactions were performed in a Proflex PCR system (Life Technologies) with conditions set at 95 °C for 2 minutes, followed by 25 cycles of 95 °C for 15 seconds, 50 °C for 15 seconds, 72 °C for 15 seconds and 1 cycle 72 °C for 5 minutes. DNA was then purified by ethanol precipitation as described in section 2.2.6.

2.2.8 PCR amplification of CHIKV 1-337

PCR reactions were performed to generate DNA fragments encompassing the CHIKV 5' genome region base pairs 1-337 (CHIKV 1-337). Reactions were performed as described in section 2.2.7 using the following primers:
CHIKV 1-337 forward primer:
CHIKV 1-337 reverse primer:
DNA was then purified by ethanol precipitation as described in section 2.2.6.

2.2.9 Overlap extension PCR

As the pUC57-M-PK and pUC57-199 constructs are based on the ZIKV N-LUC IC sequence, overlap extension PCR was used to generate DNA fragments which are compatible for ligation into the ZIKV IC construct. Prior to overlap extension PCR, two DNA precursor fragments were amplified using PCR. Precursor fragments contain a 19 bp region of overlap to allow annealing during overlap extension stage.

For the first precursor fragment, pUC57-M-PK and pUC57-210 DNA was used as a template in PCR reactions to amplify the first 491 bp of the ZIKV mutant sequence. Reactions were assembled as described in section 2.2.6, using the ZIKV mut sequence 5' forward primer and the ZIKV overlap extension reverse primer. For the second precursor fragment, ZIKV IC DNA was used as a template in a PCR reaction which amplified a 1135 bp region of the ZIKV sequence using the ZIKV overlap extension forward primer and the Upstream of AVRII reverse primer. Reactions were performed in a Proflex PCR system (Life Technologies) with conditions set at 95 °C for 2 minutes, followed by 25 cycles of 95 °C for 30 seconds, 55 °C for 30 seconds, 72 °C for 60 seconds and 1 cycle 72 °C for 5 minutes. PCR products were then separated using agarose gel electrophoresis, as described in section 2.2.1, and extracted using a GeneJET gel extraction kit (Thermo Fisher) according to the manufacturer's instructions.

For the overlap extension stage, 50 ng of the first precursor fragment and an equimolar concentration of the second precursor fragment were combined. 12 PCR reactions were assembled as described in section 2.2.7, excluding the addition of primers. An initial PCR reaction was performed using an Eppendorf mastercycler EP Gradient thermocycler with the conditions set at 95 °C for 2 minutes, followed by 6 cycles of 95 °C for 15 seconds, a heat gradient ranging from 50-70 °C for 30 seconds, 72 °C for 60 seconds and 1 cycle of 72 °C for 5 minutes. 0.5nM ZIKV mut sequence 5' forward primer and the Upstream of AVRII reverse primer were then added to the reactions. Secondary reactions were then performed using an Eppendorf mastercycler EP Gradient thermocycler, using the cycle conditions 95 °C for 2 minutes, followed by 30 cycles of 95 °C for 15 seconds, a heat gradient ranging from 50-70 °C for 30 seconds, 72 °C for 60 seconds and 1 cycle of 72 °C for 10 minutes. PCR products were then separated using agarose gel electrophoresis, as described in section 2.2.1, and extracted using a GeneJET gel extraction kit (Thermo Fisher) according to the manufacturer's instructions.

2.2.10 ZIKV DNA ligations and colony PCR screening

For ligation of mutant sequences into ZIKV IC constructs, both vector DNA (ZIKV IC) and insert DNA (pUC57-M-SLE) and products generated by overlap extension PCR described in section 2.2.9) was digested using a double digest with EcoRI-HF and AvrII (NEB). Restriction digest reactions were assembled using 5 µg of template DNA, 20 units of EcoRI-HF, 5 units of AvrII, 1 x cutsmart buffer (pH 8.0) and made to a final volume of 50 µl with nuclease free H_2O. Reactions were incubated at 37 °C using a Proflex PCR system (Life Technologies) for 3 hours. To dephosphorylate vector DNA, 5 units of calf intestinal alkaline phosphatase were added to vector DNA reactions, and reactions were incubated at 37 °C for a further 30 minutes. DNA was then purified by ethanol precipitation as described in section 2.2.6. Ligation reactions were assembled using a 5:1 insert to vector molar ratio, using 50 ng of vector DNA, 1 x T4

DNA ligase buffer and 3 units of T4 DNA ligase (NEB), made up to a volume of 20 µl with nuclease free H_20. Reactions were incubated at 16 °C overnight and the heat inactivated at 65 °C for 10 minutes using a Proflex PCR system (Life Technologies). 10 µl of ligase reactions was then transformed into Turbo competent *E. coli* cells as described in section 2.2.3. For colony PCR DNA ligations, plates were incubated at 37 °C for 16 hours, followed by a further 6 hours at 30 °C.

For colony screening, PCR reactions were assembled as described in section 2.2.7, using ZIKV mut sequence 5' forward primer and the Upstream of AVRII reverse primer. Colonies were streaked onto plates and then added to PCR reactions. PCR reactions were performed using a proflex thermocycler with the conditions set at 95 °C for 2 minutes, followed by 35 cycles of 95 °C for 15 seconds, 55 °C for 30 seconds, 72 °C for 30 seconds and 1 cycle of 72 °C for 10 minutes. PCR products were then separated using agarose gel electrophoresis and sequenced using the ZIKV Outer amplicon forward primer. Colonies with the correct sequence were used to inoculate 20 ml starter cultures of soy broth supplemented with 12.5 µg/ml chloramphenicol. Starter culture were used to inoculate 250 ml culture of soy broth/ 12.5 µg/ml chloramphenicol, which was incubated overnight at 37 °C. Plasmid DNA was isolated and purified using a GeneJET Plasmid Maxiprep Kit (Thermo Scientific) following the manufacturer's instructions. DNA purity and quantity were assessed using a Nanodrop 1000 (Thermo Scientific) spectrophotometer.

2.2.11 Restriction digest of CHIKV DNA constructs

Prior to *in vitro* transcription reactions, CHIKV ICRES, Fluc-SGR, Fluc-Rluc-(GDD>GAA)-SGR DNA was linearised in a restriction enzyme reaction using 10 µg of template DNA, 20 units Not-I HF (NEB), 1 x cutsmart buffer (pH 8.0) and made to a final volume of 50 µl with nuclease free H_2O. Reactions were incubated at 37 °C overnight using a Proflex PCR system (Life Technologies). Linear DNA was then purified using a Wizard SV gel and PCR clean-up system (Promega) according to the manufacturer's instructions. DNA purity and quantity were assessed using agarose gel electrophoresis and a Nanodrop 1000 (Thermo Scientific) spectrophotometer.

2.2.12 *In vitro* RNA transcription

For generation of capped CHIKV and ZIKV RNA, 1 µg of template DNA plasmid was used as a template for production of 5' [m7G(5')ppp(5')G] capped (m7G capped) RNA, using an SP6 mMessage mMachine kit (Life Technologies) or an SP6-Scribe ™ standard RNA IVT kit (Lucigen), according to the manufactures instructions. Following transcription, DNA template was removed by DNase I (Life Technologies) treatment and purified using LiCl precipitation solution (7.5 M LiCl, 50 mM EDTA, pH 8.0) (invitrogen) and water added to a final volume of 50 µl and final LiCl concentration of 2.5 M. Samples were chilled at -20 °C for 30 minutes and then centrifuged at 13,000 x g for 15 minutes at 4 °C. Supernatant was aspirated and pellets dried before resuspension in nuclease free water.

For uncapped ZIKV IC and ZIKV 1-772 RNA, 1 µg of PCR DNA was used as a template for uncapped RNA transcription using an SP6 RiboMAX™ Express Large-Scale RNA Production System (Promega) according to the manufacturer's instructions. Template DNA was removed

by addition of 2 units of RQ1 RNase-Free DNase and RNA was purified using LiCl precipitation as described previously.

For T7_Ren RNA, 1 µg of cDNA template was linearised using XhoI (NEB) (as descried in section 2.2.11). DNA was combined with 1 X RNAPol Reaction Buffer (NEB), supplemented with 4 mM each of ATP, UTP, CTP and GTP, 3 mM cap analogue, 14 mM $MgCl_2$, 0.32 units of Pyrophosphatase, Inorganic (yeast) (NEB), 160 units of RNase inhibitor (promega) and 200 units of T7 polymerase. Reactions were incubated at 37 °C for 3 hours. Following transcription, remaining DNA template was removed by DNase I (Life Technologies) treatment and RNA was purified by ethanol precipitation.

For all *in vitro* transcription products, RNA purity and quantity were determined using a Nanodrop 1000 (Thermo Scientific) spectrophotometer and denaturing agarose MOPS gel electrophoresis.

2.3 CHIKV LNA binding assays using ^{32}P

2.3.1 LNA and CHIKV 1-337 ^{32}P end labelling

For end labelling of LNA oligonucleotides, 60 pmoles of LNA was mixed with 2 µl 10 x PNK buffer, 10 units of polynucleotide kinase, 15 µCi of ATP [γ-^{32}P] and made up to a volume of 20 µl with RNase free H_2O. Reactions were incubated at 37 °C for 30 minutes, followed by incubation at 65 °C for 20 minutes to inactivate the enzyme. LNA oligonucleotides were precipitated by addition of 3 volumes of ethanol and 0.1 volumes NaOAc (pH 5.2), followed by incubation at -80 °C for 60 minutes. LNA oligonucleotides were pelleted by centrifugation at 16,000 x g for 30 minutes and washed with 0.5 ml 70% ethanol, before being resuspended in 60 µl 0.5 x TE buffer.

For end labelling CHIKV 1-337, 20 pmoles of LNA was mixed with 2 µl 10 x polynucleotide kinase (PNK) buffer, 10 units of polynucleotide kinase, 5 µCi of ATP [γ-^{32}P] and made up to a volume of 20 µl with RNase free H_2O. End labelling reactions and ethanol precipitations were carried out as previously described and RNA was resuspended in 20 µl 0.5 x TE buffer.

2.3.2 Electromobility shift assays

For LNA EMSAs, 3.125 pmoles of CHIKV 1-337 RNA in 0.5 x TE buffer was heated to 95 °C for 2 minutes and incubated on ice for 2 minutes. 3.3 x RNA folding buffer (330 mM HEPES (pH 8.0), 20 mM $MgCl_2$ and 330 mM NaCl) was added to a final volume of 10 µl and reactions incubated for 20 minutes at 37 °C. 1.25 pmoles of ^{32}P labelled LNA oligonucleotides were added (a 2.5:1 ratio of RNA:LNA) and reactions were incubated at 37 °C for 30 minutes. Reactions were analysed using 8% native PAGE gel electrophoresis alongside 4 pmoles of ^{32}P labelled CHIKV 1-337 RNA at 180 V for 100 minutes in 1 x TBE (Thermo Fisher). Samples were analysed using the following native loading buffer: 30% Glycerol, 0.25% Xylene Cyanol, 0.25% Bromphenol Blue, in nuclease free water (pH 7.0). The gel was fixed for 1 hour in fix solution (12% Methanol, 10% Acetic acid) and dried for 1 hour in a gel drier. Gels were exposed to a Fujifilm imaging plate for 16 hours and visualised using a Fujifilm FLA-5000 phosphoimager

illumination laser 635. For dose response LNA-162 binding assays, binding reactions were performed as above, using a range of CHIKV 1-337 RNA concentrations from 0.25 pmoles to 6.25 picomoles.

2.4 Sub-genomic replicon transfection and cell viability assays

2.4.1 LNA dose response luciferase assay

Huh7 cells were seeded in 24 well plates at $5x10^4$ cells per well and maintained overnight in 1 ml of DMEM until monolayers reached 80-90% confluency. Cells were washed once in 1 x PBS before addition of 400 µl opti-MEM reduced-serum media and 100 µl transfection media was added dropwise. Transfection media was prepared according to the manufacturer's instructions (Invitrogen), using 1 µl Lipofectamine 2000, 250 ng of Fluc-SGR RNA, 50 ng *Renilla* luciferase RNA, and appropriate concentrations of LNA, before being made up to 100 µl using opti-MEM reduced-serum media. For LNA transfected wells, LNA was added to the RNA mix prior to the addition of lipofectamine 2000. Monolayers were maintained for 6 hours after transfection, before being washed once in 1x PBS and lysed in 100 µl 1 x passive lysis buffer (Promega). Lysates were stored at -20 °C, prior to analysis. To measure luciferase expression, 50 µl Firefly Dual-luciferase substrate (Promega) was added to 20 µl cell lysate and quantified using a FLUOstar Optima luminometer (BMG labtech) after which, 50 µl *Renilla* Dual-luciferase substrate was added to the same sample and expression levels were measured.

2.4.2 CHIKV translation assay

Huh7 cells were seeded at $5x10^4$ cells per well in a 24 well plate and maintained overnight in 1 ml of DMEM until monolayers reached 80-90% confluency. Transfection media was prepared according to the manufacturer's instructions. 500 ng of WT or GAA mutant CHIKV dual-luciferase replicon RNA was combined with 1 µl lipofectamine 2000 and 100 µl opti-MEM per well. For LNA transfected wells, LNA was added to the RNA mix prior to the addition of lipofectamine. Cells were washed once in 1 x PBS before addition of 400 µl opti-MEM reduced-serum media and 100 µl transfection media added dropwise. 4 hours post transfection, cells were washed with PBS and 1 ml of DMEM was added. Cells were maintained until 4, 6 and 12 hours post transfection, washed with PBS and lysed for 20 minutes, using 100 µl passive lysis buffer 2.0 (Biotium). Lysates were stored at -20 °C until analysis. To measure luciferase expression, 50 µl Firefly luciferase assay buffer 2.0 (Biotium) was added to 20 µl cell lysate and quantified using a FLUOstar Optima luminometer (BMG labtech). Then, 50 µl *Renilla* luciferase assay buffer 2.0 (Biotium) was added to the same sample and expression of *Renilla* luciferase was quantified using the same method as for Firefly luciferase.

2.4.3 LNA cell viability assay

Huh7 cells were seeded in 96 well plates at $1x10^4$ cells per well and maintained overnight in 100 µl of DMEM. The following day, monolayers at 90% confluency were washed once with 1 x PBS before addition of 25 µl opti-MEM reduced-serum media and 25 µl of transfection media in a dropwise manner. Transfection media was prepared according to the manufacturer's instructions using 0.2 µl of Lipofectamine 2000 and made up to 25 µl using opti-Mem. Where appropriate, LNA oligonucleotides were added to the transfection media in concentrations ranging from 10 nM to 10 µM. After transfection monolayers were maintained for 6 or 24 hours, before the media was removed and replaced with 100 µl of 1 mg/ml Thiazolyl Blue Tetrazolium Bromide (MTT) (Sigma) and incubated at 37 °C in 5% CO_2 for 30 minutes. After incubation, MTT solution was replaced with 100 µl DMSO and the plate was shaken at 60 rpm for 5 minutes. Absorbance at 570 nm was determined using an Infinite F50 microplate reader (Tecan). Absorbance was normalised and expressed as a percentage of untreated control cells.

2.4.4 ZIKV oligonucleotide and LNA co-transfection

Vero cells were seeded in 24 well plates at $5x10^4$ cells per well and maintained overnight in 1 ml of DMEM until monolayers reached 80-90% confluency. Transfection media was prepared according to the manufacturer's instructions (Invitrogen), using 250 ng ZIKV Nano-luc replicon RNA per well, 1 µl lipofectamine per well and ZIKV oligonucleotide/LNA (to a final concentration of 100 nM), made up to a volume of 100 µl using opti-MEM reduced-serum media. For ZIKV DNA oligonucleotide/LNA transfected wells, ZIKV oligonucleotide/LNA was added to the RNA mix prior to the addition of lipofectamine 2000. Monolayers were maintained for 4 hours post transfection, before cells were washed using 1 x PBS and 500 µl DMEM was added to the cells. At 6, 24 and 48 hours post transfection, 80 µl of cell supernatant was collected and added to 20 µl 5 x Passive lysis buffer (Promega). At 48 hours, cells were washed with PBS and monolayers were lysed using 100 µl 1 x Passive lysis buffer. Lysates were stored at -20 °C, prior to analysis. To measure luciferase expression, 20 µl of Nano-Glo Luciferase substrate (Promega) was added to 20 µl cell lysate and quantified using a FLUOstar Optima luminometer (BMG labtech).

2.5 Virus production

2.5.1 Production of CHIKV

1.2×10^6 BHK-21 cells in 400 µl ice-cold DEPC-PBS were electroporated with 1 µg 5'-capped *in vitro* transcribed RNA in a 4 mm electrocuvette, with a single square wave pulse at 260 V for 25 ms using a Bio-Rad electroporator. Cells were then incubated on ice for 5 minutes and seeded into a T175 flask in 20 ml DMEM. After 24 h, supernatant was aspirated and titred by plaque assay as described below.

2.5.2 Quantification of CHIKV titre by plaque assay

BHK cells were seeded at 1 x 10^5 cells per well in 12-well plates and maintained overnight in 1 ml DMEM. Huh7 cell supernatant was thawed on ice and a 10-fold dilution series from 10^{-0} to 10^{-7} was prepared, using 20 µl virus stock in 180 µl 2% DMEM. Monolayers of 80% confluency were washed once in 1 x PBS and 150 µl of virus dilution added to monolayers. Plates were rocked for 10 minutes before incubation at 37 °C in 5% CO_2 for a further 50 minutes. After incubation, 1 ml of overlay media containing 0.8% methylcellulose in DMEM was added. Plates were then incubated at 37 °C in 5% CO_2 for 48 hours. Following this, the overlay media was removed and monolayers washed in 1 x PBS. Monolayers were then fixed in a 5% formaldehyde solution of 30 minutes, before staining using 0.05% Crystal violet solution. Plaque forming units per ml (PFU/ml) were calculated using the following equation:

$$\text{PFU/ml} = \frac{\text{Average number of plaques at dilution}}{\text{Inoculum in ml x Dilution factor}}$$

2.5.3 ZIKV Production

For production of infectious ZIKV, 3 x 10^5 Vero cells or 6 x 10^5 C6/36 cells were seeded in 6 well plates and maintained overnight. The following day, monolayers were washed with 1 x PBS and 2.5 ml opti-MEM reduced-serum media (Life Technologies) added. Cells were transfected with 500 µl of transfection media, containing 5 µl Lipofectamine 2000 (Life Technologies) and 2.5 µg of infectious WT or mutant ZIKV RNA according to manufacturer's instructions. For Vero cells, cells were incubated at 37 °C in 5% CO2 for 5 days. For C6/36 cells, cells were incubated at 28 °C for 5, 7 and 9 days. For both cell lines, supernatant was harvested and stored at -80 °C.

2.5.4 Quantification of ZIKV titre by plaque assay

For ZIKV plaque assays, 1 x 10^5 Vero cells per well were seeded in 12-well plates and maintained overnight in 1 ml DMEM. Plaque assays were set up as described in section 2.4.2 and incubated at 37 °C in 5% CO_2 for 5 days. Monolayers were fixed in a 5% Formaldehyde solution of 30 minutes, after which they were stained using 0.05% Crystal violet solution. Plaque forming units per ml (PFU/ml) were calculated using the following equation:

$$\text{PFU/ml} = \frac{\text{Average number of plaques at dilution}}{\text{Inoculum in ml x Dilution factor}}$$

2.6 CHIKV LNA experiments using infectious virus

2.6.1 LNA oligonucleotide transfection of CHIKV infected Huh7 cells

Huh7 cells were seeded in 12 well plates at a confluency of $1x10^5$ cells per well overnight in 1 ml of DMEM until monolayers reached 80-90% confluency. The following day, cells were infected with CHIKV (MOI 1), rocked for 10 minutes and incubated for a further 50 minutes. Monlayers were then washed with 1 x PBS and maintained in 1 ml of DMEM for 3 hours. During this incubation period, LNAs were prepared for transfection using 2 µl of Lipofectamine 2000 per well, according to the manufacturer's instructions, using a total volume of 200 µl opti-MEM reduced-serum media per well. At the end of the 3-hour incubation, cells were washed with 1 x PBS and 800 µl of opti-MEM reduced-serum media was added. Monolayers were then transfected with 600 nM or 3 µM LNA. Monolayers were incubated with the transfection mix for 4 hours, before being washed with 1 x PBS and maintained in 1 ml DMEM until time points were taken. Cell supernatant was harvested and CHIKV titre was quantified by plaque assay as described in section 2.5.2.

2.6.2 Quantification of CHIKV Genome copies using qRT-PCR

LNA transfections were set up as described in section 2.6.1. After transfection, monolayers were incubated for a further 16 hours, then washed with 1 x PBS and lysed using 500 µl TRI Reagent® Solution (Applied Biosystems). Lysates were incubated at room temperature for 30 minutes to allow deactivation of infectious CHIKV. Following deactivation, 100 µl of Chloroform was added to the lysates, which were vortexed and then incubated at room temperature for 2-3 minutes. Samples were then centrifuged at 16,000 x g at 4 °C for 30 minutes before the upper phase was aspirated and placed in a separate tube. 250 µl isopropanol was then added and samples were inverted 4-6 times before incubation at room temperature for 5 minutes. Samples were then centrifuged at 12,000 x g at 4 °C for 10 minutes and pellets were washed using 500 µl ice cold 75% ethanol. Samples were then centrifuged at 7500 x g at 4 °C for 5 minutes and the remaining ethanol was aspirated. Pellets were resuspended in 20 µl of nuclease free water and quantified using a NanoDrop 1000 spectrophotometer (Thermo scientific). 1 µg of extracted RNA was used to generate cDNA using an RNA-to-cDNA kit (appliedbiosystems) according to the manufacturer's instructions.

Quantitative PCR was performed using the qPCRBIO SyGreen Blue Mix Lo-ROX (PCR Biosystems), with primers amplifying a 131 bp region of the CHIKV E1 sequence, 100 ng of cDNA template and the following PCR program: 95 °C for 2 mins, 40 x (95 °C for 5 seconds, 60 °C for 30 seconds), dissociation curve 60-95 °C, as pre-defined by the Mx3005P thermal cycler (Agilent Technologies). For standards to quantify copy numbers in the respective samples, *In vitro* transcribed CHIKV ICRES RNA was reverse transcribed and serial dilutions from 10^{-2} to 10^{-7} were used. Cell supernatant was also harvested for each condition and CHIKV titre quantified by plaque assay for comparison.

2.6.3 LNA transfection of CHIKV infected cells time course

LNA transfections were set up as described in section 2.6.1 using 3 µM LNA. Monolayers were incubated with the transfection mix for 4 hours after which, supernatant was harvested and stored at -80 °C. Cells were then washed with 1 x PBS and 1 ml of DMEM was added. Following a further 5 hours incubation, 250 µl of supernatant was harvested and stored at -80 °C. Further time points were taken at 16, 24 and 48 hours post infection. Harvested media was replaced with 250 µl DMEM at each time point. Released CHIKV titre for each time point was quantified by plaque assay as described earlier.

2.6.4 Infection of LNA transfected Huh7 cells

Huh7 cells were seeded in 12 well plates at a confluency of $1x10^5$ cells per well and incubated overnight in 1 ml of DMEM, until monolayers reached 80-90% confluency. The following day, monolayers were washed with 1 x PBS and 800 µl of opti-MEM was added. LNA transfection mix were prepared according to the manufacturer's instructions, using 2 µl of Lipofectamine 2000 per well, 3 µM LNA oligonucleotides and a total volume of 200 µl opti-MEM reduced-serum media per well. Monolayers were then transfected with 200 µl transfection mix and incubated for 4 hours. Following incubation, monolayers were washed with 1 x PBS and infected with CHIKV (MOI 10), rocked for 10 minutes, and incubated for a further 50 minutes. Monolayers were then washed with 1 x PBS and maintained in 1 ml of DMEM until harvest. Supernatant was harvested at 6 and 24 hours post infection and released CHIKV titre was determined by plaque assay as described previously.

2.6.5 Dose response transfection of LNA oligonucleotides in CHIKV infected Huh7 cells

LNA transfections were set up as described in section 2.6.1, with the exception that cells were transfected with either 0, 0.03, 0.075, 0.15, 0.3. 0.6. 1.25 or 3 µM LNA-162 or scrambled LNA. Cells were incubated with the transfection mix for 4 hours, then washed with 1 x PBS and maintained in 1 ml DMEM for a further 16 hours. Cell supernatant was harvested and CHIKV titre was quantified by plaque assay as described previously.

2.6.6 Assessment of CHIKV protein expression by western blot

LNA transfections were set up as described in section 2.6.1, with 3 µM LNA. Monolayers were incubated with the transfection mix for 4 hours, then washed with 1 x PBS and maintained in 1 ml DMEM for a further 16 hours. Following incubation, monolayers were lysed in IP lysis buffer (25 mM Tris-HCl (pH 7.4), 150 mM NaCl, 1 mM EDTA, 1% NP-40 Alternative and 5% glycerol) and incubated at room temperature for 30 minutes to ensure complete lysis had taken place. Protein levels were quantified using a Pierce™ BCA Protein Assay Kit (Thermo scientific) according to the manufacturer's instructions. Equal amounts of protein lysate were loaded onto an SDS-PAGE gel and separated by electrophoresis at 120V for 90 minutes. Protein was transferred onto an Immobilon-FL PVFD transfer membrane (MERCK) using a TE77X semi-dry transfer (Hoefer), in Towbin buffer (25 mM Tris (pH 8.3), 192 mM glycine, 20% (v/v) methanol) at 15 V for 60 minutes. Membranes were blocked using 1:1 diluted Odyssey® Blocking Buffer (LI-COR) in PBS for 30 minutes.

Membranes were probed with primary anti-capsid (1:1000, rabbit polyclonal, in-house) or anti-nsp1 (1:1000, rabbit polyclonal, in-house) antibody, in addition to the housekeeping protein actin (clone AC-15, mouse monoclonal, Sigma) in diluted Odyssey® Blocking Buffer in PBS (LI-COR) overnight at 4 °C. The following morning, primary antibody was removed and membranes were washed 3 times using 1 x PBS. Membranes were next stained with secondary antibodies (IRDye® 800CW Donkey anti-Mouse; IRDye® 680LT Donkey anti-Rabbit; Li-Cor) for 1 hour at room temperature. Membranes were then washed 3 times using 1 x PBS and dried before being imaged were imaged using an Odyssey® Fc Imaging System (Li-Cor).

2.7 ZIKV infectious virus assays

2.7.1 ZIKV infections

For ZIKV infections, $1x10^5$ Vero cells or $2x10^5$ C6/36 cells were seeded in 12 well plates or $5x10^4$ cells per well in 24 well plates. The following day, cells were washed once in 1 x PBS and infected with ZIKV in 200ul of 2% FBS in PBS. Cells were rocked for 10 minutes and the incubated at 37 °C (Vero) or 28 °C (C6/36) for a further 50 minutes. Cells were then washed with 1 x PBS and 1 ml of either DMEM (Vero) or Leibovitz's L-15 (C6/36) was added. Cells were maintained until harvest, where the supernatant was aspirated and stored at -80 °C.

2.7.2 ZIKV Infectious centre assays

For ZIKV infectious centre assays, 4 x 10^5 Vero cells were seeded in a 6 well plate and incubated overnight. Huh7 cells were washed with 1 x PBS, trypsinised and pelleted by centrifugation at 1000 x g for 2 minutes. In duplicate, cells were washed in ice cold DEPC-PBS and resuspended to 2 x 10^5 cells in 400 µl DEPC-PBS. 1 µg of either WT or mutant ZIKV RNA was added to an electroporation cuvette and the 400 µl cells suspension was mixed with the RNA. The Huh7 cell/ RNA mix was electroporated using a single square wave pulse at 250 V for 25 ms and then incubated on ice for 2 minutes. 2 x electroporated cell suspensions were pooled into a final volume of 1.2 ml in DMEM and serially diluted using 1 in 10 dilutions from 10^{-1} to 10^{-6}. Vero cell monolayers were washed with 1 x PBS and 1 ml of diluted cell suspension

was added to the wells. Cells were then incubated at 37 °C for 2 hours, after which the media used to apply the transfected cell inocula was removed and monolayers washed with 1 x PBS. 1 ml of overlay media containing 0.8% methylcellulose in DMEM was added and monolayers were incubated for 5 days, then fixed in a 5% Formaldehyde solution of 30 minutes. Monolayers were stained using 0.05% Crystal violet solution and plaque forming units per µg RNA (PFU/ µg RNA) were calculated using the following equation:

$$\text{PFU/ µg RNA} = \frac{\text{Average number of plaques at dilution}}{\text{Inoculum in ml x Dilution factor}}$$

2.7.3 Assessment of ZIKV protein expression by western blot

To assess ZIKV protein expression levels, $1x10^5$ Vero cells or $2x10^5$ C6/36 cells were seeded in 12 well plates and maintained overnight. The following day cells were transfected with WT or mutant ZIKV RNA as described in section 2.5.3 and incubated for 24-72 hours. After incubation, cells were lysed in IP lysis buffer and incubated at room temperature for 30 minutes to ensure complete lysis had taken place. Protein levels were quantified using a Pierce™ BCA Protein Assay Kit (Thermo scientific) according to the manufacturer's instructions. Equal amounts of protein lysate were loaded onto an SDS-PAGE gel and separated by electrophoresis at 120 V for 90 minutes. Proteins were transferred to membranes as described in section 2.5.6.

Membranes were probed with primary anti-NS5 (1:1000, rabbit polyclonal, in-house) or anti-NS3 (1:1000, rabbit polyclonal, in-house) in 1:1 diluted Odyssey® Blocking Buffer in PBS (LI-COR) overnight at 4 °C. The following morning, primary antibody was removed and membranes were washed 3 times using 1 x PBS. Membranes were next stained with secondary antibodies (IRDye® 680LT Donkey anti-Rabbit; Li-Cor) for 1 hour at room temperature. Membranes were then washed 3 times using 1 x PBS and dried before being imaged were imaged using an Odyssey® Fc Imaging System (Li-Cor).

2.7.4 Extraction and sequencing of ZIKV RNA

For analysis of ZIKV M-SLE mutations, C6/36 cells were transfected with either WT or M-SLE ZIKV RNA, as described in section 2.5.3, and cell supernatant collected 9 days post transfection. Cell supernatant was also collected from untransfected cells as a control. To extract total RNA, 200 µl of cell supernatant was combined with 800 µl TRI Reagent® Solution (Applied Biosystems) and incubated for 30 minutes to deactivate ZIKV. RNA was extracted as described in section 2.6.2 and reverse transcribed using Superscript III reverse transcriptase (Thermo Fisher) according to the manufacturer's instructions. Briefly, 1 µg of total RNA was combined with 1 µl of 50 µM random hexamers, 1 µl of 10mM dNTPs, made up to a final volume of 13 µl with nuclease free H_2O and incubated at 65 °C for 5 minutes, followed by a 1-minute incubation on ice. Next, 4 µl First strand buffer, 1 µl 0.1mM DTT and 1 µl Superscript III reverse transcriptase were added per reaction. Reactions were then incubated in a Proflex PCR system with conditions set at 25 °C for 5 minutes, 50 °C for 60 minutes and 70 °C for 15 minutes.

For analysis of ZIKV M-PK and M-SLD mutations, Vero cells were seeded at $5x10^4$ cells per well in 24 well plates and infected with either WT, M-PK or M-SLD, as described in section 2.7.1,

(MOI of 0.005, due to the low titre of M-PK virus stocks). At 24 and 48 hours post infection, cells were washed with 1 x PBS and 250 µl TRI Reagent® Solution (Applied Biosystems) was added, followed by incubated for 30 minutes at room temperature. Following deactivation, 50 µl of chloroform was added to the lysates, which were vortexed and then incubated at room temperature for 2-3 minutes. Samples were then centrifuged at 16,000 x g at 4 °C for 30 minutes before the upper phase was aspirated and placed in a separate tube. 125 µl isopropanol was then added and samples were inverted 4-6 times, before incubation at room temperature for 5 minutes. Samples were then centrifuged at 12,000 x g at 4 °C for 10 minutes and pellets were washed using 500 µl ice cold 75% ethanol. Samples were then centrifuged at 7500 x g at 4 °C for 5 minutes and remining ethanol was aspirated. Pellets were resuspended in 20 µl of nuclease free water and quantified using a NanoDrop 1000 spectrophotometer (Thermo scientific). 1 µg of extracted RNA was used to generate cDNA using an RNA-to-cDNA kit (Applied Biosystems) according to the manufacturer's instructions.

cDNA was then used as a template for a 2-step PCR reaction. For the first step, 2 µl of ZIKV cDNA was combined with 0.5 nM forward outer amplicon primer, 0.5 nM reverse outer amplicon primer, 1 µM dNTPs, 1 x GC buffer (NEB), 4 units of Phusion DNA polymerase (NEB) and nuclease free water was added to a final volume of 50 µl. Reactions were performed in a Proflex PCR system (Life Technologies) with conditions set at 95°C for 5 minutes, followed by 20 cycles of 95 °C for 30 seconds, 55 °C for 30 seconds, 72 °C for 1 minute and 1 cycle 72 °C for 5 minutes.

The secondary PCR reaction was set up using 2 µl of the primary PCR reaction as the template, with 0.5 nM forward inner amplicon primer, 0.5 nM reverse inner amplicon primer, 1 µM dNTPs, 1 x GC buffer (NEB), 4 units of Phusion DNA polymerase (NEB) and nuclease free water was added to a final volume of 50 µl. Reactions were performed in a Proflex PCR system (Life Technologies) with conditions set at 95 °C for 5 minutes, followed by 35 cycles of 95 °C for 30 seconds, 58 °C for 30 seconds, 72 °C for 1 minute and 1 cycle 72 °C for 5 minutes. DNA was purified using a Wizard SV Gel and PCR Clean-Up System (Promega) according to the manufacturer's instructions and DNA integrity, purity and size were confirmed by 2% agarose gel electrophoresis. DNA was then sequenced using Sanger sequencing (Eurofins), using the forward inner amplicon primer.

2.7.5 ZIKV LNA transfections using ZIKV N-Luc infectious virus

Vero cells were seeded at $5x10^4$ cells per well in 24 well plates and incubated overnight. Cells were then infected (MOI 0.1) with WT ZIKV N-luc IC, as described in section 2.7.1, and incubated for 3 hours. During this time, LNA oligonucleotides were prepared for transfection according to the manufacturer's instructions, using 1 µl of Lipofectamine 2000 per well, 3 µM or 5 µM LNA and total volume of 100 µl opti-MEM reduced-serum media per well. At the end of the 3-hour incubation, cells were washed with 1 x PBS and 400 µl of opti-MEM reduced-serum media was added. 100 µl transfection mix was added to the monolayers and cells were incubated for 4 hours, before being washed with 1 x PBS and maintained in 500 µl DMEM until harvest. At 24 and 48 hours post transfection, 80 µl of cell supernatant was collected and added to 20 µl 5 x Passive lysis buffer (Promega). Lysates were stored at -20°C, prior to analysis. To measure luciferase expression, 20 µl of Nano-Glo Luciferase substrate (Promega)

was added to 20 µl cell lysate and quantified using a FLUOstar Optima luminometer (BMG labtech).

2.8 SHAPE Reactions

2.8.1 SHAPE primer selection

For ^{32}P primer end labelling, 10 nM primer was mixed with 20 units of polynucleotide kinase (NEB), 2 µl of 1x FS buffer (Invitrogen), 10 µCi of ATP [γ-^{32}P] and made to a final volume of 20µl. Reactions were incubated at 37 °C for 30 minutes, followed by incubation at 65 °C for 20 minutes. DNA was then precipitated by addition of 0.2 volumes of 5M NaCl, 0.1 volume of 100mM EDTA, 20 volumes of absolute ethanol and incubated at -80 °C for 30 minutes. DNA was pelleted by centrifugation at 16,000 xg for 40 minutes and the supernatant aspirated. The DNA pellet then was resuspended in 50 µl 1mM HEPES buffer.

For primer extensions, 3 µl of labelled primer was added to 300 ng ZIKV 1-772 RNA and incubated at 95 °C for 5 minutes, followed 35 °C for 5 minutes, and incubation on ice for 2 minutes. 6 µl of a master mix containing 4 µl 5x superscript III buffer (Invitrogen), 17 mM DTT and 1.7 mM dNTPs was added to each reaction, followed by incubation at 55 °C for 1 minute. 200 units of Superscript III (Invitrogen) was added to each reaction, followed by incubation at 55 °C for 30 minutes. 1 µl of 4M NaOH and 32 µl acid stop mix (4:25 (v/v) 1 M unbuffered Tris-HCl and stop dye (85% formamide, 0.5 TBE, 50 mM EDTA, pH 8.0, containing bromophenol blue and xylene cyanol tracking dyes) was then added to each reaction followed by incubation at 95 °C for 5 minutes. Reactions were analysed by 7.5% urea-acrylamide gel electrophoresis at 180V for 90 minutes. Gels were fixed for 1 hour in fix solution (12% methanol, 10% acetic acid) and dried for 1 hour on a gel drier. Gels were exposed to a Fujifilm imaging plate for 36 hours and visualised using a Fujifilm FLA-5000 phosphoimager illumination laser 635.

2.8.2 RNA folding and NMIA Treatment

20 µg of uncapped ZIKV 1-772 or ZIKV IC RNA was heated to 95 °C and incubated on ice for 2 minutes, before addition of 45 µl 3.3 x RNA folding buffer (330 mM HEPES (pH 8.0), 20 mM $MgCl_2$ and 330 mM NaCl) and 80 units of RNase inhibitor (Promega). Reactions were then made up to a volume of 150 µl with nuclease free water and incubated at 28 °C or 37 °C for 30 minutes, before being divided into two 72 µl fractions designated "+" for NMIA treated reactions and "-" for negative control reactions. 8 µl of 50 mM NMIA was added to the "+" fractions and 8 µl of neat DMSO was added to the "-" fractions, before incubation at 28 °C or 37 °C for 50 minutes. RNA was precipitated by addition of 0.1 volumes of 5 M NaOAc, 2 µl 100 mM EDTA, 20 µg Glycogen, made up to a volume of 375 µl with ice cold absolute ethanol and incubated at -80 °C for 30 minutes. RNA was pelleted by centrifugation at 16,000 x g for 30 minutes at 4 °C, washed with 70% ethanol and resuspended in 10 µl of 0.5 x TE buffer pH 8.0 (Ambion).

2.7.3 SHAPE Primer extension

1 µl of 10 µM 6-carboxyfluorescein (6-FAM) labelled primer was added to 5 µl of either NMIA treated or DMSO treated ZIKV 5′ RNA fragment and made up to 12 µl with nuclease free water. Reactions were then incubated in a Proflex PCR thermos cycler set at 85 °C for 1 minute, 60 °C for 10 minutes and 35 °C for 10 minutes. A reverse transcription master mix containing 1 x FS buffer, 12 mM DTT, 40 units of RNase inhibitor, 1.3 mM dNTPs, 200 units of Superscript III, made up to a volume of 8 µl with nuclease free water, was preheated to 37 °C and added to each reaction before incubation at 52 °C for 30 minutes. Then, 1 µl of 4 M NaOH was added to each reaction, followed by incubation at 95 °C for 3 minutes. 2 µl 2 M HCl was then added to each reaction and reactions were incubated for 2 minutes on ice. DNA was then ethanol precipitated as previously described, washed with 70% ethanol, resuspended in 40 µl deionized formamide and incubated to 65 °C for 10 minutes, vortexed and stored at -80 °C.

2.8.4 SHAPE Ladder preparation

For ddATP sequencing ladders, reactions were performed as for NMIA and DMSO reactions using 10 µg of ZIKV 1-754 per reaction, with the following additional changes: 1 µl of 10 µM Hexachloro-Fluorescein labelled primer was added in place of 6-FAM labelled primer; 1 µl 10mM ddATP was added to the reverse transcription master mix.

2.8.5 Sample preparation and data analysis

Prior to analysis, 10 µl of sequencing ladder was added to NMIA and DMSO reactions. Reactions were analysed by capillary electrophoresis fragment size analysis using the DNA Sequencing & Services provided by the University of Dundee. Capillary electrophoresis sequencing data was analysed using the QuSHAPE software package using default settings (Weeks laboratory, University of California, [291]). Briefly, the NMIA treated (+ve) signal data is subtracted from the DMSO (-ve) signal, in order to account for non-specific terminations. Data is then ranked based on intensity and the top 2% of values are discarded as outliers. The total reactivity is normalised to the average of the next 8% of products, which allows different data sets to be compared. Three biological repeats were carried out using both SHAPE primers and the data from the two primers was combined, resulting in a file contain reactivity values for the entire sequence of interest. NMIA reactivity data was then overlaid onto to a structural model drawn using the VARNA software package based on thermodynamic predictions made using “RNAstructure” prediction software Version 6.2 (Mathews lab, University of Rochester, [292]) and the mfold RNA folding form web server version 2.4 [293].

Chapter 3 LNA-162 inhibits CHIKV genome replication

3.1 Introduction

Since the RNA replication elements (RREs) within the CHIKV 5' genome region are known to be essential for replication and are distinct from host RNA structures, RREs within the CHIKV genome have the potential to act as targets for antiviral agents. Modified oligonucleotides, complimentary to primary nucleotide sequences within individual RNA structures, were designed to disrupt both RNA-RNA and RNA-protein interactions essential for the function of these RREs and therefore inhibit virus replication. This hypothesis was investigated using locked nucleic acid oligonucleotides, as these are commercially available and have several properties which make them ideal for this application. Compared with DNA oligonucleotides, LNA oligonucleotides bind to target nucleotide sequences with greater specificity, higher melting temperatures (T_m), are resistant to endonuclease activity increasing half-life in the cellular environment and can be design such that activation of RNase H activation is not induced [282,283].

RREs in alphaviruses have not previously been targeted, however LNA oligonucleotides have previously been used as antiviral agents in a number of divergent RNA viruses. A conserved RRE within the 3' region of the HCV genome, termed SL9266, forms a range of pseudoknot and kissing loop interactions. Targeting SL9266 with LNA oligonucleotides caused inhibition of HCV translation and significant inhibition of replication, in both sub-genomic replicon and infectious virus systems [294]. The HIV genome is a homo-dimer of two RNA strands, which require the interaction of RNA stem-loops termed dimerization initiation sites (DIS). LNA oligonucleotides have been used to block DIS interactions, resulting in inhibition of HIV replication [295]. In addition, the HIV TAR element, an RRE essential for transcription of the integrated genome, has been targeted using LNA oligonucleotides [296,297]. In Ebola virus, LNA oligonucleotides targeting conserved regions close to the translation start sites of the nucleoprotein, viral protein 24 and viral protein 35 genes were effective at reducing protein translation in an *in vitro* system and in human cells [298]. In addition, LNA oligonucleotides have been used to inhibit virus replication by targeting host micro RNAs (miRNA). MiR-122 is a liver-specific miRNA which functions in the HCV lifecycle. Miravirsen, an LNA inhibitor which binds to a stem-loop structure within miR-122, reached phase 2 clinical trials for treatment of HCV [299,300].

From analysis of published literature, targeting RREs within the CHIKV genome has not been investigated as an antiviral approach and nothing has been published regarding the relative availability of individual RNA elements during active virus replication. In this chapter, binding of LNA oligonucleotides to specific RREs within the 5' UTR and adjacent nsp1 coding region of CHIKV is characterised, in the context of a folded RNA molecule. Using a of variety of approaches the effects of antisense LNA oligonucleotides on CHIKV replication are determined, validating the concept that such structures are available for binding and can be targeted in the context of the infectious CHIKV lifecycle. In addition, the lifecycle stage at which an antisense LNA oligonucleotide impairs CHIKV replication was determined.

3.1.1 CHIKV RNA structures

Using a combination of selective 2′ hydroxyl acylation analysed by primer extension (SHAPE) constrained thermodynamic modelling and structure based reverse genetics approaches, members of the Tuplin laboratory previously characterised RREs within the first 300 nts of the CHIKV genome, demonstrating that they are essential for efficient virus genome replication in both mammalian and mosquito hosts [2]. Two RREs (SL3 and SL47) were identified within the 5′ UTR and five are present in the adjacent nsp1 encoding region of ORF1 (SL85, SL102, SL165, SL194 and SL246) [2]. These RREs function during genome replication within the positive sense genomic RNA, at the stage of negative strand synthesis.
SL47 is highly conserved structure which functions in both human and mosquito cells, in a structure dependent manner. SL47 has been proposed to play a role in CHIKV template specificity during the initiation of negative-strand synthesis [2,301]. SL85 and SL102 are involved in replication in mammalian cells again, in a structure dependent manner. SL85 may also form potential local pseudoknot interactions, between SL85 apical region and an adjacent upstream genome region, overlapping with the AUG start codon [2].
SL165 and SL194 correspond to the 51 nt nsP1 conserved sequence element (CSE), which is highly conserved in structure and sequence across the alphavirus genus [109–111,301]. In CHIKV, the 51 nt element enhances replication in human cells but does not have significant effect on virus replication in *A. albopictus* derived cells - in contrast to VEEV and SINV where the 51 nt element enhances replication in both vertebrate and invertebrate derived cells. Reverse genetic studies demonstrate that SL165 functions through a structure-dependent and sequence independent mechanism. In contrast, the function of SL194 is dependent on both the primary sequence of its terminal unpaired loop and the structure of the heteroduplex stem - with the terminal loop acting as a CHIKV-specific signal motif [2].

3.1.2 Rational for the design of LNA oligonucleotides

In order to disrupt the function of RNA structures in CHIKV, a panel of antisense LNA oligonucleotides was designed to bind specifically to primary sequence within RREs. Since the RNA structures within the 51 nt CSE, SL165 and SL194, are highly conserved and known to be essential for CHIKV replication, it was reasoned that these structures would be good candidates for LNA targeting. LNA-162 targets the top-loop and duplex of SL165, targeting nucleotides 162-181; LNA-202 targets the top-loop and duplex of SL194, targeting nucleotides 202-220 (Fig. 3.1).

LNA-225 targets nucleotides 225-244, targeting a conserved region of the genome between SL194 and SL246 (CR). In order to control for non-specific effects on virus replication, associated with LNA hybridisation, LNA-760 (targeting nucleotides 760-779 within a non-structured region of the nsp1 coding region) was designed as a negative control.

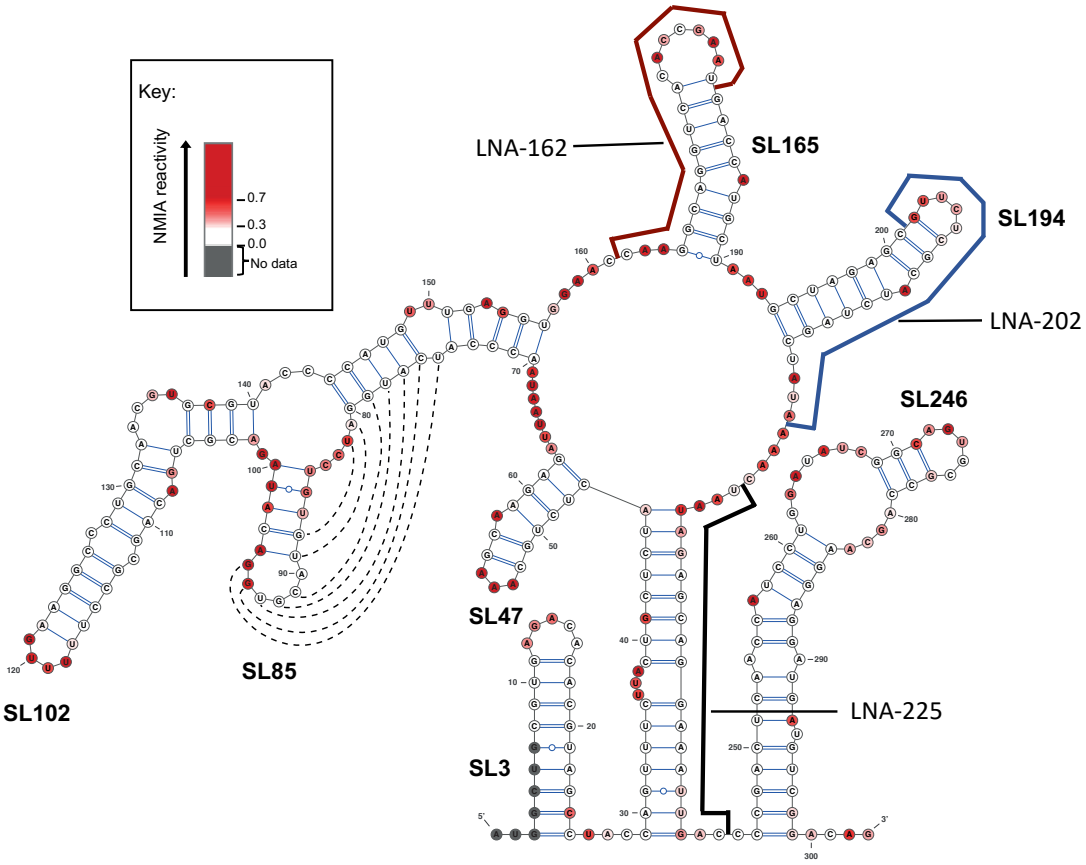


Figure 3.1 Binding regions for targeted antisense LNA oligonucleotides against CHIKV
37 °C SHAPE reactivities for individual nucleotides overlaid onto a 37 °C thermodynamically derived model of RNA folding, generated using SHAPE-directed constraints. Binding regions for LNA-162 (Red line), LNA-202 (Blue line) and LNA-225 (Black line). The CHIKV SHAPE map was produced by Kendal et al. [2].

3.1.3 Factors associated with LNA design

LNA oligonucleotides are a powerful tool for molecular research, however, in order to be used effectively a number of design factors must be considered. LNA nucleotides contain an additional bridge connecting the 2' oxygen and 4' carbon, locking the nucleotide in an ideal conformation for Watson-Crick base pairing, and in an LNA mixmer oligonucleotide, LNA nucleotides are interspaced within a DNA backbone. LNA oligonucleotides are designed to have a high affinity for the intended target, whilst limited self-hybridisation and secondary structure formation. The position of LNA nucleotides within the DNA backbone sequence, length and sequence of the oligonucleotide are all highly important for effective design.

In order to maximise binding affinity, LNA oligonucleotides were optimised using the LNA T_m predictor (Qiagen). Compared with DNA oligonucleotides, LNA oligonucleotides have different melting properties and thus specialised algorithms are required for T_m prediction. T_m Predictor (Qiagen) uses a modified nearest-neighbour thermodynamic model combined with data from measurements of oligonucleotides ranging for 12-27 bp to predict the T_m of an LNA oligonucleotide at a neutral pH. The predicted T_m of LNA oligonucleotide binding is dependent on the type of nucleic acid being targeted, differing between DNA or RNA targets.

Anti-CHIKV LNA oligonucleotides were designed to maximise the T_m for RNA binding, aiming for an RNA T_m of between 75-90 °C.

Due to the improved properties of LNA nucleotides, secondary structures can form easily and LNA oligonucleotides may be self-complementary. Using the LNA Oligo optimiser (Qiagen) it is possible to predict the formation of such elements, which informs the design of the LNA oligonucleotide. The LNA Oligo optimiser hybridisation score predicts the melting temperature for interactions between different molecules of the same LNA oligonucleotide (i.e. the temperature at which dimers or chains of the LNA oligonucleotide will not be stable) and the secondary structure score predicts the melting temperature for secondary structures within the LNA oligonucleotide. LNA oligonucleotides were designed to have a hybridisation and secondary structure scores score of less than 28 °C, to ensure that LNA oligonucleotides were viable for use in cell culture experiments. The inclusion of three 5' and 3' terminal LNA nucleotides are another design consideration. This prevents nuclease degradation in a cellular environment and therefore improves the half-life of oligonucleotide. As an addition consideration, LNA oligonucleotides were designed to prevent activation of RNase H, which will degrade RNA/DNA hybrids. LNA oligonucleotides were therefore deigned to have no more than four consecutive DNA bases, preventing RNase H activation.

3.2 LNA oligonucleotides bind to a folded RNA molecule

3.2.1 Analysis of LNA panel binding to CHIKV RNA

Although LNA oligonucleotides were designed to bind the primary RNA sequence, and theoretically should be able to bind to the intended sequence, secondary and tertiary RNA interactions may prevent LNA oligonucleotides from accessing the intended sequence. For example, the LNA binding site may be inaccessible due to a pseudoknot interaction. It is therefore important to determine that LNA oligonucleotides are able to bind RREs in the context of a folded RNA molecule, in a setting where RNA and LNA concentration can be controlled at a high level. To compare the ability of LNA oligonucleotides to bind their intended RNA targets, native electromobility shift assays (EMSAs) were utilised. Unlike denaturing gel electrophoresis, native EMSAs allows both RNA and LNA: RNA complexes to maintain a folded state on a gel. The CHIKV 1-337 RNA is 337 nts in length and therefore separates at a much higher position on the gel compared with the LNA oligonucleotides (around 20 nts) therefore, LNA-RNA complexes will separate at a different position on a gel compared with free LNA. This makes it possible to distinguish between LNA oligonucleotides which can bind the CHIKV RNA and LNA oligonucleotides which are not able to bind the RNA.

LNA oligonucleotides were 5´end labelled using γ-^{32}P ATP and visualised on a native PAGE gel (Fig. 3.2). Radiolabelled LNA were incubated with folded CHIKV 1-337 RNA, using a 1:2.5 ratio of LNA to RNA. EMSAs confirmed that LNA-162, -202 and -225 are able to bind their intended RNA targets, whereas the control LNA-760 and scrambled-LNA were unable to bind the RNA (Fig. 3.3A). This indicates that LNA oligonucleotides are able to specifically bind to their intended target sequences and that non-specific interactions do not play a significant role in LNA-RNA interactions. If an LNA is able to bind the RNA target, the resulting LNA/RNA complex will separate at a higher position on the gel compared with unbound ^{32}P labelled LNA, relative to the length of the target RNA. Since the bound LNAs are only around 20 nucleotides in

length, LNA/RNA complexes should separate to a similar position on a native gel to ^{32}P labelled CHIKV 1-337 RNA molecules. As a control for the position of the shift, a ^{32}P labelled CHIKV RNA molecule was loaded alongside the samples containing ^{32}P labelled LNA.

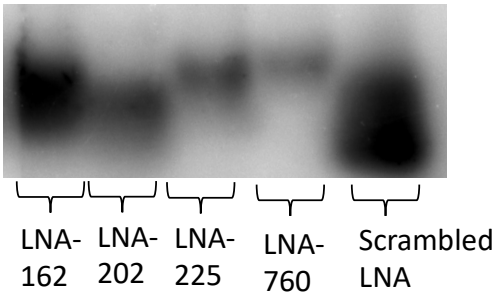


Figure 3.2 Visualisation of LNA oligonucleotides
LNA oligonucleotides were 5´end labelled using γ-^{32}P ATP and visualized on a 7.5% native PAGE gel. Differences in the levels of γ-^{32}P signal can be attributed to efficiency of labelling and ethanol precipitation, which was used to purify the labelled LNA oligonucleotides. LNA oligonucleotides are either 19 nucleotides (LNA-202, Scrambled-LNA) or 20 nucleotides (LNA-162, LNA-225, LNA-760). Since LNA oligonucleotides were separated using a native gel, it is possible that the folding conformations of the different LNA oligonucleotides influences the rate of migration through the gel.

To investigate the specificity of the LNA-162 binding interaction, a dose response experiment was performed, using increasing concentrations of RNA at ratios 1:0, 5:1, 1:1, 1:2.5, 1:5 LNA: RNA, using 1.25 pmol of LNA (Fig. 3.3B). As more RNA was added, the percentage of LNA in complex increased, demonstrating a dose-dependent effect (Fig. 3.3C). This, combined with the lack of binding seen for negative control LNA oligonucleotides, demonstrates that the binding effect of LNA-162 to the CHIKV RNA is a specific interaction.

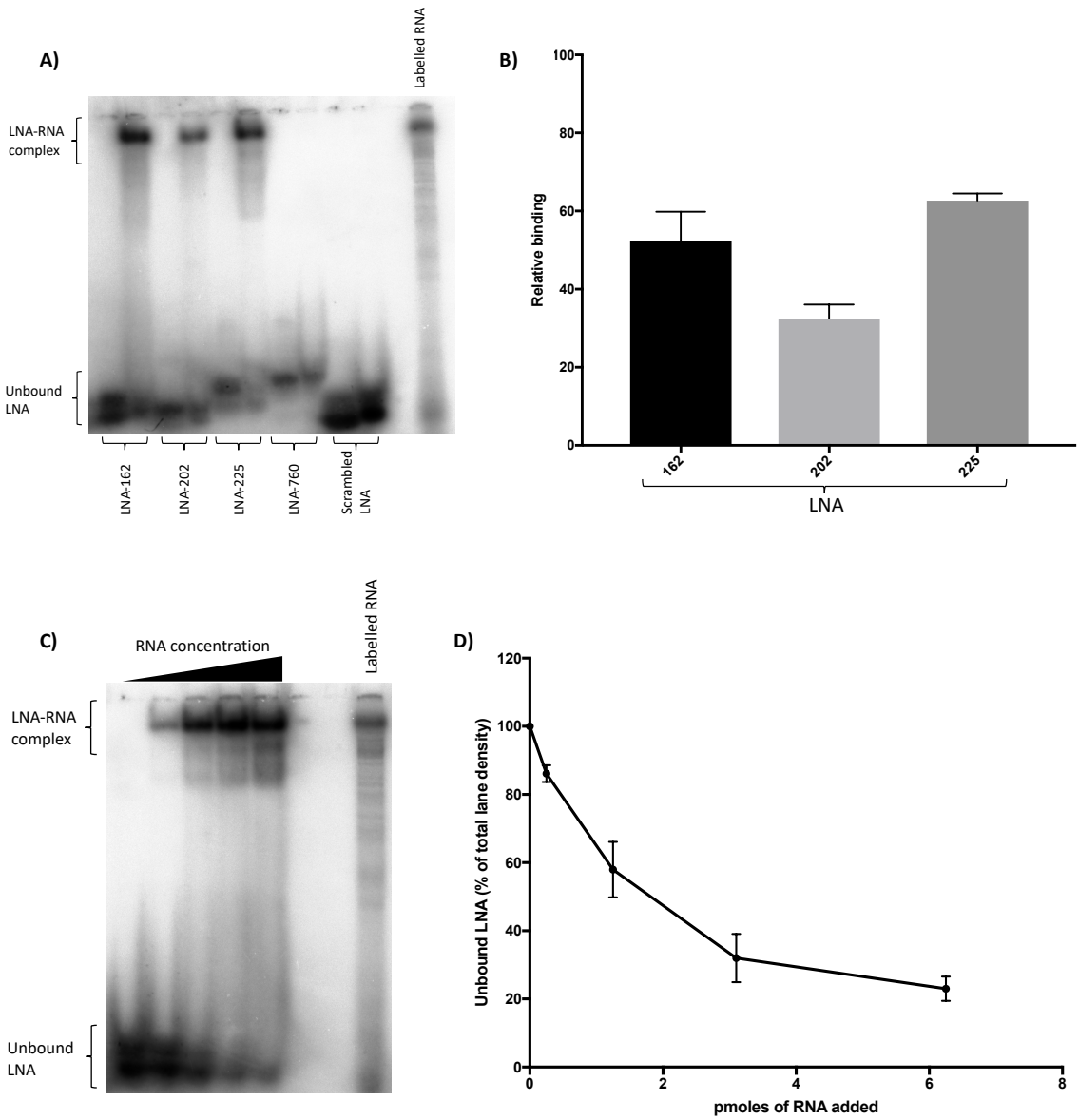


Figure 3.3 LNA-RNA binding assays for CHIKV RNA 5' region

A) *LNA oligonucleotides were end labelled with ^{32}P prior to incubation with folded CHIKV 1-337 RNA. LNA-162, -202 and -225 are able to bind target RNA structures. Both control LNA-760 and scrambled-LNA showed no binding.* ***B)*** *Binding efficiencies for LNA oligonucleotides were quantified by densitometry, as a percentage of total lane density, normalised to the corresponding lane containing no RNA (n = 2).* ***C)*** *^{32}P labelled LNA-162 was incubated with increasing concentrations of folded CHIKV RNA. Increasing the concentration of RNA added results in a greater shift of labelled LNA. For both* ***A*** *and* ***C****, ^{32}P labelled CHIIKV 1-337 RNA was included as a size marker to indicate that the ^{32}P signal from the LNA had shifted to the correct position. Since the LNA/RNA complex has only marginally increased molecular weight compared with the RNA alone, both migrate a similar distance on a native EMSA.* ***D)*** *Binding efficiency for LNA-162 was quantified using densitometry as relative levels of levels of unbound LNA, shown as a percentage of total lane density (n = 3). Error bars represent standard error of the mean.*

3.3 LNA-162 inhibits infectious CHIKV

3.3.1. LNA-162 inhibits CHIKV

Having shown that LNA oligonucleotides were able to bind their target in a native EMSA, the effects of LNA transfection on infectious virus replication was next investigated. LNA-162, -202 and -225 were designed to bind functional RREs SL165, SL194 and the CR respectively. The control LNA-760 was designed to bind a non-structured region of the nsp1 coding region, to act as a negative control for binding to the RNA sequence. A scrambled LNA was also included as a control for transfection of LNA oligonucleotides into cells. Huh7 cells were infected with CHIKV (MOI 1) and incubated for 3 hours to allow the formation of replication spherules. 600 nM LNA oligonucleotides were then transfected into the cells and CHIKV titre was assessed at 24 hours post infection (Fig 3.4B). Compared with mock transfected cells, transfection of Scrambled-LNA, LNAs-202, -225 and -760 had no significant effect on released virus titre. In contrast, transfection of LNA-162 significantly reduced the released CHIKV titre (P values less than or equal to (P ≤) 0.05), representing an approximately 80% reduction in CHIKV titre compared with transfection of scrambled-LNA. To investigate the effects of higher concentrations of LNA-162, CHIKV infected Huh7 cells (MOI 1) were transfected with 3 µM LNA-162, which again resulted in a significant reduced viral titre (P ≤ 0.0001) (Fig. 3.4C). LNA-162 was assessed for cellular toxicity, with a 20% decrease in cell viability deemed as the maximum acceptable decrease in cell viability. At 6 hours, transfection of 10 µM LNA-162 resulted in a 20% decrease in cell viability compared with untreated cells whereas, at 24 hours, transfection of 10 µM LNA-162 resulted in a 30% decrease in cell viability compared with untreated cells (P ≤ 0.01). Transfection of 5 µM LNA-162 resulted in a 10% decrease in cell viability at 6 hours and a 20% decrease in cell viability at 24 hours (appendix Fig. 1). Consequently, concentrations of LNA-162 equal to or less than 5 µM were deemed acceptable for use in cellular assays. It was therefore concluded, that since transfection of equivalent concentrations of scrambled-LNA did not affect CHIKV replication, LNA-162 is able to specifically interact with CHIKV during the virus lifecycle and disrupt the virus replication process.

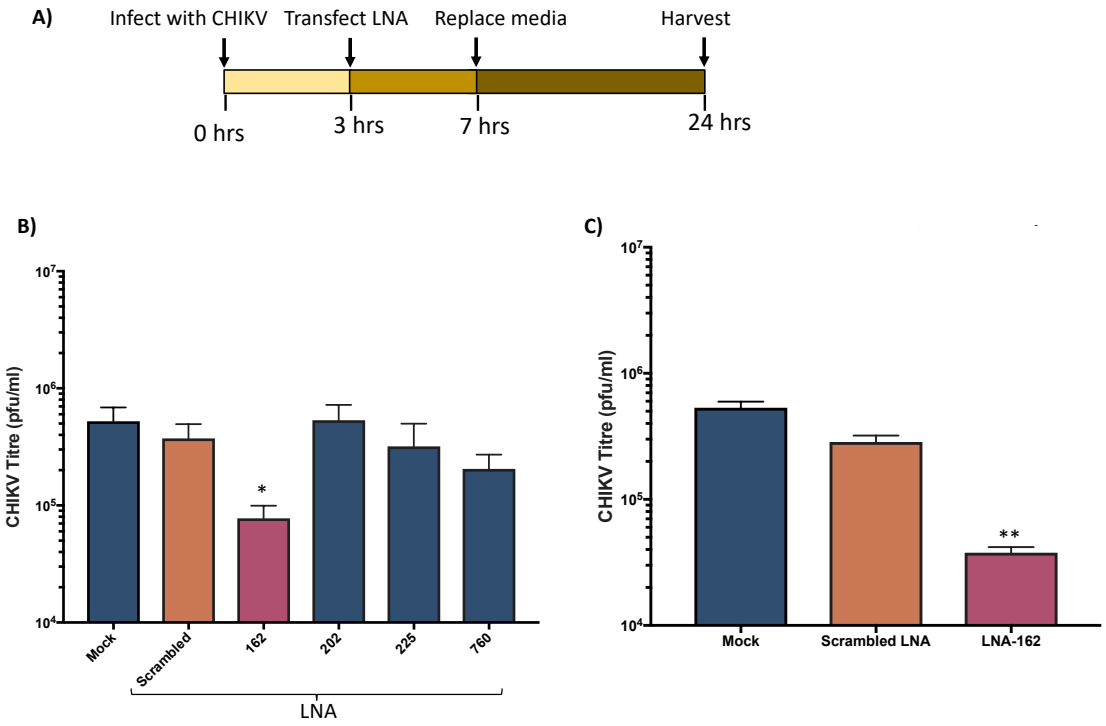


Figure 3.4 LNA-162 inhibits CHIKV

A) *Schematic of experimental design.* ***B)*** *Huh7 cells were infected with CHIKV (MOI 1) and then transfected with 600nM LNA oligonucleotides. Supernatants were harvested 24 hours post infection and virus titre was determined by plaque assay.* ***C)*** *Huh7 cells were infected at an MOI of 1, before transfection with 3 µM LNA-162 or Scrambled LNA. Supernatants were harvested 24 hours post infection and virus titre was determined by plaque assay. * = P ≤ 0.05, ** = P ≤ 0.0001 (Two-tailed T-test vs scrambled LNA) (n = 3). Error bars represent standard error of the mean.*

Having shown that LNA-162 can inhibit CHIKV, the effect of transfecting a range of different concentrations of LNA-162 on CHIKV replication was investigated (Fig. 3.5). Transfection of 3 µM, 1.25 µM and 600 nM LNA-162 transfection resulted similar levels of significant inhibition ($P \leq 0.001$). Transfection of 300 nM LNA-162 showed significant inhibition, but to a lesser extent than the higher concentrations used ($P \leq 0.005$). Concentrations below 300 nM did not significantly reduce viral titre compared with scrambled LNA.

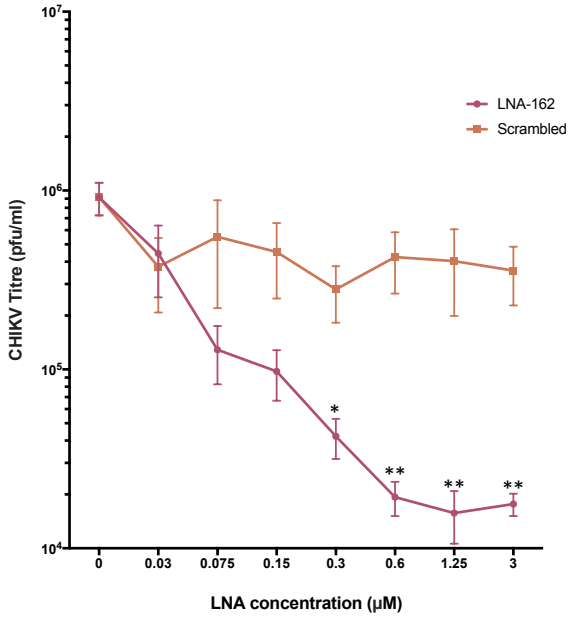


Figure 3.5 LNA-162 and scrambled LNA dose response curve

*Huh7 cells were infected (MOI 1) and cells were transfected with different concentrations of LNA-162 or scrambled-LNA at 3 hours post infection. Supernatant was collected at 24 hours post infection and virus titre was determined by plaque assay. *= P ≤ 0.005, ** = P ≤ 0.001, (Two-tailed T-test vs scrambled LNA) (n = 3). Error bars represent standard error of the mean.*

Having shown that LNA-162 is able to inhibit CHIKV replication the effect of transfecting 3 µM LNA-162 at different stages of the lifecycle was investigated. Huh7 cells were infected with CHIKV (MOI 1) and then transfected with LNA-162. Cell supernatant was collected at 7, 12, 16, 24 and 48 hours post infection (Fig. 3.6). At all of the time points tested, LNA-162 transfection caused a significant reduction in CHIKV titre compared with transfection of scrambled-LNA. At 7 hours, a small but significant difference ($P \leq 0.01$) was observed, which reflects the low amounts of virus present at this time point. At 12 hours post infection, transfection of LNA-162 produced a significant reduction in CHIKV titre ($P \leq 0.05$). By 16 and 24 hours post infection, LNA-162 transfected cells showed an 85-90% reduction in virus titre compared with scrambled-LNA transfection ($P \leq 0.001$). By 48 hours, the effect on CHIKV replication was slightly diminished, which may reflect the half-life of the LNA in the cells, but remained significant ($P \leq 0.05$).

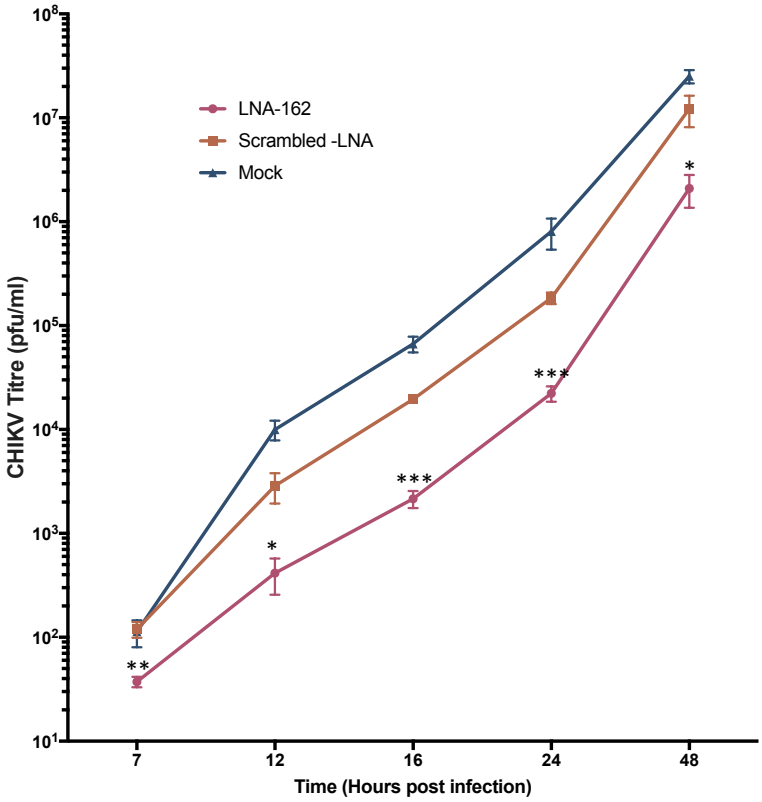


Figure 3.6 CHIKV growth curve from 7 to 48 hours post infection
*Huh7 cells were infected with CHIKV (MOI 1) and incubated for 3 hours before transfection with 3 µM of either scrambled-LNA or LNA-162. Released CHIKV titre was quantified by plaque assay. * = P ≤ 0.05, ** = P ≤ 0.01, *** = P ≤ 0.001 (Two-tailed T-test vs scrambled LNA) (n = 3). Error bars represent standard error of the mean.*

Additionally, another experiment was performed, where Huh7 cells were first transfected with 3 µM LNA-162 and then infected with CHIKV (MOI 10) (Fig. 3.7). At 6 hours post infection, there was no significant difference between LNA-162 and scrambled transfected cells however, by 24 hours post infection, LNA-162 transfected cells produced significantly less CHIKV compared with Scrambled-LNA (P ≤ 0.0001), showing comparable levels of inhibition to experiments where the LNA was added post-infection. The lack of difference between conditions at 6 hours may be a result of the experimental set up, where CHIKV replication has not yet begun in earnest and the viral titre reflects high levels of input virus, therefore LNA-162 has not affected replication.

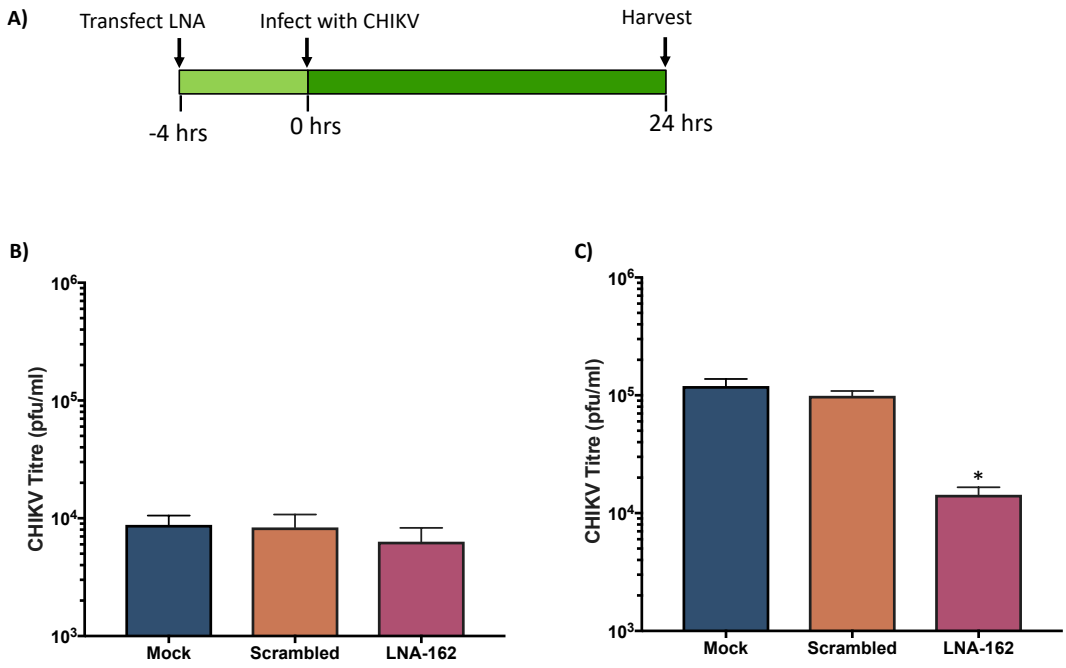


Figure 3.7 CHIKV infection of LNA transfected cells.
A) *Schematic of experimental time scale.* ***B-C)*** *Huh7 cells were transfected with 3 µM LNA-162 or Scrambled LNA and infected with CHIKV (MOI 10). CHIKV was harvested at 6* ***(B)*** *and 24* ***(C)*** *hours post infection and virus titres were determined by plaque assay. * = P ≤ 0.0001 (Two-tailed T-test vs scrambled LNA) (n = 3). Error bars represent standard error of the mean.*

3.3.2 Transfection of LNA-162 into CHIKV infected cells reduces viral protein expression

To investigate level of CHIKV protein expression, Huh7 cells were infected with CHIKV (MOI 1) and then transfected with 3 µM LNA-162 or scrambled-LNA. Cells were lysed at 24 hours post infection and protein expression was analysed by western blot (Fig. 3.8). Huh7 cells transfected with LNA-162 showed a marked reduction in expression of both the structural capsid protein and the non-structural nsp1 protein. This evidence supports the findings of the previous infectious CHIKV experiments, and supports the conclusion that LNA-162 is able to significantly impair CHIKV replication.

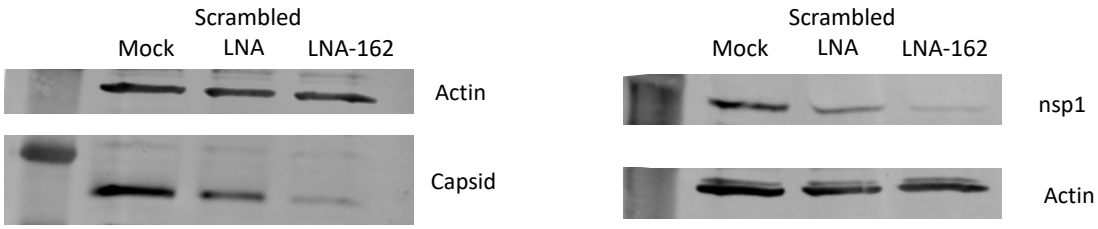


Figure 3.8 LNA-162 inhibits CHIKV protein expression.
Huh7 cells were infected with CHIKV (MOI 1) and transfected with 3 µM LNA-162, scrambled-LNA or mock transfected at 3 hpi. Intracellular expression levels of CHIKV capsid and nsp1 were determined by western blot at 24 hours post infection. Membranes were stained with anti-actin as a loading control to ensure equivalent amounts of cell lysate were loaded into each well. Antibodies used as follows: primary anti-capsid (1:1000, rabbit polyclonal, in-house), anti-nsp1 (1:1000, rabbit polyclonal, in-house) and anti-actin (clone AC-15, mouse monoclonal, Sigma) and secondary antibodies (IRDye® 800CW Donkey anti-Mouse; IRDye® 680LT Donkey anti-Rabbit; Li-Cor).

3.3.3 Viral genome copies are reduced in cells transfected with LNA-162

In order to begin narrowing down the lifecycle stage that LNA-162 is affecting, the CHIKV genome copy number in infected cells was quantified. Huh7 cells were infected with CHIKV (MOI 1), and transfected with 3 µM LNA-162 or scrambled LNA. Total RNA was extracted at 24 hours post infection. CHIKV genome copies were quantified by qRT-PCR (Fig. 3.9). Cells transfected with LNA-162 had significantly fewer CHIKV genome copies compared with Scrambled-LNA transfected cells ($P \leq 0.0005$), indicating that LNA-162 may affect the genome replication stage of the lifecycle. CHIKV supernatant was collected for the wells where RNA was extracted, as a comparison of relative levels of inhibition. Plaque assay results showed a high degree of similarity to qRT-PCR results ($P \leq 0.005$), confirming the validity of the qRT-PCR results and further indicating that LNA-162 is an effective inhibitor of CHIKV replication.

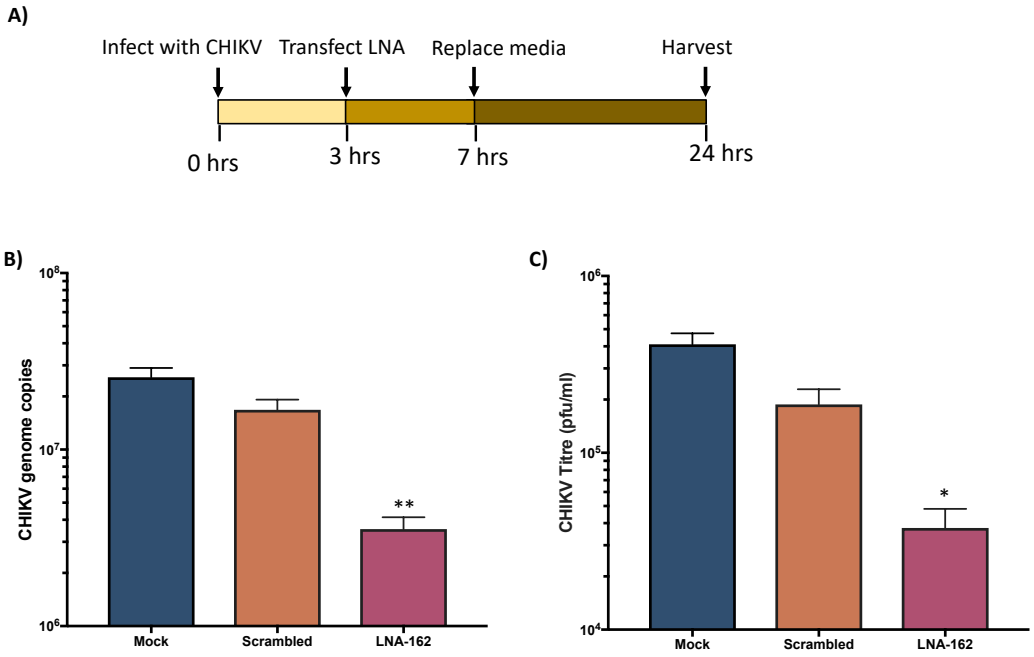


Figure 3.9 CHIKV infection of LNA transfected cells

A) *Diagram of experimental workflow* ***B)*** *Total RNA was extracted from Huh7 cells infected with CHIKV (MOI 1) and transfected with 3 µM of either LNA-162 or scrambled-LNA. RNA was quantified by reverse transcription followed by qPCR (n = 3).* ***C)*** *Released CHIKV collected from the same wells used for qRT-PCR experiments was collected at 24 hours post infection and quantified by plaque assay. * = P ≤ 0.005, ** = P ≤ 0.0005 (Two-tailed T-test vs scrambled LNA) (n = 3). Error bars represent standard error of the mean.*

3.4 LNA-162 Inhibits the sub-genomic replicon

3.4.1 LNA-162 inhibits the CHIKV sub-genomic replicon in a dose-dependent manner

In order to further investigate which stage of the virus lifecycle was being impaired by LNA-162, a CHIKV sub-genomic replicon system (CHIKV SGR) was utilised. CHIKV SGR expresses the non-structural CHIKV proteins (nsp1-4) from the first ORF, but expresses a firefly luciferase reporter from the second ORF in place of the viral structural proteins (Fig. 3.10). Since the replicon contains all the viral proteins associated with genome replication, this system can be used as a model for viral genome replication.

The effect of LNA oligonucleotides on CHIKV replicon replication was assayed in Huh7 cells at 6 hours post transfection (Fig. 3.11). For all data presented, LNA oligonucleotides were co-transfected with CHIKV SGR and a 5′ capped *Renilla* luciferase RNA, using lipofectamine 2000. LNA oligonucleotides were transfected into Huh7 cells at a range of concentrations from 20 nM to 250 nM, in order to investigate dose-dependent effects on CHIKV replication. At the lowest concentrations of LNA transfected, 20 nM and 30 nM, LNA-162 did not have a significant effect on CHIKV genome replication. The reduction in replication became significant when LNA-162 was co-transfected at concentrations above 50 nM LNA-162 and all subsequent higher concentrations of LNA-162 tested significantly impaired CHIKV genome replication.

In contrast, co-transfection of a scrambled-LNA control at equivalent concentrations to LNA-162 did not affect replicon replication up to a concentration of 100 nM. At concentration above 100 nM a small amount of inhibition was observed however, at these concentrations there was a reduction in the levels of luciferase activity detected from the *Renilla* expressing transfection control. This may suggest that at concentrations of LNA above 100 nM, transfection efficiency had been reduced and therefore explains the lower firefly luciferase levels detected at these concentrations. In order to account for the reduction in transfection efficiency, the firefly expression signal was normalised to the *Renilla* luciferase signal. In the normalised data, transfection of LNA-162 at concentrations above 30 nM resulted in significant inhibition of CHIKV replication, with a greater effect on replication seen at higher concentrations of LNA-162, demonstrating a dose depend effect. Significant inhibition of CHIKV replication was seen at 30 nM ($P \leq 0.05$), 50 nM ($P \leq 0.01$), 75-200 nM ($P \leq 0.001$) and 250 nM ($P \leq 0.01$). By contrast, co-transfection of equivalent concentrations of scrambled-LNA produced no effect on CHIKV genome replication, demonstrating that LNA-162 acts through a specific mechanism and that the effect on CHIKV genome replication is not due to non-specific interactions. In addition, transfection of 100 nM LNA-202, LNA-225 and LNA-760 had no effect on CHIKV genome replication (appendix Fig. 2)

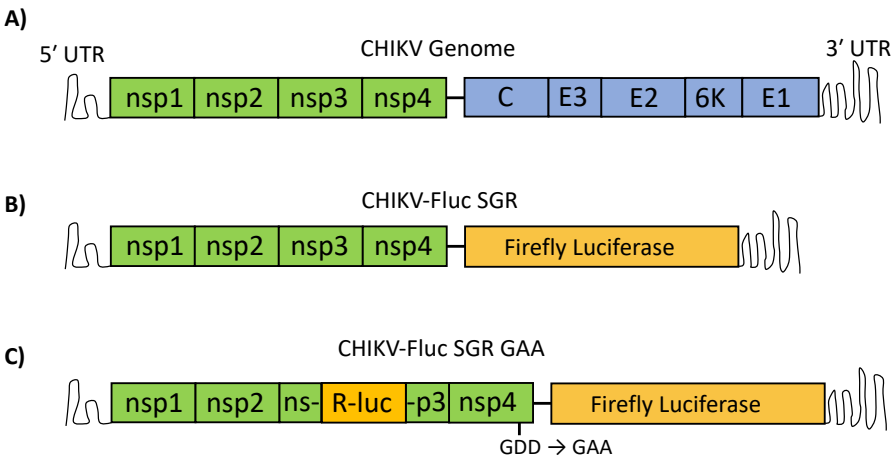


Figure 3.10 CHIKV sub-genomic replicon schematics

A) *Full length CHIKV genome.* ***B)*** *Firefly luciferase expressing CHIKV replicon (CHIKV-Fluc SGR), where the second OFR is replaced with a Firefly luciferase coding region. Luciferase expression can be used as a meaure of CHIKV genome replication.* ***C)*** *A replication deficient CHIKV replicon (CHIKV-Fluc SGR GAA) where an Renilla luciferase (R-luc) sequence is inserted into the nsp3 coding region. Levels of Renilla luciferase expression can be used as a measure of CHIKV translation.*

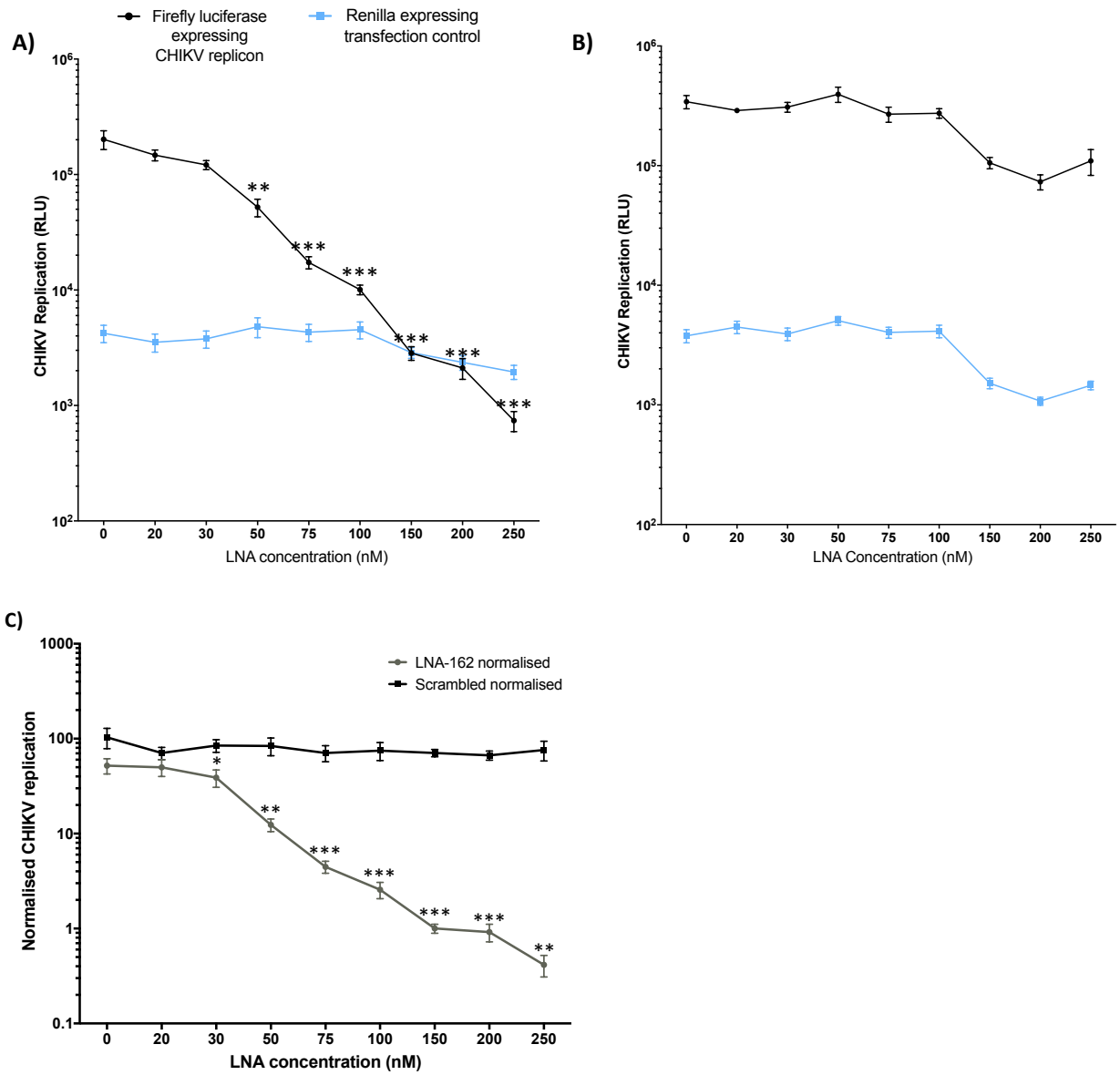


Figure 3.11 LNA- 162 inhibits CHIKV genome replication

LNA-162 inhibits CHIKV genome replication. ***A-B)*** *Huh7 cells were co-transfected with CHIKV SGR, a Renilla luciferase expressing transfection control and either LNA-162* ***(A)*** *or scrambled-LNA* ***(B)*** *at concentrations ranging from 1-250nM. Cells were lysed at 6 hours post infection and levels of firefly and Renilla luciferase were determined.* ***C)*** *Huh7 cells were co-transfected with CHIKV SGR, a Renilla luciferase expressing transfection control and either LNA-162 or scrambled-LNA at concentrations ranging from 1-250nM. Cells were lysed at 6 hours post infection and levels of firefly and Renilla luciferase were determined. Firefly luciferase signal was normalised to Renilla transfection control signal. * = P ≤ 0.05, ** = P ≤ 0.01, *** = P ≤ 0.001. (Two-tailed T-test vs scrambled LNA) (n = 3). Error bars represent standard error of the mean.*

3.4.2 The effect of LNA-162 on replicon translation

Since LNA-162 binds to the CHIKV RNA, it is possible that the oligonucleotide could interfere with the process of translation. In order to investigate this, a translation reporter system was utilised. The Fluc-Rluc-(GDD>GAA)-SGR construct contains an active site mutation (GDD>GAA) which prevents replication. Since this construct also contains a *Renilla* luciferase gene within the nsp3 coding region, expression of *Renilla* can be used as a measure of translation. Huh7 cells were co-transfected with Fluc-Rluc-(GDD>GAA)-SGR and either 100 nM LNA-162, scrambled-LNA or no LNA. WT CHIKV SGR was also included in the experiment for comparison. Luciferase expression was measured at 4, 6 and 12 hours post transfection (Fig. 3.12) (4 and 12-hour time points shown in appendix Fig. 3). Transfection of LNA-162 had no significant effect on translation at all time points, indicating that LNA-162 does not affect translation efficiency.

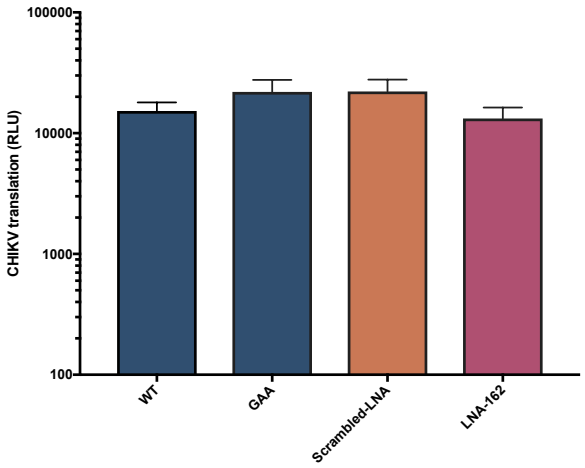


Figure 3.12 LNA-162 does not inhibit CHIKV Translation
The effect of LNA transfection on translation were determined using a replication deficient CHIKV replicon system (CHIKV_Rep(GDD>GAA)) in Huh7 cells. Expression of Renilla luciferase, representing translation, was determined at 6-hours post transfection. (n = 3). Error bars represent standard error of the mean.

3.5 Discussion

There are currently no licenced treatments available which specifically impair CHIKV replication. Since CHIKV RREs are known to be essential for viral genome replication, targeting these RNA structures is one method which could lead to specific antiviral therapies. In order to investigate the availability of RREs for targeting during CHIKV replication, and if disruption of RREs using an agent which acts *in trans* is sufficient to impair replication, antisense LNA oligonucleotides were utilised. Despite the knowledge that CHIKV RREs are required for efficient replication, at the outset of this project it was uncertain whether RREs would be available for binding within the RNA replication complex. Factors such as binding of *trans* activating factors to RREs or tertiary RNA-RNA interactions may prevent LNA binding. In

addition, since RNA replication takes place within membrane bound spherules, LNA oligonucleotides may be unable to access the site of RNA replication.

Of the LNA oligonucleotides investigated in this thesis, one oligonucleotide, LNA-162 (specific to RRE SL165) was shown to inhibit CHIKV genome replication, demonstrating that the RRE SL165 is accessible to *trans* acting agents during virus replication and is a viable target for antiviral agents. Presumably, the oligonucleotide acts by disrupting the essential folding conformation of the stem loop structure or via disruption of essential *trans* activating RNA binding proteins. Alphavirus RNA is replicated in membrane bound replication complexes, which contain both the viral RNA and non-structural proteins [302] and previous studies have indicated that replication complexes are most active in RNA synthesis at around 4 hours post infection [303,304]. LNA-162 was able to significantly inhibit CHIKV replication when transfected into infected cells at 3 hours post infection, indicating that the LNA is able to gain access to the viral RNA within the replication complex. Since alphavirus RNA replication initially takes place at the plasma membrane [305], it is possible that LNA-162 may bind prior to the formation of membrane-bound replication spherules. Electron microscopy images of SINV infected BHK cells shows that early replication complexes are connected with cytoplasmic side of the plasma membrane [306], and LNA oligonucleotides could feasible enter where replication complexes are open to the cytoplasm. In addition to LNA oligonucleotides overcoming entry into the replication complex, individual RREs may be unavailable for binding due to the presence of higher order RNA-RNA interactions or binding of *trans* activating factors, such as cellular or viral proteins.

Inhibition of replication by LNA-162 indicates that SL165 is available for binding within the replication complex and that the hybridisation strength of LNA-162 is likely sufficient to outcompete other factors which could prevent binding. In order to investigate the lifecycle stage at which LNA-162 acts, a sub-genomic replicon system was utilised, which acts as a model for genome replication. Inhibition of the sub-genomic replicon indicated that LNA-162 impairs CHIKV replication at the level of genome replication. It as also confirmed that LNA-162 does not significantly impair translation of ORF-1, indicating that inhibition is due to LNA-162 acting directly on the viral RNA to prevent the function of the RRE SL165, rather than through steric hinderance of the translation machinery.

Interestingly, although both LNA-202 and LNA-225 were shown to bind to the CHIKV genome by EMSA, neither oligonucleotide was shown to significantly impair CHIKV replication in both infectious virus or replicon systems. Since SL194 is known to be essential for genome replication, this would indicate that the lack of antiviral activity is due to the target sequences being inaccessible within the replication complex. Higher order RNA/RNA interactions or *trans* activating RNA binding proteins could block the LNA-RNA when in the context of an infected cell. It may also be possible that different anti-sense LNA oligonucleotides are variable in their ability to enter the replication complex or that these oligonucleotides do not have sufficient binding strength to outcompete *trans* activating factors. Such results demonstrate the importance of investigating antiviral agents in the context of a cellular system, as binding of agents to antiviral agents within an *in vitro* (or *in silico*) system will not necessarily lead to the desired effect when the complexity of the wider virus-cellular interaction is considered.

Since the first use oligonucleotides to inhibit viral replication in the 1970s [280], several antisense oligonucleotide drugs have been approved for use in a range of conditions. For example, Nusinersen is an antisense oligonucleotide used in treatment of spinal muscular atrophy. Nusinersen functions by modulating alternative splicing of the survival of motor neuron 2 gene (SMN2), functionally converting it to the SMN1 gene. This increases the level of the SMN protein, halting disease progression [307,308]. In addition miravirsen, an LNA oligonucleotide treatment targeting host miRNA MiR-122, reached phase II clinical trials for in the treatment of HCV [299]. Although miravirsen did not enter the clinic, likely due to the development of direct acting antivirals [309], this demonstrates the viability of this approach.

Viral RREs make attractive targets for novel anti-viral agents due to the essential roles which they play in viral replication and their specificity to the virus, which may limit off target side effects [2]. LNA oligonucleotides have previously been used to target RNA structures in a number of divergent RNA viruses. In HIV-1, LNA oligonucleotides have been used to block the dimerization initiation site, preventing formation of the genome homodimer [295] and in HCV, LNA oligonucleotides targeting RNA stem-loop structures block interactions between the viral *cis*-acting replication element and 3' non-coding region, causing inhibition of viral translation [294]. Since RREs in CHIKV have been shown here to be viable targets for antiviral agents, it is possible that this approach can potentially be applied to many viruses which contain functional RREs.

In summary, RRE SL165 within the CHIKV genome can be specifically targeted using antisense LNA oligonucleotides, resulting in inhibition of viral genome replication. This demonstrates that the viral replication complex is accessible to oligonucleotides and that the viral RRE SL165 is available for binding within the RNA replication complex, whilst other RREs such as SL194 may be inaccessible. These findings further our understanding of CHIKV RNA replication and provide a rational for the use of antiviral agents targeting RREs in CHIKV. Following on from these experiments, SL165 could provide the basis for further screening of small molecules, for example by using small molecule microarrays, which may improve binding specificity and achieve greater effect on CHIKV replication [310].

Chapter 4 ZIKV SHAPE mapping

4.1 Introduction

RNA structures found within the 5' UTR and adjacent capsid coding region are critically important for the lifecycle of Flaviviruses and have been shown to play important roles in RNA replication [152,206]. Compared with other flaviviruses, ZIKV has 58%-60% nucleotide sequence identify with WNV and DENV [153] and RNA elements such as SLA are often highly conserved between different flaviviruses [152]. The first 498 bp of the ZIKV genome were selected for study, in the context of both full-length genome and truncated ZIKV 1-754 bp transcripts at both 28 °C and 37 °C. There are several reasons for investigating this region. Firstly, the most highly conserved and well-studied Flavivirus RNA structures SLA, SLB and cHP are contained within this region [152,311]. Since functional studies of these structures have been carried out in other, closely related viruses, it is important to establish the similarity of these structures between ZIKV and related viruses, to ensure that the same models of replication can apply. For example, if the ZIKV SLA was structurally divergent to other flaviviruses, then it may function in a different way. Secondly, as RNA structural arrangement is sensitive to temperature variations, RNA structures may form differently at 28 °C compared with 37 °C, reflecting the differing body temperatures of the arbovirus invertebrate and mammalian hosts. Thirdly, this region contains the conserved sequence elements responsible for genome cyclisation, allowing the cyclised state of the RNA molecule to be determined. The ZIKV 1-754 RNA molecule acts as a negative control for genome cyclisation, since it lacks the 3' cyclisation sequences.

Prior to the use of SHAPE mapping, techniques such as enzymatic digestion and chemical probing were used to determine if nucleotides were base-paired or non-base paired. Using the enzymatic method, RNA molecules were first end labelled with a radioactive marker and then partially cleaved, using ribonucleases (RNases) which specifically cleave either single stranded or double stranded RNA. For example, in a native RNA buffer, a single stranded cleaving enzyme (e.g. RNase T1) will digest only single stranded regions of a folded RNA molecule. This causes a specific pattern of digestion with fragments of varying length, and this can be compared to an equivalent RNA molecule sequencing ladder which has been digested in a denaturing buffer (i.e. where all nucleotides can be digested by the single stranded specific enzyme). By resolving both native and denatured reactions on a denaturing polyacrylamide gel, the position of base paired and single stranded nucleotides can be determined, as shown by the methodology published by Nilsen, 2011 [312]. Chemical reagents can be used in a similar manner to interrogate RNA structure. For example, ethylnitrosourea alkylates compounds can phosphorylate oxygen atoms in single stranded RNA regions which do not form tertiary interactions. This results in the formation of a phosphodiester, which is hydrolysed, causing the RNA to cleave. RNA fragments generated by cleavage can subsequently be analysed by gel electrophoresis as previously described [313]. Traditional chemical and enzymatic approaches have a number of limitations. Firstly, the reagents can react non-specifically, causing some base pairs, such as G-U pairs, to be identified as reactive [314]. In addition, a given reagent can typically only react with a subset of the four RNA nucleotides, for example RNase T1 can only react with G nucleotides [312].

This means that multiple reagents need to be used to examine all of the nucleotide positions in a given RNA sequence, which may also result in conflicting data for individual nucleotides [314]. The size of the transcript which can be analysed is also limited to the resolving size of the denaturing gel and requires the use of radio labelled primers.

To avoid these short comings, RNA structures in ZIKV were mapped using selective 2'-hydroxyl acylation analysed by primer extension (SHAPE) chemistry. SHAPE involves the use of reagents, such as N-methyl-nitroisatoic anhydride (NMIA), which covalently bind to the 2'-hydroxyl groups of the ribose backbone in nucleotides which are structurally flexible (Fig 4.1). The workflow for SHAPE reactions/ data processing is summarised in figure 4.2 and 4.3. Single stranded (unconstrained) nucleotides are more likely to display conformations that enhance the nucleophilicity of the 2' hydroxyl group, therefore NMIA reacts selectively with non-base paired RNA nucleotides, forming adducts which can be detected using reverse transcription. As the reverse transcriptase enzyme terminates at the position of the adducts, this generates a pool of multiple cDNAs, the length of which is determined by the nucleotide position at which termination occurred [315]. The cDNA is then quantified to give a measure of the reactivity of each nucleotide within the RNA sequence. The position of each nucleotide is determined by comparing the NMIA signal with a sequencing ladder derived from the same RNA sequence and each nucleotide position is assigned a normalised reactivity score on a scale of 0 to 5, where a score of 1 is the average reactivity for a highly reactive nucleotide [291]. Nucleotides with a normalised reactivity score of greater than 0.7, up to a score of 5, are classed as highly reactive, indicating that these nucleotides are structural flexible and are therefore non-base paired (nucleotides with a score of greater than 5 are classified as having a reactivity of 5). A reactivity score between than 0.3-0.7 indicates that the nucleotide is partially reactive, indicating that the nucleotide is sometimes base paired and sometimes non-base paired. This may suggest that the RNA region is structurally flexible and adopts multiple conformations. By contrast, nucleotides with a score of less than 0.3 have low reactivity, indicating that these nucleotides are constrained and therefore are base paired.

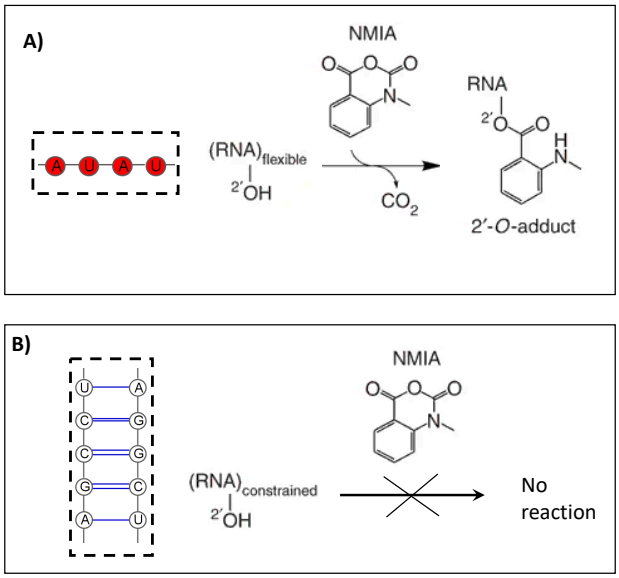


Figure 4.1 RNA modification by N-methyl-nitroisatoic anhydride (NMIA)
A) *NMIA reacts preferentially with the 2' -hydroxyl group of non-basepaired RNA nucleotides, which are flexible, forming 2' -O-adducts.* ***B)*** *NMIA does not react with base-paired nucleotides, where the nucleotides are constrained. Adapted from Wilkinson et al* [316].

The SHAPE technique has several advantages over the traditional chemical and enzymatic probing techniques. As NMIA is able to react with all of RNA bases, a single reagent can be used to interrogate every base, therefore the entire RNA sequence can be analysed in a single reaction. Due to high sensitivity, NMIA reactivity is able to detect tertiary RNA structures, such as pseudoknot interactions. The technique was further improved by the development of high-throughput SHAPE, whereby the radioactive primers are replaced with fluorescently labelled primers and then analysed by capillary electrophoresis [317]. This approach allows longer sequences to be investigated, as sample reads are not limited to the size of a gel, and 96 samples to be consecutively investigated, increasing the efficiency of data analysis.

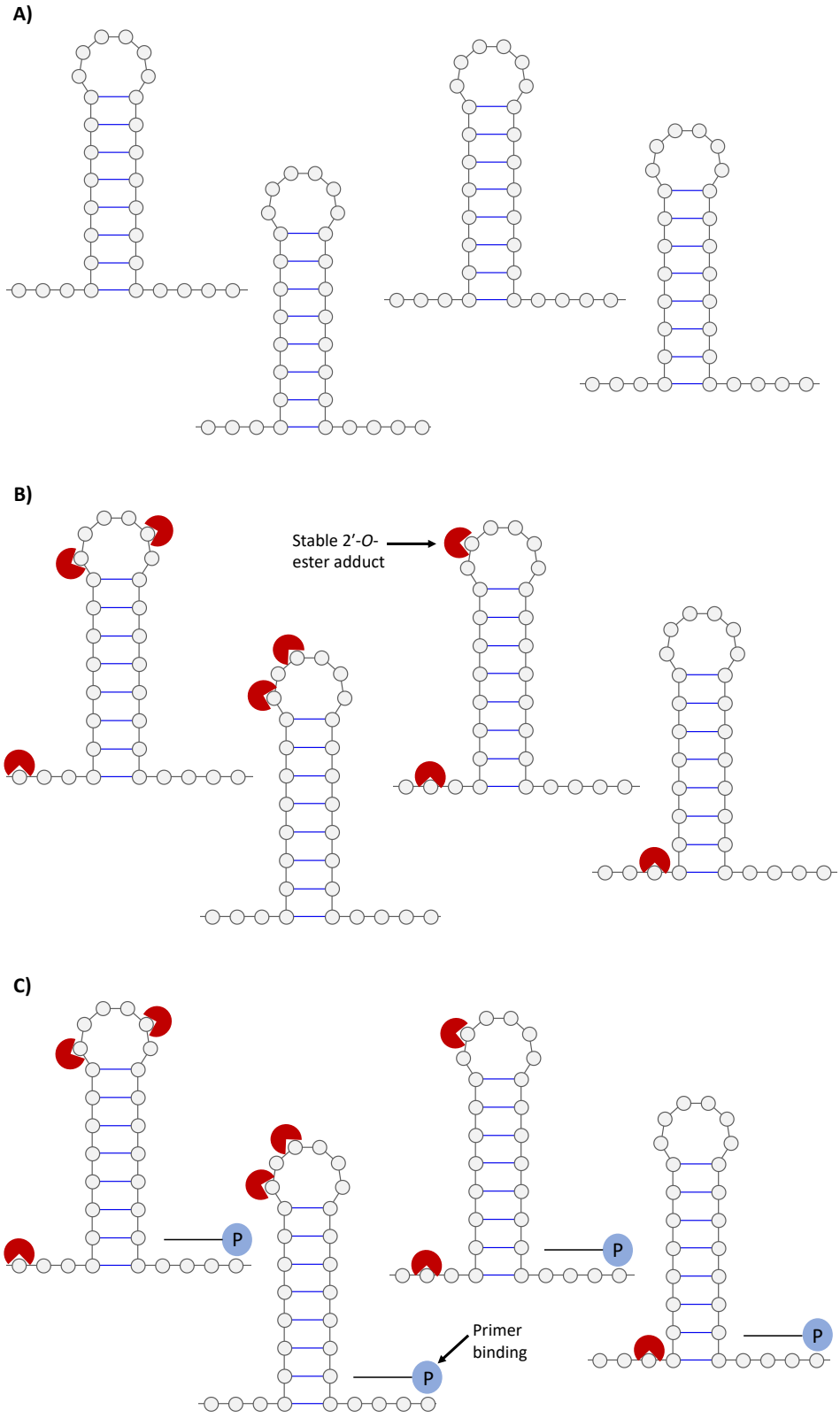
RNA structural models can also be produced using *in silico* thermodynamic prediction tools, such as mfold, which can predict RNA secondary structure based on minimum free energy of folding. Whilst such models can provide some insight, they may not include additional folding conformations associated with higher order interactions or biological conditions. In order to produce more accurate models of RNA secondary structure, SHAPE reactivity data can be combined with *in silico* algorithms, such as restrained sample (Rsample). Rsample is incorporated into the "RNAstructure" software package, in which thermodynamic models can be constrained using reactivity values obtained from SHAPE reactions. The algorithm analyses the agreement between experimentally acquired data and the thermodynamic prediction, producing a more accurate representation of the secondary structure. In addition, Rsample considers that multiple copies of the same RNA can fold into different conformation, which may be important in a biological context [318]. This approach has been shown to be highly accurate, as the predicted secondary structure for 16S rRNA was shown to have 97% similarity with the base pairs visualised in the crystal structure of the 30S ribosomal subunit [319]. The combination of SHAPE data and prediction algorithms also correctly predicted the secondary

structure of RNAs ranging from 34-530 nts, some of which included pseudoknots, with >90% accuracy [320] and was shown to have >90% accuracy in predicting base pairing within difficult-to-model RNAs, where the structure had previously been solved [321].

The remaining 10% discrepancy between previously established structures and SHAPE-derived structures may be attributed to several properties of the SHAPE method. SHAPE reactivity reflects the structural conformation in solution at the time of probing, meaning that if an RNA is partially misfolded or adopts multiple conformations, then this will be reflected in the reactivity data [321]. In addition, nucleotides within base paired regions may be NMIA reactive in some RNA molecules, for example nucleotides within the closing pairs of helices and adjacent nucleotides are more prone to acetylation. Nucleotides within base paired regions can also be constrained in a conformation in which the 2'-hydroxyl group is accessible for NMIA binding [322]. Despite this, the combination of SHAPE reactivities and structure prediction algorithms is a highly powerful and relatively high-throughput technique for the determination of RNA structure, especially when data is combined from several biological repeats.

This chapter describes the optimisation of a ZIKV infectious clone system, a DNA system which allows manipulation of the ZIKV genetic sequence and recovery of ZIKV mutant viruses. Since this system was novel to the research group at the beginning of this project, a significant amount of optimisation was required. Using RNA derived from this system, SHAPE mapping was performed to investigate RNA structures within the 5' region of the ZIKV genome, encompassing the 5' UTR and the adjacent capsid coding region, which were unknown when the project began. Prior to SHAPE analysis, multiple primers were investigated for efficacy. SHAPE analysis was performed to determine the RNA structure and genome conformation, either linear or circular, of full-length ZIKV genome RNA using *in vitro* SHAPE, at both 28 °C and 37 °C. In addition, to investigate genome cyclisation, the RNA structure of ZIKV 1-754, which lacks cyclisation sequences, at both 28 °C and 37 °C was determined and compared with SHAPE data for full-length RNA.

During the course of this project, two publications reported RNA structural data for ZIKV. Liu et al. reported *in vitro* SHAPE reactivity for the 5' UTR and adjacent capsid coding region of ZIKV [225] and Li et al. reported *in cell* SHAPE reactivity for the entire ZIKV genome [212]. These publications provide a useful point of comparison with the data presented in this thesis.


A)
B)
Stable 2′-O-
ester adduct
C)
P
P
P
P
Primer
binding

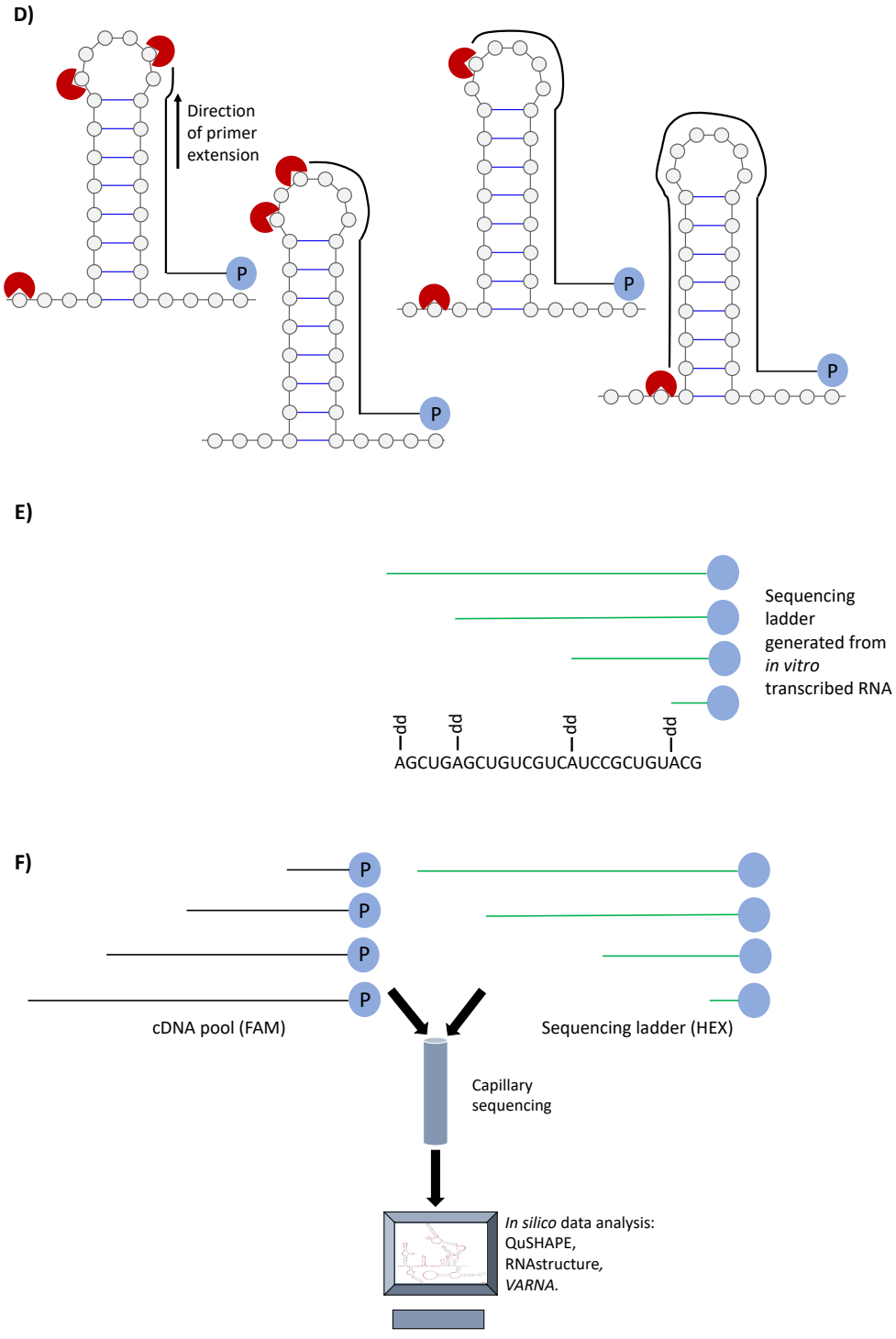


Figure 4.2 SHAPE reaction process

A) *RNA is heated to 95 °C and then incubated at either 28 °C or 37 °C to allow formation of RNA structures.* ***B)*** *RNA is treated with NMIA, which binds to non-base-paired nucleotides, forming abducts.* ***C)*** *Fluorescent primer binds to RNA.* ***D)*** *Primer extension and reverse transcription. Extension is blocked by abducts.* ***E)*** *Sequencing ladder is generated from primer extension and reverse transcription of in vitro transcribed RNA.* ***F)*** *Fluorescent cDNA pool (corresponding to the position of modified nucleotides) and sequencing ladder are combined and sequenced by capillary. Sequence files are analysed in silico.*

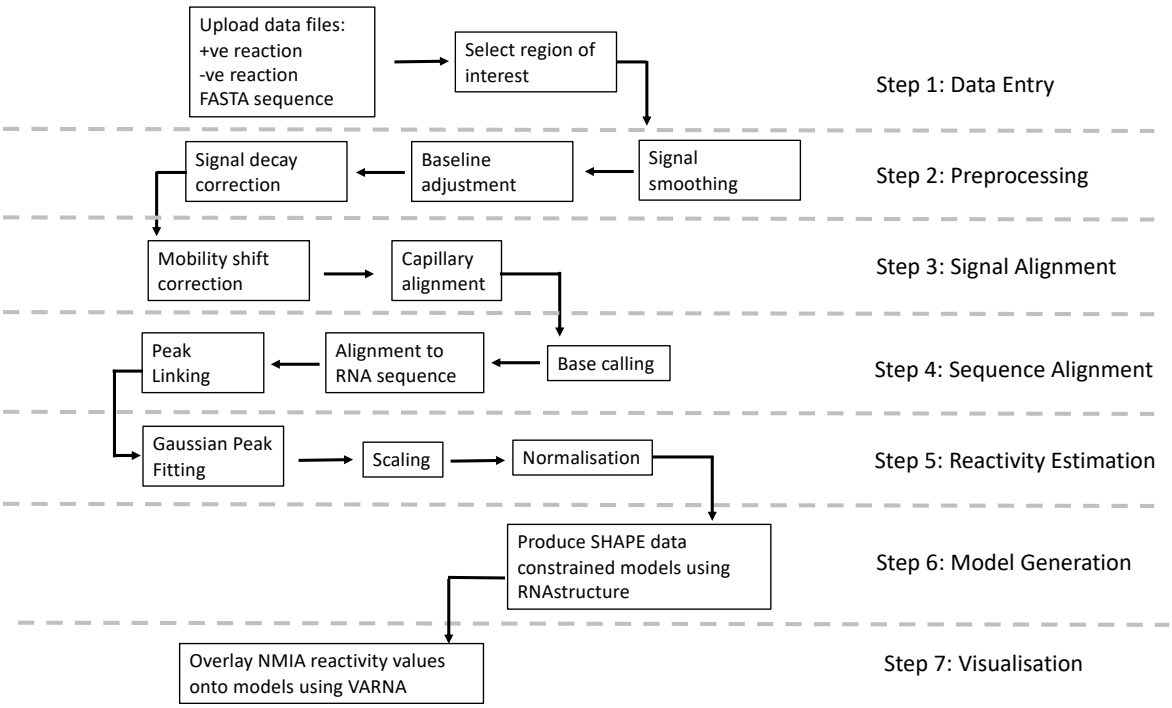


Figure 4.3 SHAPE data analysis workflow

Step 1: *Data from NMIA treated (+) and control DMSO (-) reactions and a FASTA file containing the RNA sequence file are uploaded to QuSHAPE and the region of interest is selected.* ***Step 2:*** *Preprocessing operations remove high-frequency noise and baseline offset. Signal decay correction converts the fluorescence signal intensities to probabilities of primer termination.* ***Step 3:*** *The use of different fluorescent labels (FAM and HEX) and the use of different capillaries result in different retention times. The mobility shift correction operations align pairs of signals within each capillary and capillary alignment aligns the signals across two capillaries.* ***Step 4:*** *The base calling operation classifies all of the peaks in the sequencing signal from the DMSO (-) capillary as either specific peaks produced by ddATP and non-specific peaks corresponding to the other three bases. The signal from the DMSO signal (FAM) is then aligned with the RNA sequence signal (HEX). The peak linking operation then links the (-) sequencing data to the corresponding peaks in the (+) and (-) signals.* ***Step 5:*** *Reactivity estimation is determined uses Gaussian integration, where the area for each peak is correlated with the primer termination probability for the corresponding nucleotide in the RNA sequence. The scaling operation determines the magnitude of the scaling parameter, which determines the spread of the distribution and establishes units of measurement. Normalisation computes the probability of adduct formation at each nucleotide (P_{add}). P_{add} is then normalized, such that the average intensity for highly reactive nucleotides is 1.0 and where zero indicates no reactivity. Data was then capped, such that reactivates of >5 were given a score of 5 and nucleotides which had no data were given a score -0.1.* ***Step 6:*** *RNA structural models are generated using the "RNAstructure" prediction algorithm, where the prediction is constrained by the SHAPE reactivity values. This process creates dot-bracket files, which represent the predicted RNA secondary structure.* ***Step 7:*** *Dot brackets files are used to create RNA structure maps using VARNA and the SHAPE reactivity values are overlaid on the maps for visualisation of SHAPE reactivity.*

4.2 Optimisation of ZIKV DNA system

In order to study the replication of ZIKV in cell culture, an infectious cDNA ZIKV clone designated pCCI-SP6-ZIKV was utilised. pCCI-SP6-ZIKV is based on a consensus sequence of 5 ZIKV isolates from the Brazilian lineage and was kindly provided by Professor Andres Mertis (University of Tartu, Estonia). The cDNA infectious clone system allows the use of reverse genetic approaches to investigate the phenotypic effects of point mutations, which affect both the structure and/or sequence of the RNA stem loops and tertiary interactions.

pCCI-SP6-ZIKV is a low copy number plasmid and exhibits toxicity in *E. coli*, resulting in a high potential for DNA recombination. Purification of pCCI-SP6-ZIKV therefore requires modifications to standard *E. coli*-based DNA purification methods. Initial attempts to transform pCCI-SP6-ZIKV, using high efficiency 5-alpha F' Iq competent *E. coli* (NEB) on LB agar, failed to recover colonies. Growth of pCCI-SP6-ZIKV transformed colonies was supported on tryptic soy agar and transformation efficiency was improved through the use of high efficiency turbo competent *E. coli* cells (NEB). DNA was successfully purified from turbo competent cells using a GeneJET Plasmid Maxiprep kit (Thermo fisher) however, agarose gel electrophoresis confirmed that that the DNA quality was poor.

In consultation with Prof. Merits, further modifications to the DNA purification protocol were made. In order to activate the second origin of replication within the plasmid, DNA was transformed into EPI 300 cells and incubated with 0.1% arabinose. Cells were then lysed and lysate was filtered through a gauze, before DNA was precipitated with isopropanol. Pellets were then processed using a GeneJET Plasmid Miniprep kit (Thermo fisher) to purify the DNA. This resulted in greatly improved DNA integrity and purity.

For generation of ZIKV IC RNA, initially, DNA was linearised using NotI-HF and the linear DNA was used as a template for *in vitro* transcription. However, this approach proved unreliable due to the inherent instability of the plasmid. To solve this problem, the ZIKV IC plasmid was used as the template for a PCR reaction. PCR reactions were carried out using a high-fidelity polymerase, amplifying the entire ZIKV IC genome and resulting in high quality templates. Full-length PCR DNA was first purified using a Wizard SV Gel and PCR Clean-Up System (Promega) however, yields of purified DNA were very low. To solve this issue, PCR DNA was purified the using ethanol precipitation, which resulted in improved DNA yield.

4.3 Optimisation of ZIKV SHAPE primers

Primers were designed to bind two regions within the ZIKV RNA sequence, allowing SHAPE reactions to resolve the extreme 5' end and an upstream region of the ZIKV genome. SHAPE primers 575, 583 and 590 were designed to bind within the middle of the capsid coding region, resulting in a 247-262 bp extension product. Primer 802 binds at the 3' end of the Capsid coding sequence, producing an extension product of 474 bp. Primers 824 and 826 binds to the 5' end of the prM coding sequence, producing an extension product of 496-498 bp. Test primers were end labelled using an ATP [γ-^{32}P] marker for visualisation.

In vitro extension reactions were performed (Fig. 4.4), using ZIKV 1-754 as a template, in order to determine which of the primers investigated produced the best extension products for use in SHAPE reactions. Products were analysed using urea-acrylamide gel electrophoresis and visualised using photo-stimulated luminescence. Primers 590, 583 and 575 all produced products of a similar size. Compared with primers 575 and 583, primer 590 produced higher levels of ZIKV RNA and was therefore selected for use in SHAPE experiments. Primers 802 and 824 produced truncated transcripts compared with primer 826, therefore primer 826 was taken forward for use in SHAPE experiments.

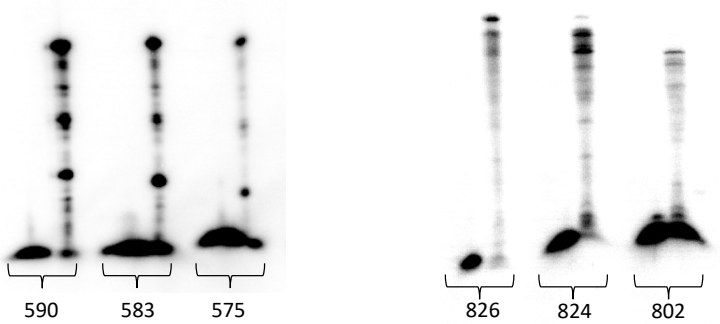
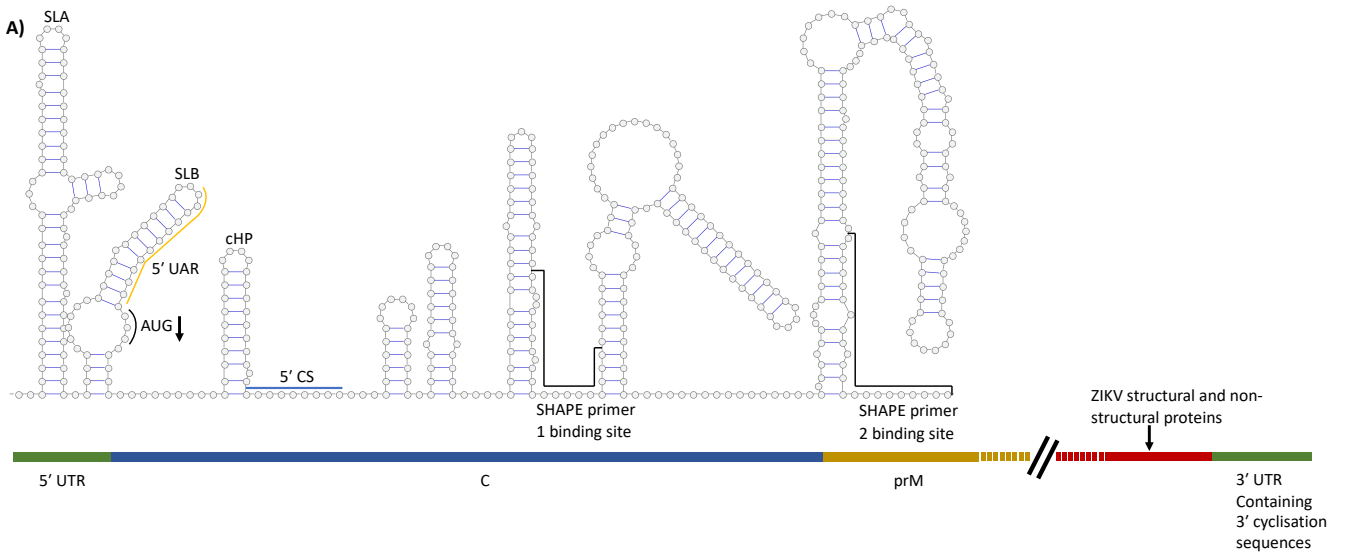
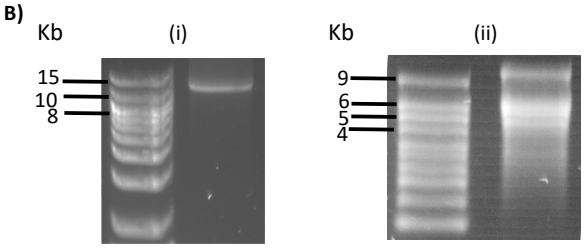


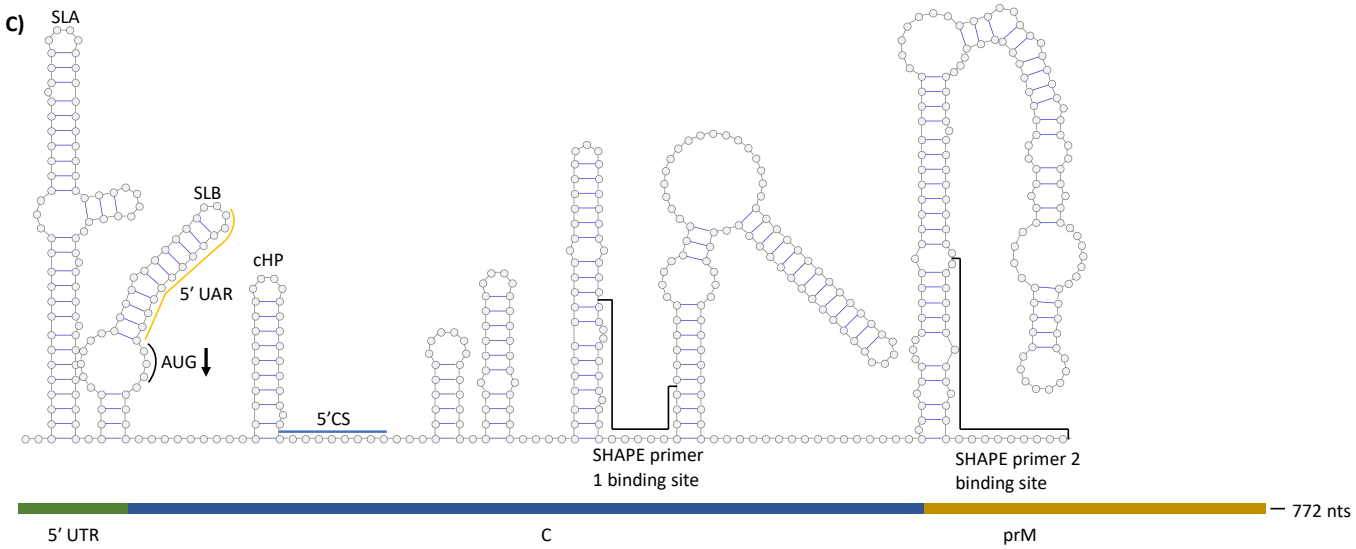
Figure 4.4 ZIKV RNA in vitro extension reactions
300 ng of ZIKV 1-772 RNA was incubated with γ-32P labelled primers and visualised using denaturing Urea gels. Labelled primers were run alongside extension reactions. Primer 590 and 826 were selected for use in SHAPE experiments as these primers resulted in extension reactions with the longest extension products and highest yield of RNA. Binding sites for primer 590 (primer binding site 1) and 826 (primer binding site 2) are shown in figure 4.5.

4.4 *In vitro* SHAPE mapping of ZIKV

RNA structures within the 5' region of the ZIKV genome, were characterised using *in vitro* SHAPE mapping, based on the reactivity of individual RNA nucleotides to the chemical reagent NMIA and structure prediction algorithms. SHAPE reactions were performed on both full-length RNA template, representing the full genomic RNA, and shorter RNA molecules encompassing only the 5' UTR and adjacent coding region, and therefore lack the 3' hybridisation sequences required for genome cyclisation. The intention behind this was to investigate if the circular genome conformation is favoured by full length genomic RNA molecules in an RNA folding buffer, in the absence of potential *trans* activating factors (such as the NS5). Additionally, SHAPE reactions were performed at both 28 °C and 37 °C, reflecting mosquito and mammalian body temperature respectively. Any differences which may occur in RNA structures at the two temperatures may reflect host-specific adaptations.


A)
SLA
SLB
cHP
5′ UAR
AUG
5′ CS
SHAPE primer
1 binding site
SHAPE primer
2 binding site
ZIKV structural and non-
structural proteins
5′ UTR
C
prM
3′ UTR
Containing
3′ cyclisation
sequences


B)
Kb
(i)
15
10
8
Kb
(ii)
9
6
5
4


C)
SLA
SLB
cHP
5′ UAR
AUG
5′CS
SHAPE primer
1 binding site
SHAPE primer 2
binding site
772 nts
5′ UTR
C
prM

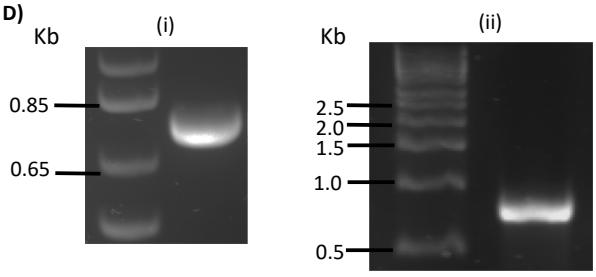
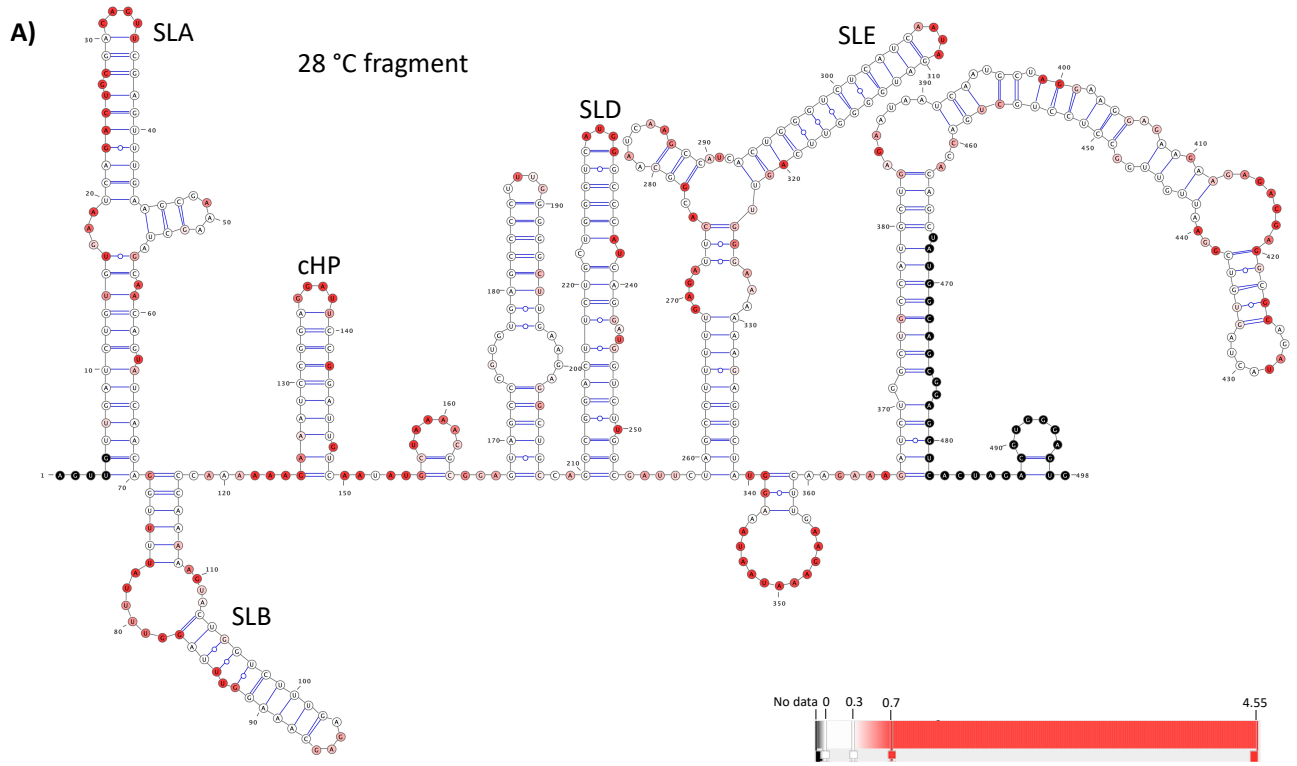

D)
Kb
(i)
0.85
0.65
Kb
(ii)
2.5
2.0
1.5
1.0
0.5

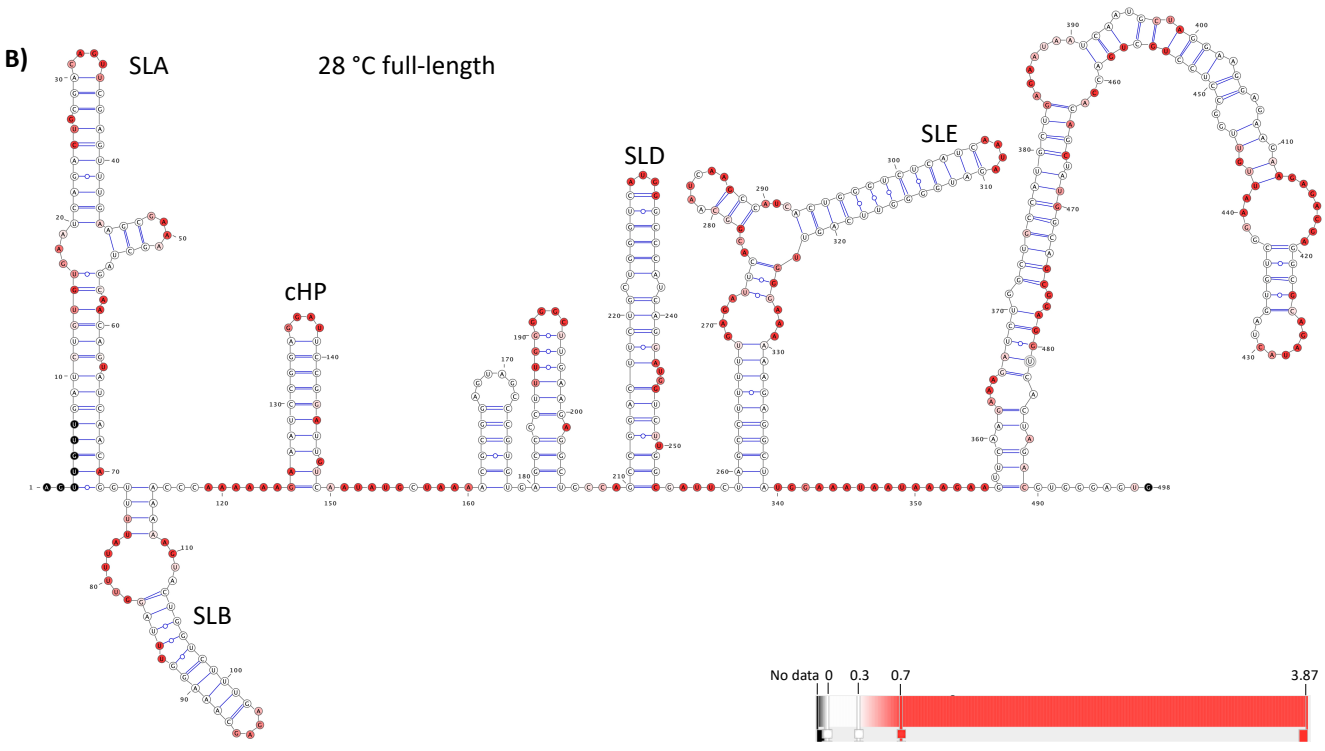
Figure 4.5 Schematic of 5' 772 nt and full-length ZIKV transcripts used for in vitro SHAPE experiments

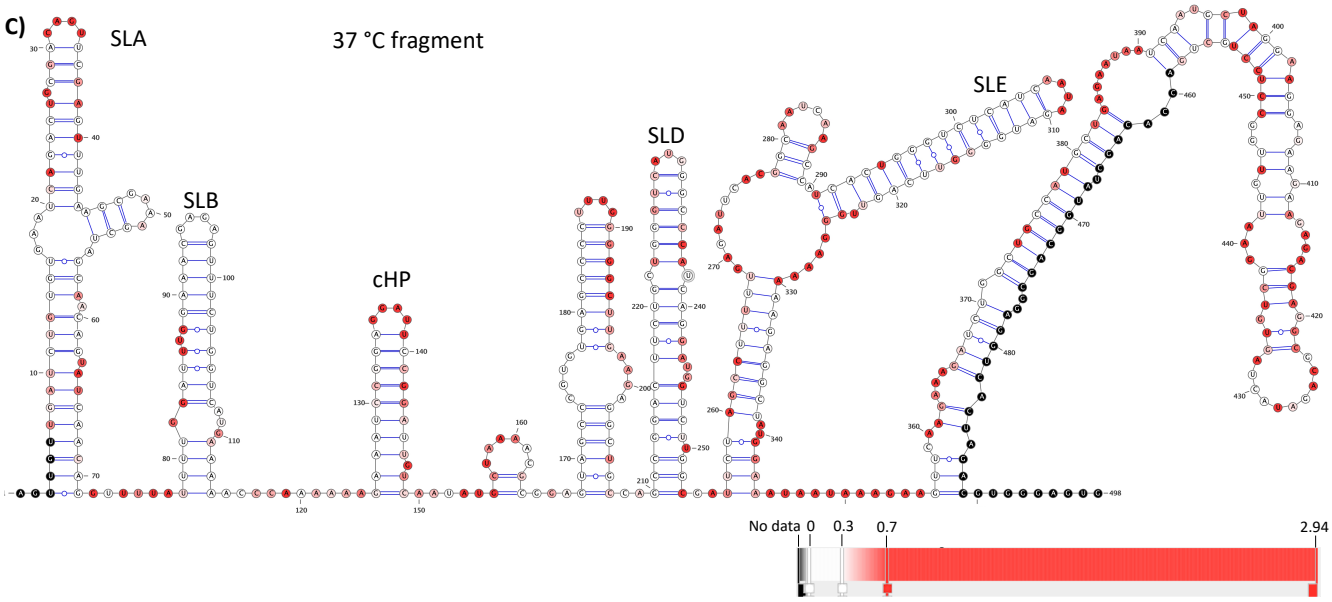
A) *The full length ZIKV genome transcript, containing the complementary 3' cyclisation sequences.* ***B)*** *(i) 1 % agarose gel of full-length ZIKV DNA PCR product (1 Kb plus ladder, Thermo Fisher Scientific). (ii) 1 % MOPS gel of full-length ZIKV RNA in vitro transcription product (Millenium^{TM} RNA Marker, Thermo Fisher Scientific).* ***C)*** *The 5' 772 nt fragment transcript contains the first 772 nts of the ZIKV genome and therefore is unable to hybridise with complementary 3' cyclisation sequences to form the circular genome conformation.* ***D)*** *(i) 2 % agarose gel of 5' 772 nt fragment DNA PCR product (1 Kb plus ladder, Thermo Fisher Scientific). (ii) 2 % MOPS gel of 5' 772 nt fragment RNA in vitro transcription product (Millenium^{TM} RNA Marker, Thermo Fisher Scientific).*

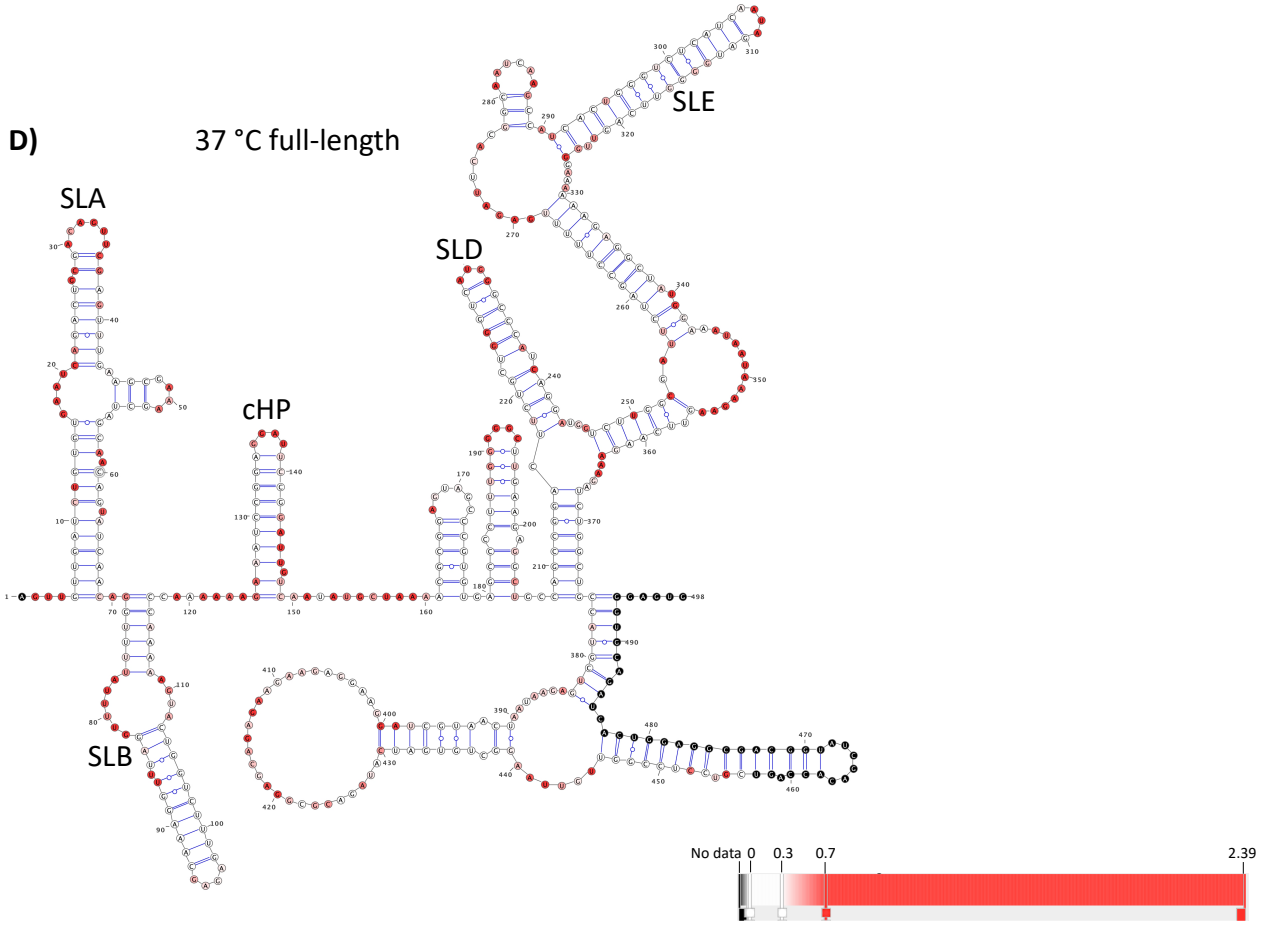
SHAPE reactions were performed using an RNA folding buffer (330 mM HEPES (pH 8.0), 20 mM $MgCl_2$ and 330 mM NaCl) published by Weeks et al., which the Tuplin group have previous used to perform SHAPE reactions using CHIKV RNA [2,316]. Weeks et al. report that NMIA reactivity is dependent on pH and recommend using the folding buffer at pH 8.0 however, acknowledge that in practice a pH range of between 7.5 and pH 8.0 works well [316]. Optimisation experiments performed by the Tuplin group prior to the start of this project tested a pH range of pH 6.8 to pH 8.0 (a broad range to account for different cell types and intracellular compartments), with no difference in experimental outcome observed (unpublished data). Therefore to maintain consistency with the published literature [2,316] , the experiments in this thesis were preformed using the RNA folding buffer at pH 8.0.

SHAPE fragment analysis was carried out by capillary electrophoresis and data analysis was performed using QuSHAPE. Reactivity values were then used to constrain thermodynamic prediction algorithms using "RNAstructure", generating structural models of the RNA regions. Reactivity values represent a composite of data combined from the two primers used. Data obtained using primer 590 was used for the upstream nucleotides and primer 826 data was used for the downstream region, with a short region of data overlap around nucleotide 160. SHAPE reactivity values were overlaid onto RNA structure models for visualisation, using the VARNA visualisation applet for secondary structure.


A)
SLA
28 °C fragment
SLE
SLD
cHP
SLB
No data 0
0.3
0.7
4.55


B)
SLA
28 °C full-length
SLE
SLD
cHP
SLB
No data 0
0.3
0.7
3.87


c)
SLA
37 °C fragment
SLB
cHP
SLD
SLE
No data 0
0.3
0.7
2.94


D)
37 °C full-length
SLA
SLE
SLD
cHP
SLB
No data 0
0.3
0.7
2.39

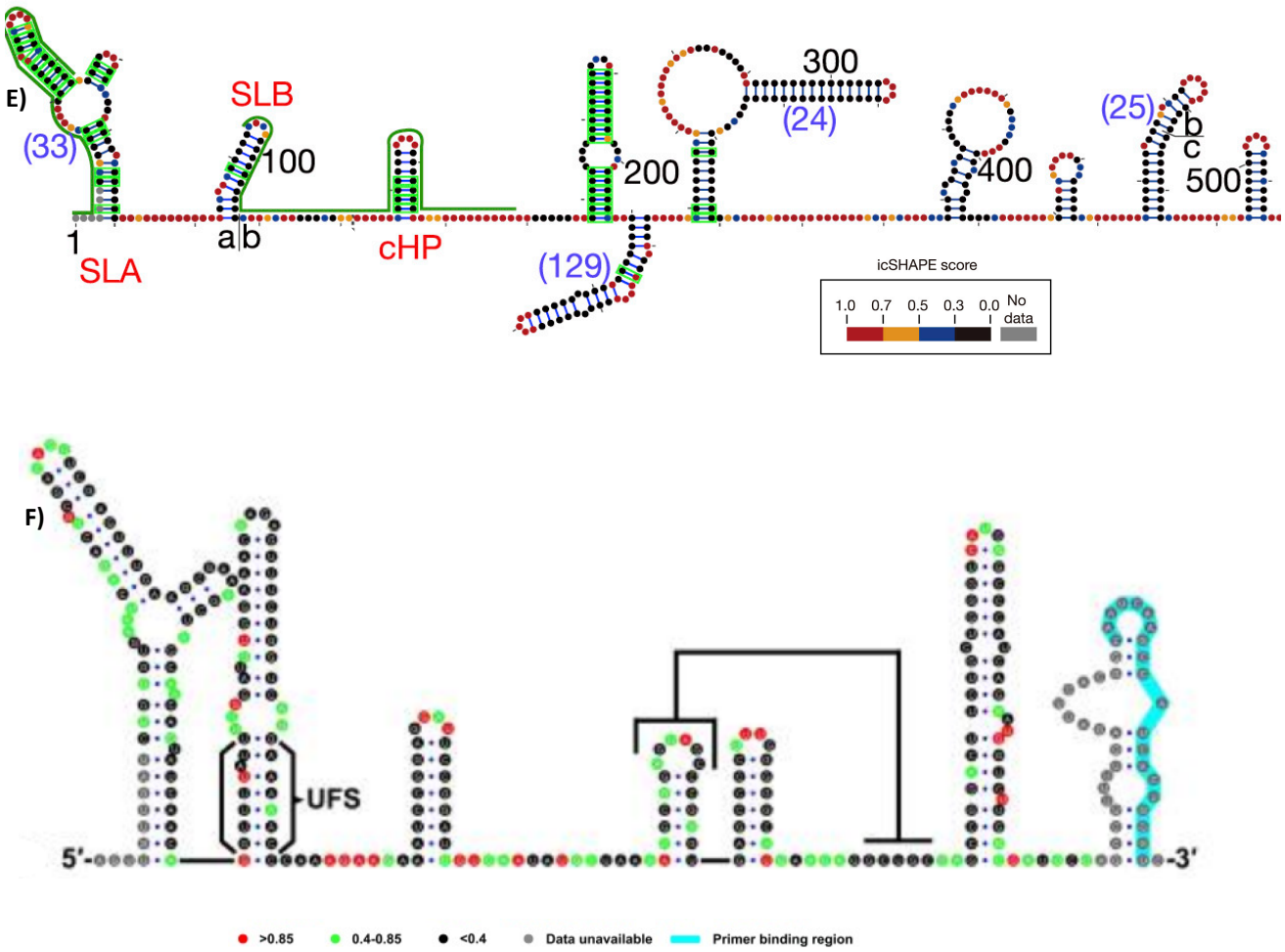


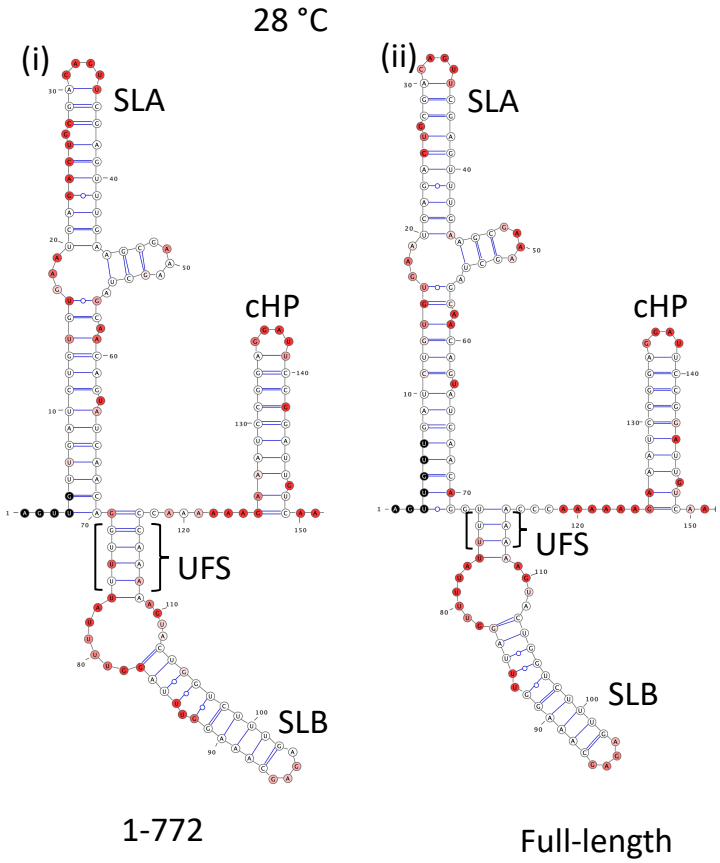
Figure 4.6 Full models of the ZIKV 5′ UTR and adjacent coding region

A) *28* °C *ZIKV RNA 1-772.* ***B)*** *28* °C *full-length RNA.* ***C)*** *37* °C *ZIKV RNA 1-772.* ***D)*** *37* °C *full-length RNA. Models are generally consistent between conditions, with full-length, 37* °C *models showing variation base pair formation at the base of SLD. Despite this variation, the top region of SLD remains consistent with the other models.* ***E)*** *Published in cellulo SHAPE map of ZIKV (Adapted from Li et al.* [212]*).* ***F)*** *Published in vitro SHAPE map of ZIKV (Adapted from Liu et al.* [225]*).*

At the extreme 5’ end of the genome, the first 1-4 nucleotides of the genome are non-base-paired (or predicted to be non-based-paired in the absence of data) in all conditions tested (Fig 4.7). The unpaired nucleotides are immediately followed by stem loop A (SLA), which has high structural conservation with other mosquito-borne flaviviruses such as Dengue virus and West Nile virus [152]. Stem loop A (SLA) forms at the 5’ end of the ZIKV RNA and is described as having a Y-shape, consisting of a primary stem loop with a short side stem. Overlaid SHAPE reactivity values for SLA showed a high level of agreement with the “RNAstructure” generated structural model, suggesting that the SHAPE reactivity for this region is highly accurate, with only minor differences between the conditions tested. In the 37 °C full length reactions, SLA is comprised of 66 nucleotides, from nucleotide 5-69. High NMIA reactivity was observed at the apical top loop of SLA, indicating that the top loop may have a more open confirmation compared with the structural predictions. This may be important for interactions with mammalian specific proteins or for the formation of higher order interactions with other regions of the genome. Li et al. suggest possible long range interactions between SLA and an upstream region of the ZIKV genome at 37 °C [212]. In the 37 °C fragment and 28 °C full length “RNAstructure” generated models, SLA consists of nucleotides 3-71 and is 70 nucleotides long. Due to a lack of data for the extreme terminal nucleotides (i.e. nucleotides 1-3 for the ZIKV genome), it is possible that the modelling for this region is not accurate. For the 37 °C fragment reactions, nucleotides at the base region of SLA were highly reactive, indicating that there may be some flexibility in this region. In the 28 °C fragment “RNAstructure” generated model, SLA consists of nucleotides 4-70.

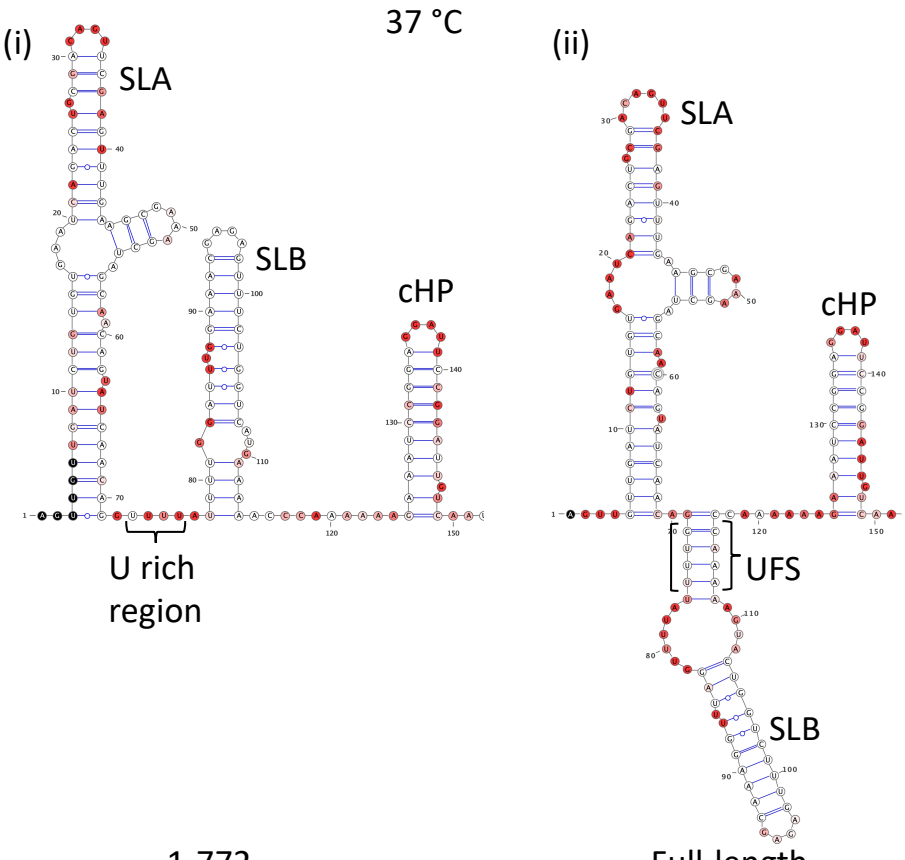
In all “RNAstructure” generated models, the SLA side stem comprised of 10 nucleotides, from nucleotide 45-54. In both the 28 °C and 37 °C full length reactions, the non-base-paired nucleotides at the side stem terminus were highly reactive. The same nucleotides were less reactive in the 37 °C fragment reactions and mostly unreactive in the 28 °C fragment reactions. This could indicate higher order interactions taking place in the 28 °C fragment reactions.

A)



1-772 Full-length

B)



1-772 Full-length

Figure 4.7 RNA structural models for the 5′ region of ZIKV, extreme 5′ end region
A) *28* °C *(i) ZIKV 1-772 and (ii) full-length RNA.* ***B)*** *37* °C *(i) ZIKV 1-772 and (ii) full-length RNA. The 5′ -UAR-flanking stem (UFS) is shown at the base of SLB, with the exception of the 37* °C *fragment model where the U rich region is shown. The UFS region in the full-length 28* °C *model is shortened compared with the other models.*

The next RNA structure is stem loop B (SLB), within which the 5′ upstream of AUG region (5′ UAR) is located. The 5′ UAR has been demonstrated to hybridise with the complementary 3′ UAR to form long range interactions during genome cyclisation. In addition, the base of SLB contains a Uracil rich region which has been designated the 5′ -UAR-flanking stem (UFS). Disruption of the UFS by mutagenesis results in reduced viral titre and the UFS has been proposed to play a role in the switch between linear and circular genome conformations [225]. In the 37 °C full length, 28 °C full length and 28 °C fragment "RNAstructure" generated models, SLB consists of a short base, consistent with the UFS region described by Liu et al [225] followed by an open region of non-base paired nucleotides and then a stem with a highly reactive apical loop. SHAPE data overlays are highly consistent with "RNAstructure" generated models for each of these conditions, meaning that there is high confidence in the structural models. In cell SHAPE of whole cells lysates performed by Li et al shows the U rich region as non-base-paired and therefore shows SLB as reduced in length compared to these models, lacking the UFS region and having weaker reactivity at the apical loop [212]. This may reflect differences between *in vitro* and in cell methods of SHAPE however, since Liu observed differences in phenotype between WT and UFS mutants, this region is likely to play a role in replication. Since Li et al. performed SHAPE using whole cell lysates, it is also possible that the RNA adopts a separate, UFS containing, conformation in the context of viral replication bodies, compared with the rest of the cell. SHAPE analysis of RNA contained within replication bodies would be needed to confirm this. The "RNAstructure" generated model for 37 °C fragment displayed some significant differences from the other *in vitro* models. The uracil rich region showed high levels of reactivity, resulting in a lack of UFS formation. In addition, the open region is reduced in size and the apical loop showed only very weak or no reactivity and is therefore more similar to the results seen using in cell SHAPE.

The conserved RNA structure cHP was present in all "RNAstructure" generated models as a 25-nucleotide structure, comprised of nucleotides 125-149 (Fig 4.8). High NMIA reactivity at the base of the structure in all SHAPE data overlays suggest that these nucleotides may not form part of the structure, resulting in a slightly shorter stem. Li et al. and Liu et al. report that cHP consists of 20 nucleotides, supporting the concept that these highly reactive nucleotides do not form part of cHP [212,225]. Adjacent to cHP, is a 9-nucleotide sequence termed the 5′ cyclisation sequence (5′ CS), which hybridises with a complementary 3′ CS during genome cyclisation. The high NMIA reactivity of nucleotides within this sequence indicate that the 5′ CS is non-base-paired, therefore suggesting that the ZIKV genome RNA favours the linear conformation in the absence of cellular factors. In both the 28 °C and 37 °C 1-772 "RNAstructure" generated models, a small loop is present formed by nucleotides 155-164. Since nucleotides within this region have high reactivity, this region likely to be single stranded, as shown in both of the full-length SHAPE maps.

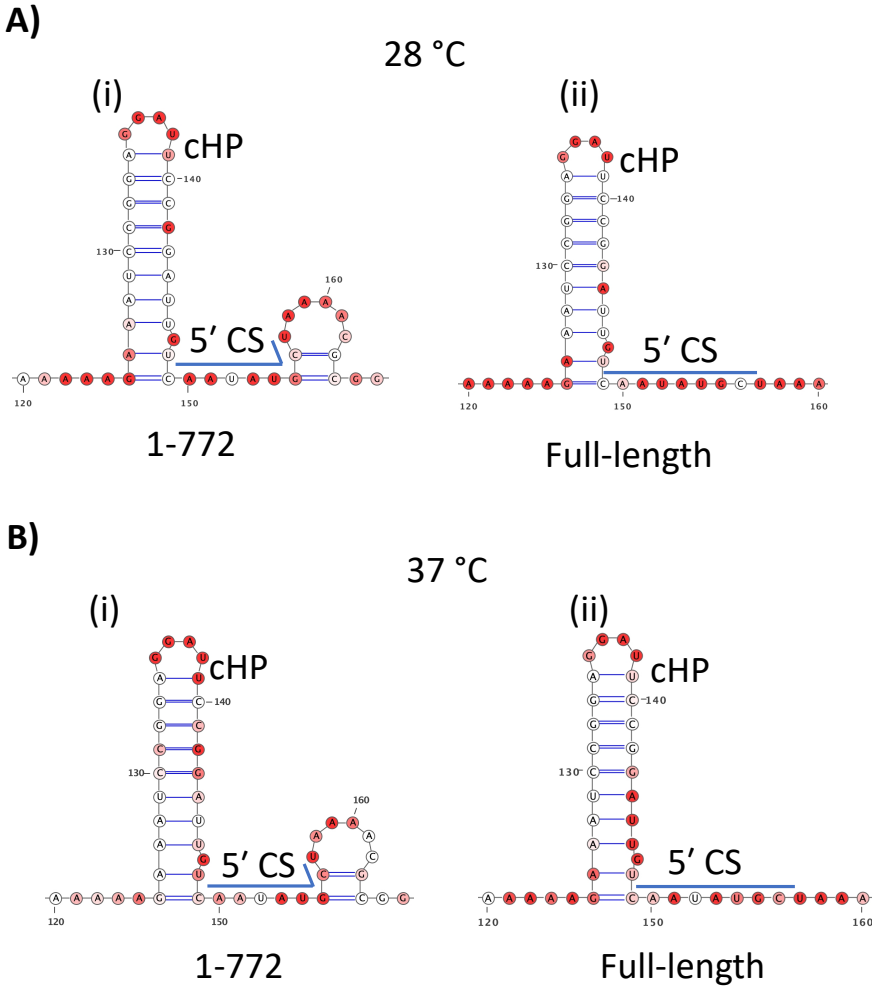


Figure 4.8 RNA structural models for the 5′ region of ZIKV, cHP region
A) *28* °C *(i) ZIKV 1-772 and (ii) full-length RNA.* ***B)*** *37* °C *(i) ZIKV 1-772 and (ii) full-length RNA. Nucleotides comprising the 5′ CS are denoted by the blue line. A small loop is present in both of the 1-772 models however, the high reactivity of these nucleotides suggest that this loop may not form.*

Previously published SHAPE data for the region downstream of the 5′ CS (nucleotides 160-230) present conflicting structural models. Liu et al. present the region as having two short stems, with a predicted pseudoknot interaction between the apical loop of the first stem and a downstream region of complementary nucleotides. In contrast, Li et al. [212] show this region as one single structure with no pseudoknot interactions. This difference means that the downstream nucleotides, which are predicted to form part of the pseudoknot interaction, are instead incorporated into the base of the elongated stem loop.

In both the 28 °C and 37 °C full length "RNAstructure" generated models, there are two stem loops, the first of which has identical base pairing to the first stem loop shown by Liu et al., however, the second of the stem loops is elongated in comparison. In the current study, for both 28 °C and 37 °C fragment "RNAstructure" generated models, a single structure is predicted, which is consistent with the data shown by Li et al. (Fig 4.9). Whilst the RNA

structure generated model for 1-772 at 28 °C is consistent with overlaid SHAPE data, in the 37 °C RNA structure generated model for 1-772, SHAPE data overlay shows high levels of NMIA reactivity within the structure, indicating that the long RNA structure predicted to be formed by nucleotides 168-207 may not be accurate.

For the next RNA structure, 37 °C fragment, 37 °C and 28 °C full length "RNAstructure" generated models show a long RNA structure formed by nucleotides 210-253 with a highly NMIA reactive apical loop, designated here as stem loop D (SLD) (Fig 4.9). The structure and length of SLD is highly consistent with the results shown by both Li et al. and Liu et al. For the 37 °C full length "RNAstructure" generated model, base pairing in this region shows variability, as nucleotides within the base of the SLD stem are paired with nucleotides 368-375. This causes a significant change in local folding architecture; whereby downstream RNA structures form as part of a side stem of SLD. Despite these differences, the top region (nucleotides 217-243) of SLD remains highly consistent with the other models and with the published literature, indicating that the top region may be the most stable and conserved region of SLD.

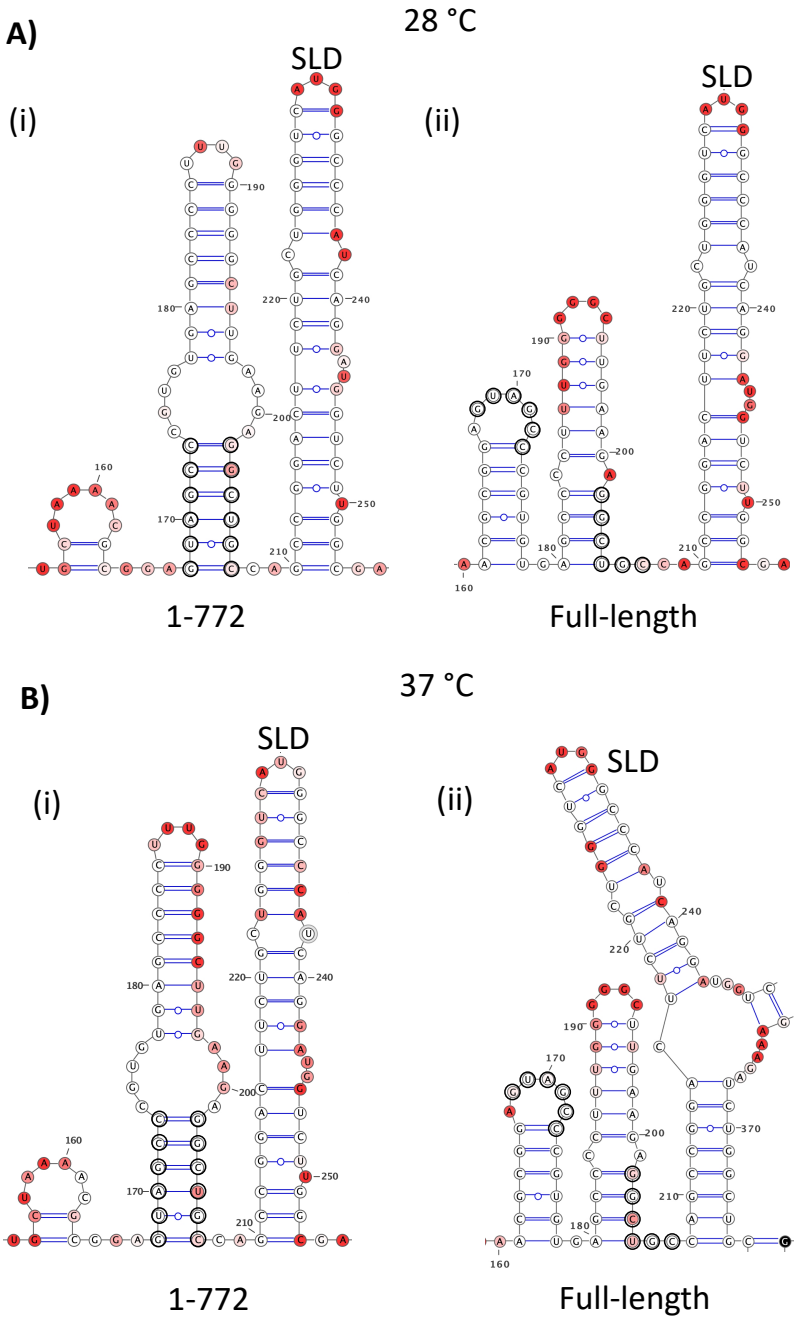


Figure 4.9 RNA structural models for the 5' region of ZIKV, SLD region
A) *28* °C *(i) ZIKV 1-772 and (ii) full-length RNA.* ***B)*** *37* °C *(i) ZIKV 1-772 and (ii) full-length RNA. Nucleotides predicted to form pseudoknot interactions by Lui et al. are indicated by circles. In the ZIKV 1-772 models, a single RNA structure is formed, with the nucleotides involved in the pseudoknot interaction forming the base of the stem region.*

The next adjacent downstream region consists of a large RNA structure with a basal stem (broadly nucleotides 259-268 based paired with nucleotides 330-344), a high NMIA reactivity region which has a largely open circle structure and a top stem region, which is designated here as stem loop E (SLE) (nucleotides 293-322) (Fig 4.10). In the 37 °C full-length and

fragment “RNAstructure” generated models, the nucleotides immediately upstream of SLE have a mostly open structure with a short hairpin. SHAPE reactivity values shows high reactivity in this region, indicating that the predicted short hairpin (nucleotides 278-289) in this region may not form, suggesting that region immediately upstream of SLE may be completely non-base-paired as reported by Li et al. [212]. In both the 28 °C fragment and full-length “RNAstructure” generated models, “RNAstructure” models predicted base pairing between nucleotides 273-275 and 324-326 however, the overlay of SHAPE reactivity values onto this region indicate that these nucleotides may be non-base-paired and therefore have a conformation similar to that seen in the 37 °C models. SLE is present in “RNAstructure” models for all conditions and was highly consistent with the SHAPE data when overlaid. The structure consists of a 30-35 nucleotide stem with a highly reactive apical loop, in the region of nucleotide 291-325. SHAPE data overlay from reactions carried out at 37 °C show greater reactivity in the stem region of this structure, indicating that the stem may have higher flexibility at this temperature.

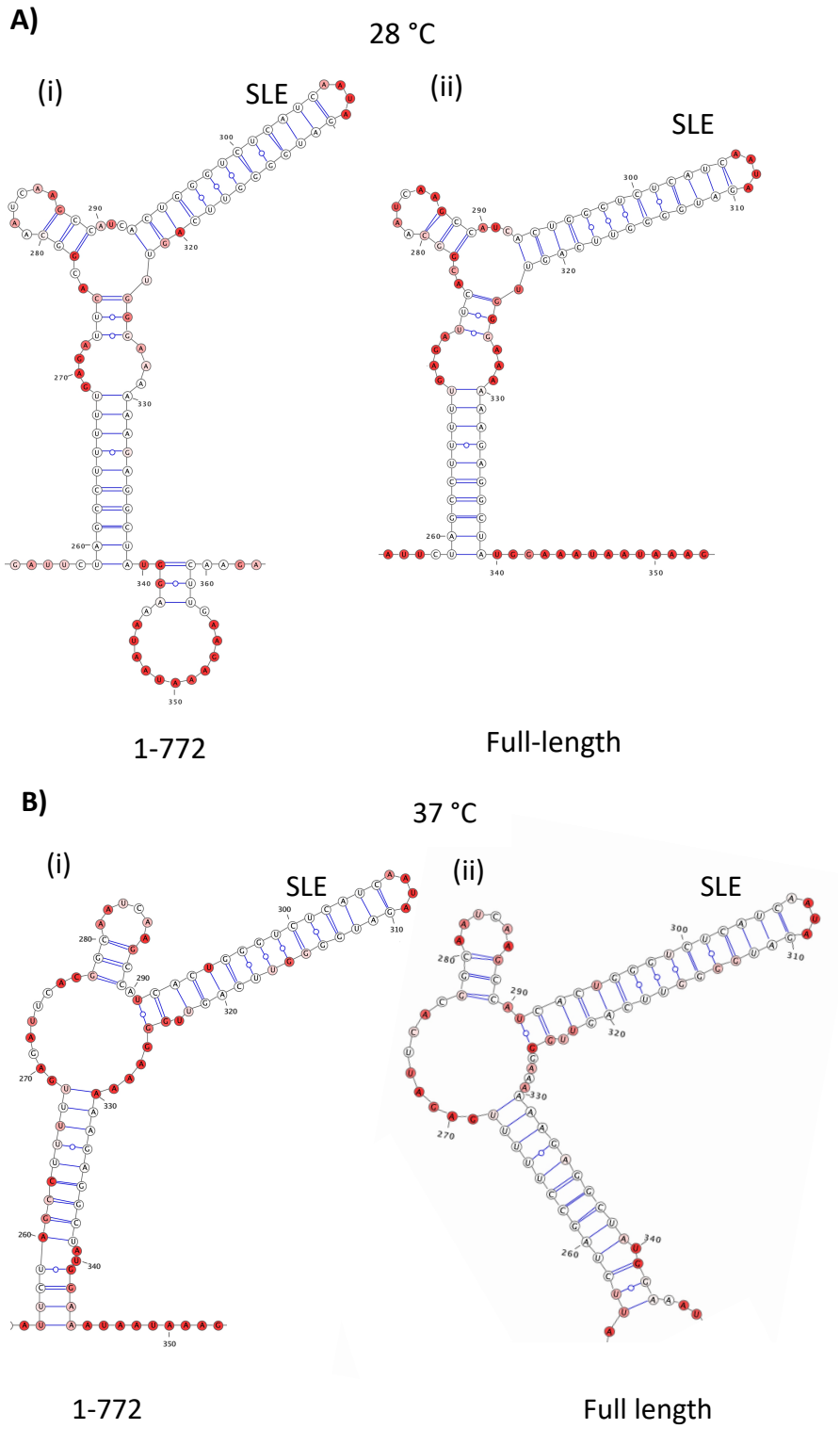


Figure 4.10 RNA structural models for the 5' region of ZIKV, SLE region
A) *28 °C (i) ZIKV 1-772 and (ii) full-length RNA.* ***B)*** *37 °C (i) ZIKV 1-772 and (ii) full-length RNA. RNA models for 28 °C show a more complex structure for nucleotides 270-290, compared with the more open structural arrangement of the 37 °C models. This may reflect different structural requirement between mammalian and mosquito hosts.*

At the 3’ end of the investigated sequence, a long RNA structure is present in the 28 °C 1-772 and full length and the 37 °C 1-772 “RNAstructure” generated models (Fig. 4.11). For the 28 °C full length model, SHAPE reactivity overlay is consistent with the “RNAstructure” model. For 1-772 at 28-°C, the overall reactivity is lower however, overlay SHAPE data is still relatively consistent with the model. For 37 °C 1-772, the nucleotides at the apical region of the RNA structure have a high level of reactivity, indicating that the apical region may have a more open structure than predicted by the model. “RNAstructure” generated models for 1-772 at 28 °C and 37 °C, and the full-length model for 28 °C, are in contrast to the model reported by Li et al. [212], in which this region consist of a short hairpin with an open loop of non-base paired nucleotides and an upstream short hairpin. The model reported by Li et al. has more similarity to the 37 °C full length model, which also contains a short hair pin with an open region, although the structure forms at an upstream position relative to the Li et al. model. In the 37 °C model, a second RNA structure is present in this region, consisting of a long duplex and an open apical top loop, which is present at the same position and has the same structure as reported by Li et al. As this region was close to the binding site of the second SHAPE primer, a number of nucleotides within this region have did not SHAPE reactivity data, therefore models produced of this region are likely to be less accurate than for the other regions investigated using SHAPE mapping.

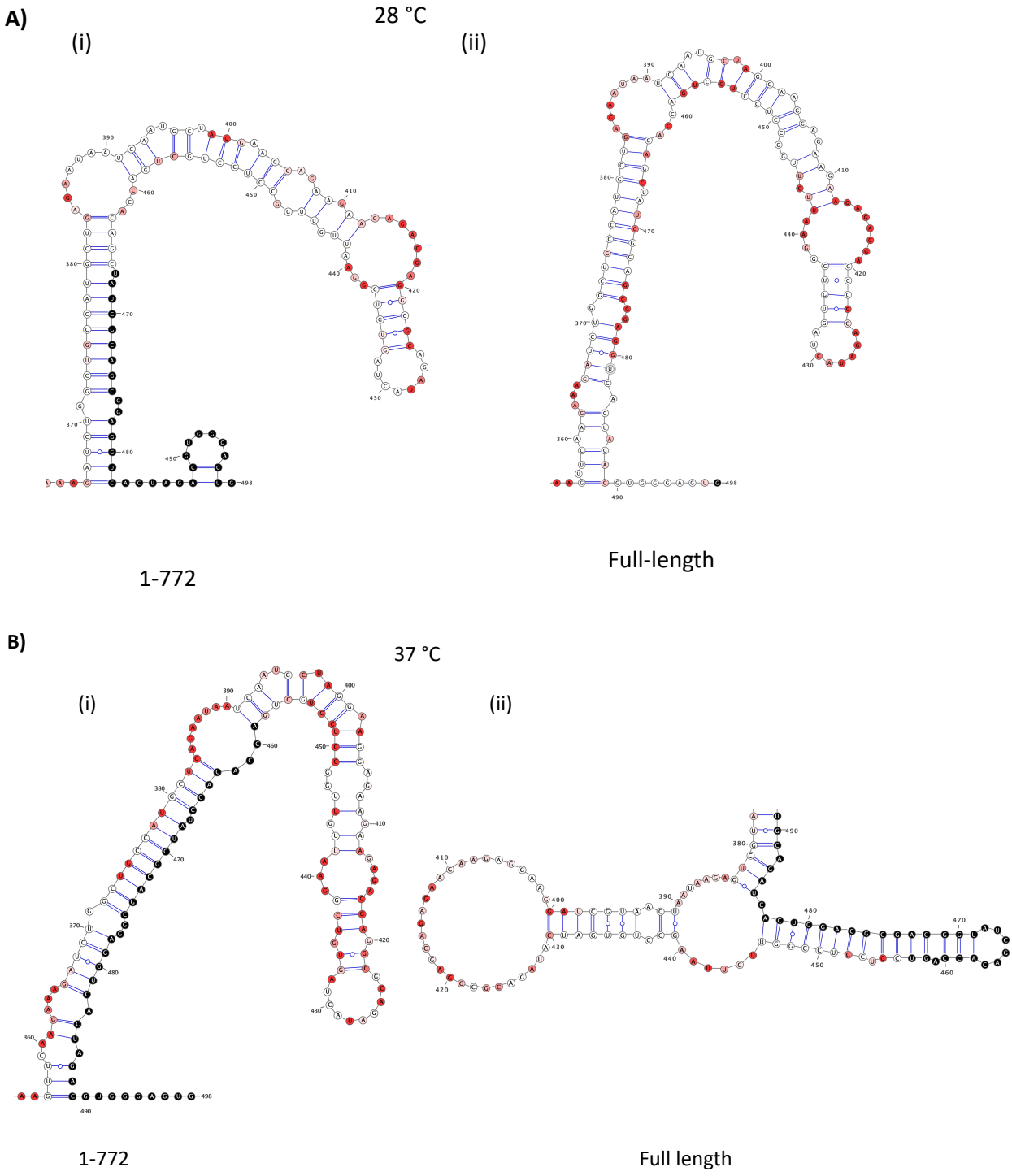


Figure 4.11 RNA structural models for the 5' region of ZIKV, 3' region of target sequence ***A)*** *28* °C *(i) ZIKV 1-772 and (ii) full-length RNA.* ***B)*** *37* °C *(i) ZIKV 1-772 and (ii) full-length RNA. A large single RNA stem is present in all models excluding the 37* °C *full length model, which contains two open regions. Since some data is missing for this region (shown in black) due to the binding site for the SHAPE primer, models for this region may not be accurate.*

4.5 Discussion

SHAPE mapping of the ZIKV 5′ UTR and adjacent coding region revealed the presence of several RNA structures. The first 3 structures are well documented and common to the insect transmitted flaviviruses [152,311]. The first notable structure is SLA and the structure of this RRE is highly conserved across the different temperatures explored and between full length and truncated RNAs. Whilst the presence of SLA is seemingly ubiquitous amongst flaviviruses, the structural conformation of SLA differs between flaviviruses transmitted by different hosts. For example, the insect-specific *Aedes* flavivirus has a shortened side stem and in tick-borne encephalitis virus SLA, the position the primary and side stem are reversed, such that the slightly elongated side-stem is found at the 5′ side of the structure [152]. Such differences are likely to play important roles in the host specific interactions of these viruses.

The structure of SLB was also highly consistent across the conditions tested. Since SLB contains the 5′ UAR sequence (which hybridises with complementary sequences in the 3′) the presence of a folded SLB indicates that the RNA is in a linear genome conformation. The ability of the flavivirus genomes to form both linear and circular genome conformations is essential for the flavivirus replication cycle. It is likely that both genome conformations are present at different stages of the virus lifecycle and that the total RNA in an infected cell is a mix of both genome conformations, perhaps in addition to lower levels of other genome conformations which are yet to be identified. It has previously been proposed that the circular genome conformation is thermodynamically favoured in the absence of *trans* activating proteins [323]. SLB was present in all ZIKV models generated, indicating that the RNA used in the *in vitro* SHAPE experiments predominantly forms the linear genome conformation. Since the linear conformation was seen at in both 28 °C and 37 °C, the switch between genome conformations is unlikely to be dependent on temperature. Since the *in vitro* SHAPE experiments lacked potential *trans* activating factors, such as the non-structural proteins, this may indicate that these factors could be required for formation of the circular genome conformation. Since the circular genome conformation is required for replication, this genome conformation may only become predominant in the presence of viral replication machinery, and be dependent on the action of *trans* activating factors, such as cellular or viral proteins. For example, DENV NS3 has been shown to bind to the 3′ SL in the 3′ UTR [324] and NS5 binds to SLA in the 5′ UTR [325], which may influence the folding conformation of the RNA.

The third conserved RRE, cHP, was also highly consistent between conditions. Since cHP has been shown to be essential for flavivirus replication and start codon selection, there is likely a high selection pressure for the cHP structure to be maintained at different temperatures and in closely related viruses. In addition to known RREs, SHAPE analysis revealed several uncharacterised RNA structures. In the region upstream of cHP (immediately upstream of nucleotide 160), SHAPE mapping results indicate either a single RNA structure (in 1-772 models) or a pair of shorter stem loops (full-length RNA models). In these models, nucleotides 168-173 and 201-207 either form the base of a larger RNA structure (1-772 model) or may potentially form a pseudoknot interaction in the full-length models, as these nucleotides have either no NMIA reactivity or low NMIA reactivity, indicating base pairing. Base pairing interactions formed by these nucleotides, either as the base of a stem or as a pseudoknot interaction, may play a role in virus replication.

Adjacent to this region, a large stem loop is present in all models, with an open loop around nucleotide 230. Whilst the 37 °C model displays differences in this stem loop compared to other models, the region formed of nucleotides 220-240 shows identical structure in all models, indicating that this structure may play a role in the virus lifecycle. Nucleotides 220-240 have been designated here as SLD. In the next upstream region, around nucleotide 260, a complex RNA structure is present, containing a base paired stem region, an open loop region and a second stem region (around nucleotides 290-325) with a non-base paired apical top loop. This second RNA stem loop is highly consistent in structure and NMIA reactivity, which may indicate that the structure plays a role in virus replication. Nucleotides 290-325 have been designated here as SLE. At the 3' end of the region of interest, long RNA structures are present in all models however, the NMIA reactivity for this region is inconsistent and data is also lacking for a significant number of nucleotides, meaning that the structural models for this region are likely inaccurate.

Although *in vitro* SHAPE mapping is a powerful technique for determining RNA structures, the technique has several limitations. The RNA used for structural analysis was generated by *in vitro* transcription, which was then folded at the relevant temperature. This process likely results in the formation of multiple conformational forms, some of which may not be representative of *in cell* RNA conformations. This concern can be addressed by performing *in cell* SHAPE, where the viral RNA is treated with SHAPE reagents in cellular environment and then purified for further analysis. The RNA models for ZIKV presented in this thesis, obtained using *in vitro* SHAPE, are highly comparable to the RNA models published by Li et al., where *in cell* SHAPE was performed [212]. The consistency between the *in vitro* and *in cell* models indicates that the ZIKV RNA models shown in this thesis are accurate and can therefore be used as a basis for mutagenesis and antisense LNA oligonucleotide design.

In addition, the process of *in vitro* transcription generates a large number of truncation products, which are included in the total RNA that is analysed. This again may result in non-native RNA conformations, which can influence the average reactivity value of nucleotides, as in the experimental data all conformations are sampled simultaneously. This issue is particularly relevant when examining full length genomic RNA, which contains a large number of truncated RNA products, due to the length of the RNA. One technique which can eliminate this limitation is the use of in gel SHAPE, in which, prior to SHAPE analysis, the mixed RNA populations are separated by native agarose gel electrophoresis. This approach has been used to separate and analyse monomeric and dimeric species of HIV-1 packaging signal RNA [326], despite the technical challenges associated with the use of native agarose gels. An additional factor to consider when using the *in vitro* SHAPE methodology is that the NMIA reactivity of individual nucleotides may be variable due to RNA breathing in solution and that flexible RNA regions, which have the potential to adopt several conformations, may be variable in NMIA reactivity. These issues may be partially addressed by utilising averaged data from multiple experimental repeats in RNA structural models, as inconsistencies between repeats will be minimised when the data is averaged.

Chapter 5 Reverse genetic analysis of novel ZIKV RNA-RNA interactions

5.1 Introduction

The ZIKV outbreak in 2015-2016 highlighted the importance of ZIKV as a human pathogen. To date, there are no licenced antiviral agents or vaccines available against ZIKV. Flavivirus RNA structures are known to be critical for efficient virus replication and therefore represent potential targets for novel antiviral agents or attenuated virus vaccine development. Having previously demonstrated that it is possible to target RNA structures in CHIKV, leading to inhibition virus replication (see chapter 3), the next aim was to apply the same hypothesis to ZIKV.

In order to impair ZIKV replication by targeting RREs, functional structures with the ZIKV genome needed to be identified. The structural conformation of the ZIKV 5' UTR and adjacent coding region was determined using the biochemical SHAPE mapping, which generated a model of RNA structures at single nucleotide resolution (see chapter 4). However, such models do not provide information about the functional role of RNA structures with ZIKV replication. RNA models derived from SHAPE mapping have previously been used as a guide for the design of mutations which disrupt base pairing within RNA structures, allowing the functional role of a given structure within virus replication to be determined. For example, disruption of the CHIKV RRE SL165 by mutagenesis impaired replication in both infectious virus and sub-genomic replicon systems, indicating that the RNA structure plays a role in viral genome replication [2].

The ZIKV 5' UTR and adjacent coding region RNA structural model contains several RNA elements which are highly conserved amongst flaviviruses, namely SLA, SLB and cHP. Since these RNA elements are structurally comparable to those found within other flaviviruses [327], it is likely that SLA, SLB and cHP carry out analogous functions in ZIKV (described in section 1.3.5.1). In addition to the conserved structures, SHAPE mapping identified several novel RNA elements including SLD, SLE and a pseudoknot interaction. In order to determine if these novel interactions play a functional role in the ZIKV lifecycle, a ZIKV infectious clone system was utilised. Using the ZIKV infectious clone system, it is possible to introduce specific mutations which disrupt structural features, whilst maintaining the WT amino acid sequence. By recovering these mutant viruses and comparing their replication to WT ZIKV, it is possible to determine if the disrupted RNA structure plays a functional role in ZIKV replication. Using this approach, the function of SLD, SLE and the pseudoknot interaction in ZIKV replication was investigated.

In order to complement the findings of the reverse genetic analysis of the ZIKV RNA structures and to validate the potential of such structures to act as targets for antiviral agents (as demonstrated for CHIKV in chapter 3), antisense oligonucleotides were designed based on ZIKV RNA structural models presented in chapter 4, with the aim of disrupting base pairing interactions within the RREs. If ZIKV replication is inhibited by agents targeting cis acting RNA

elements (shown by reverse genetics to be required for efficient ZIKV replication) this would further confirm the importance of those RNA elements in the ZIKV lifecycle. For example, targeting of a cis-replicating element (CRE) in HCV (which had previously been shown using mutagenesis to have a functional role in HCV replication [328]) using LNA oligonucleotides further validated the importance of this RNA element for HCV RNA replication [294].

RNA structures within flavivirus genomes have previously been targeted using antisense phosphorothioate oligonucleotides. In phosphorothioate oligonucleotides, one of the non-bridging oxygen atoms is replaced with a sulphur atom, increasing the nuclease resistance of the oligonucleotide however, this also reduces target binding affinity. In order to address the reduction in binding affinity, the authors used modified phosphorothioate oligonucleotides (MPO), in which the C-5 atoms of uridines and cytidines were replaced by propynyl groups (which contain a triple carbon bond, e.g. CH_3-C≡C-R) to improve binding affinity. An MPO designed to target nucleotides 99-113 in DENV-2 was shown using fluorescence microscopy to impair infectious DENV replication. The nucleotides targeted by the MPO correspond to the position of ZIKV SLB, indicating that this region may be available for binding during replication. In addition, an MPO targeting the 3′ region of the genome also impaired replication. The authors suggested that the modified phosphorothioate oligonucleotides function through the action of RNase H, thereby degrading the DENV genomic RNA [329].

In addition, RNA structures have been targeted in WNV using phosphorodiamidate morpholino oligomers (PMOs) [330]. PMOs are uncharged, nucleic acid-like molecules which contain a methylenemorpholine backbone (replacing ribose or deoxyribose rings), linked by non-ionic phosphorodiamidate groups and containing standard DNA bases. Compared with DNA oligonucleotides, PMOs have increased nuclease resistance and do not induce RNase H activity [331,332]. In order to improve cell delivery, an arginine rich peptide was covalently conjugated to the 5′ end of the PMOs, as arginine rich peptides have been shown to promote cellular uptake [333,334]. PMOs were designed to target the extreme 5′ end of the WNV genome, the AUG-1 and AUG-2 start codon and the 5′ CS. PMOs targeting the 5′ genome end and AUG-1 caused reduction in WNV titre at 48 hours post infection. In addition, a PMO targeting the 3′ CS and 3′ SL impaired infectious WNV replication. Using a reporter construct, consisting a *Renilla* luciferase reporter fused in frame with the N-terminal 31 amino acids of the C protein and flanked by the 5′ and 3′ UTRs, the 5′ end and AUG-1 PMO were shown to reduce luciferase expression at 5 hours post infection. The authors concluded that the PMOs function by inhibiting translation of the viral genome, through steric hinderance [330].

In this chapter, reverse genetics approaches are used to determine if RNA elements identified using SHAPE mapping in chapter 4 have a functional role in the ZIKV lifecycle, using a ZIKV infectious clone system. The ZIKV SHAPE map is further validated, by using the SHAPE map as a basis for the design of antisense oligonucleotides which bind and impair RNA elements with the 5′ UTR and adjacent coding region. Oligonucleotides were screened using a nano-luciferase expressing ZIKV in order to determine the effect of oligonucleotide transfection on ZIKV replication in an infectious virus system and using a nano-luciferase expressing sub-genomic replicon to determine the effect oligonucleotide transfection on ZIKV genome replication.

5.2 ZIKV mutagenesis base pair changes

In order to investigate the novel RNA secondary structures and the pseudoknot interaction identified in SHAPE analysis, mutagenesis was performed on these structures, with the aim destabilising the RNA-RNA interactions requires for RNA structure/ pseudoknot formation (Fig. 5.1). All mutations introduced were designed to maintain the correct amino acid sequence and have the potential to be compensated with further substitutions, in order to restore base pairing with an alternative sequence. Mutant sequences were purchased (Genewiz) and overlap extension PCR was used to make the purchase sequence compatible for ligation into the ZIKV IC DNA construct (appendix Fig. 4).

To investigate the potential pseudoknot interactions, four substitutions were introduced, which were predicted to prevent pseudoknot formation. Two mutations were introduced into the upstream pseudoknot region ($A_{170}U$, $C_{173}G$) and a further two mutations were introduced into the downstream pseudoknot region ($G_{203}A$, $G_{206}C$). For individual RNA structures, mutations were introduced into SLD ($U_{230}C$, $G_{233}C$, $C_{236}G$, $C_{239}A$, $C_{248}G$, $G_{251}C$, $G_{254}A$, $U_{257}C$). $U_{230}C$ changes the nucleotide sequence the top loop of SLD, with the aim of disrupting sequence specific interactions between the top loop of SLD and other regions of the genome/ viral or cellular proteins. The remaining SLD mutations were designed to disrupt base pairing within the stem region, to prevent the duplex stem from forming. Mutations were also introduced into SLE ($U_{299}G$, $C_{302}A$, $C_{305}A$, $U_{308}C$, $A_{309}C$, $A_{311}C$, $U_{317}A$, $A_{320}C$). $U_{308}C$ and $A_{309}C$ alter the nucleotide sequence of the top loop, whilst the remaining mutations were designed to disrupt base pairing with the stem region. In order to predict the effect of each set of mutations, mutant sequences were modelled using the mfold web server. The Mfold predictions indicated that the mutations were sufficient to disrupt the target RNA structures.

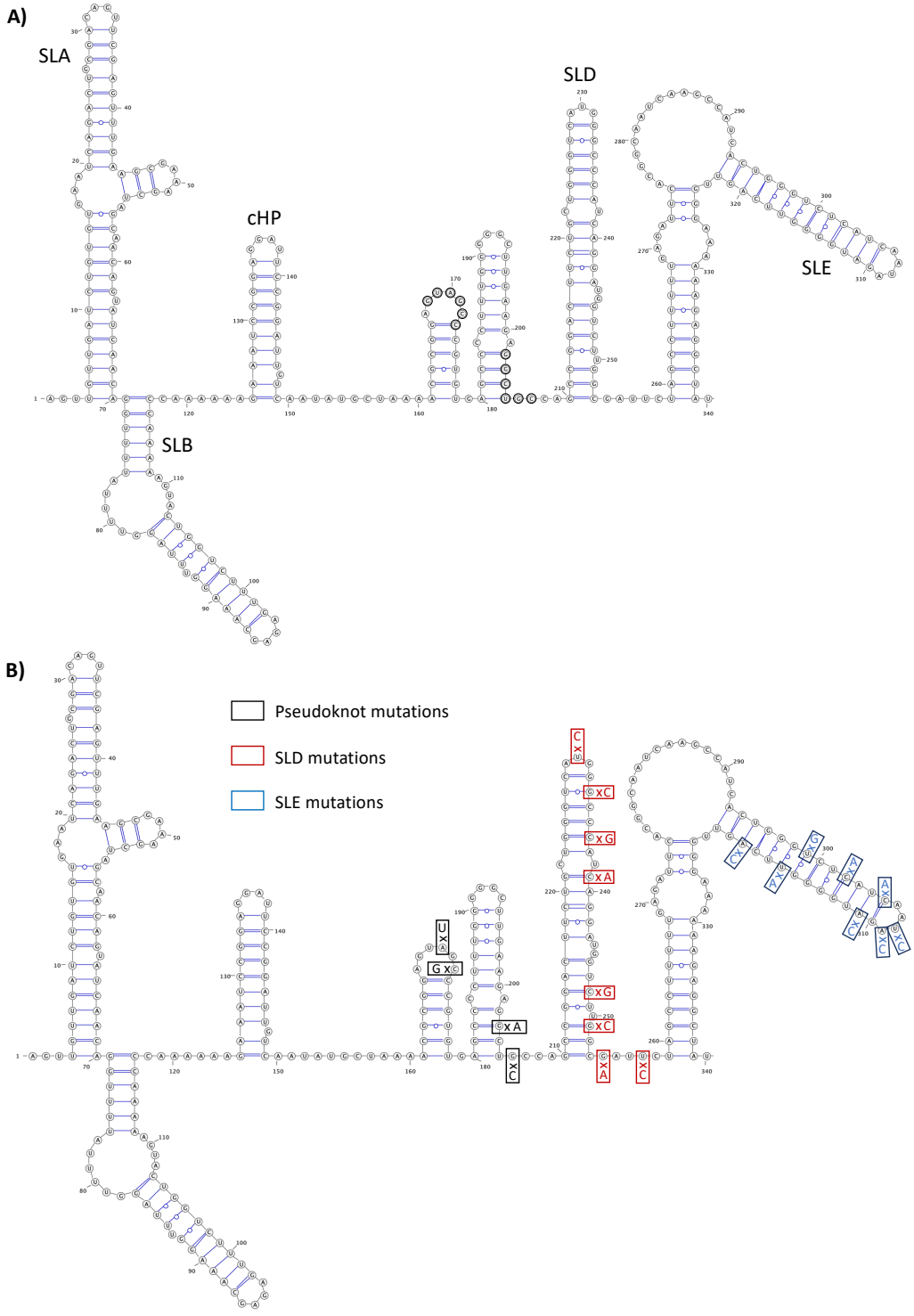

A)
SLA
cHP
SLD
SLE
SLB
B)
Pseudoknot mutations
SLD mutations
SLE mutations

Figure 5.1 RNA structural models with mutation positions
A) *Model of the RNA structures within the 5' UTR and adjacent coding region of ZIKV. Nucleotides involved in the formation of the pseudoknot interaction are indicated by black circles. The model presented here is based on SHAPE reactivity for full-length ZIKV RNA at 37* °C ***B)*** *Model of ZIKV 5' UTR and adjacent coding region showing the location of mutant residues. Pseudoknot mutation are indicated in black, SLD mutations are indicated by red circles, SLE mutations are indicated in blue. For SLD and SLE, mutations were designed to disrupt RNA base pairing and pseudoknot mutations were deign to disrupt the pseudoknot interaction, whilst maintaining the WT amino acid sequence.*

5.3 Analysis of ZIKV mutations

5.3.1 The effect of ZIKV SLE mutations on ZIKV replication

The ZIKV RNA structure SLE has not previously been characterised, therefore in order to investigate the potential functions of SLE, several mutations were introduced with the aim of destabilising SLE formation (M-SLE). To investigate the effects of SLE mutations on ZIKV growth, WT and M-SLE ZIKV was recovered from C6/36 cells. Cells were transfected with 2.5 µg of either WT or M-SLE ZIKV IC RNA and cell supernatants were collected at 5-, 7- and 9-days post transfection, after which ZIKV titre was determined by plaque assay on Vero cells (Fig. 5.2). Due to the growth rate of ZIKV in cell culture, 5 days post infection is the first time point at which useable amounts of infectious ZIKV can be collected. ZIKV titre for both WT and M-SLE transfected wells was highly comparable, indicating that for both conditions ZIKV was replicating at a similar rate.

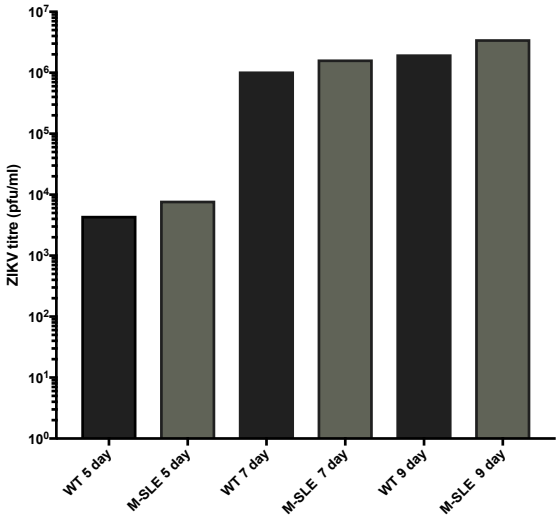


Figure 5.2 Recovery of WT and M-SLE ZIKV
C6/36 cells were transfected with 2.5 µg WT or M-SLE infectious RNA. Virus was collected at 5-, 7- and 9-days post transfection and virus titre was determined by plaque assay. M-SLE had no noticeable effect on ZIKV titre. Data from a single repeat (n=1).

C6/36 derived WT and M-SLE ZIKV was next used to infect Huh7 cells (MOI 2) and collected virus supernatant at 24, 48 and 72 hours post infection, in addition to lysing the cells for analysis of protein expression (Fig 5.3). Cells infected with WT ZIKV produced less virus compared with M-SLE by 24 hours post infection, however by 48 and 72 hours post infection. the difference between WT and M-SLE became negligible. Western blot analysis showed no difference (n = 1) in expression of both ns3 and ns5 proteins when comparing WT and M-SLE. In addition, infection of C6/36 cells with WT and M-SLE ZIKV (MOI 1) generated in Huh7 cells showed a small difference in virus titre, but in the context of the overall data this was deemed not significant (appendix Fig. 5).

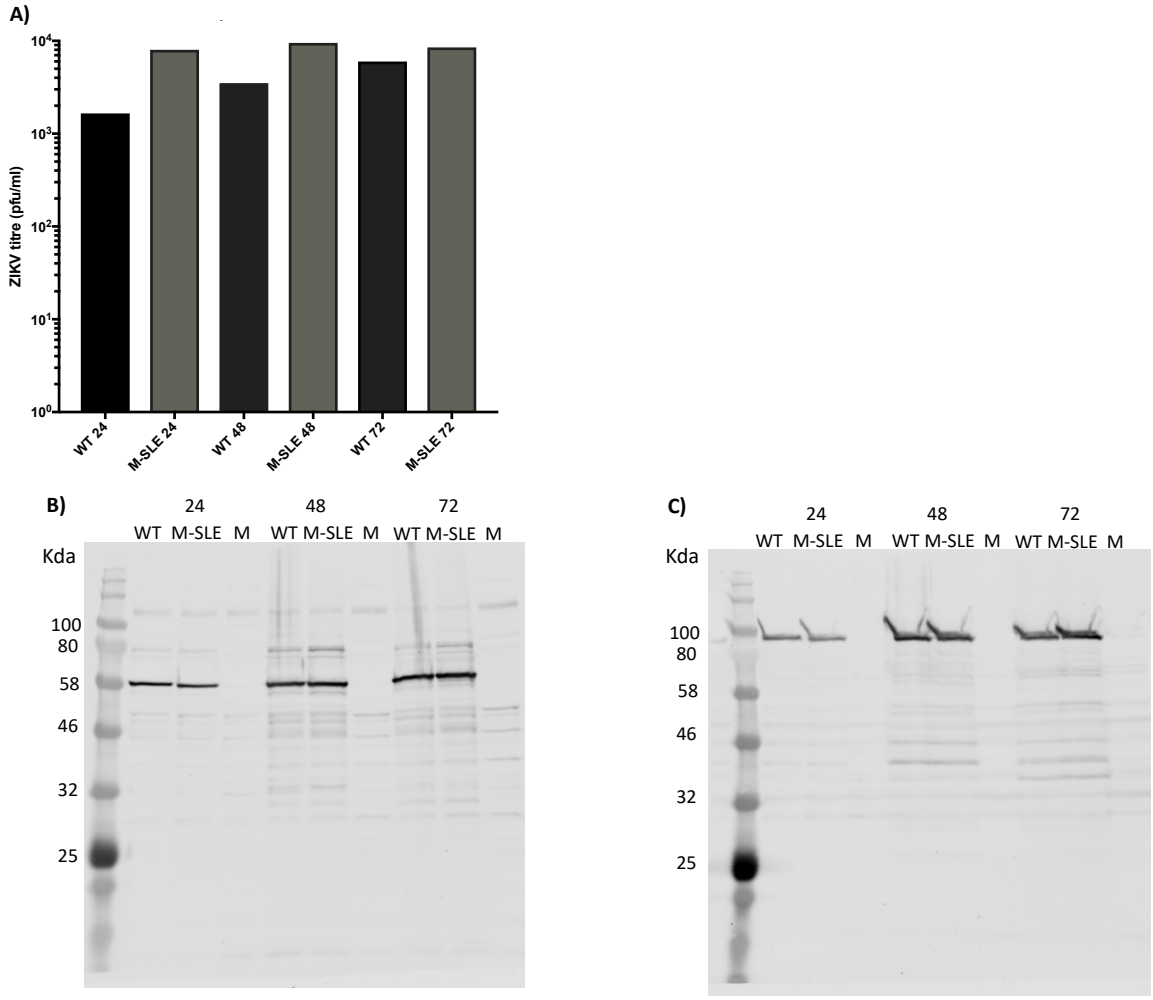


Figure 5.3 Infection of Huh7 cells with WT and M-SLE ZIKV

Huh7 cells were infected with WT and M-SLE ZIKV (MOI 2) (derived from C6/36 cells). Supernatants were harvested at 24, 48 and 72 hours and cells were lysed in IP lysis buffer. Virus titre was determined by plaque assay ***(A)*** *and levels of ns3* ***(B)*** *and ns5* ***(C)*** *expression were examined by western blot. Primary antibodies as follows: anti-nsp5 (1:1000, rabbit polyclonal, in-house) or anti-nsp3 (1:1000, rabbit polyclonal, in-house) and secondary antibody (IRDye® 680LT Donkey anti-Rabbit; Li-Cor). (n = 1).*

In order to further investigate the effects of M-SLE, Vero cells were transfected with 1 µg of either WT or M-SLE ZIKV IC RNA and virus was collected 5 days post transfection (Fig 5.4A). WT and M-SLE transfected wells produced highly similar virus titres, indicating that the virus replication rate was similar. Due to the high mutation rate of RNA viruses, it was hypothesised that the M-SLE ZIKV may contain reversion mutations which restore the WT SLE phenotype during the long incubation period needed to produce useable amounts of virus. In order to gain a more direct measure of ZIKV replication, an infectious centre assay was performed by electroporating Huh 7 cells and incubating the electroporated cells on a Vero cell monolayer (Fig 5.4B). Virus produced in the Huh 7 cells infects the Vero monolayer, resulting in the formation of plaques which can be fixed and visualized. This approach allows the replication rate of WT and M-SLE to be measured directly from the input RNA, without the need for the 5-day incubation period to generate virus stocks and the subsequent passage in cell culture to obtain the experimental result, therefore limiting the potential for reversion mutations to occur. The infectious centre assay shows highly similar levels of replication between WT and M-SLE, indicating that the mutations had no effect on ZIKV replication, suggesting that the lack of different phenotype relative to WT was not due to reversion mutations.

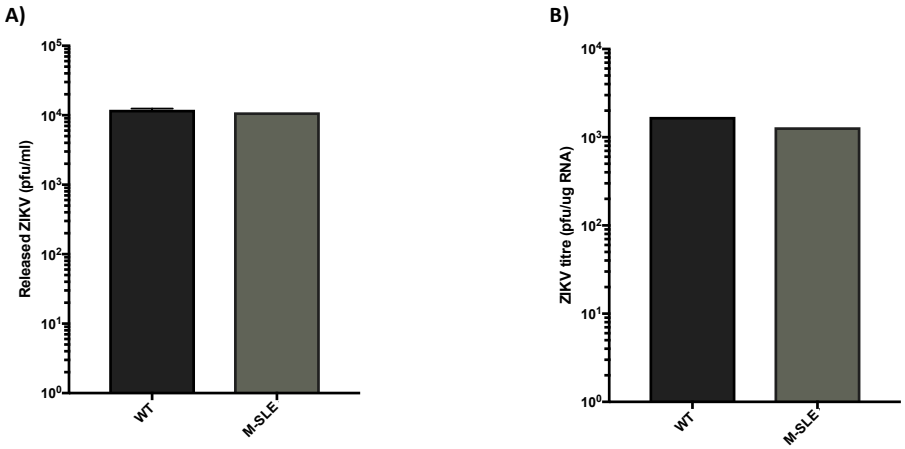


Figure 5.4 Analysis of ZIKV WT and M-SLE in mammalian cells

A) *Vero cells were transfected with WT or M-SLE 1 µg. ZIKV was collected 5 days post transfection and virus titre was determined by plaque assay.* ***B)*** *ZIKV WT and M-SLE Infectious centre assay. Huh 7 cells were electroporated with 1 µg WT or M-SLE RNA and serially diluted from 10^{-1} to 10^{-6} in DMEM. Dilutions were added to Vero cell monolayers and incubated for 5 days. Cells were fixed at 5 days and ZIKV titre was determined (n = 1).*

To confirm M-SLE sequence, total RNA was extracted from WT and M-SLE transfected C6/36 cell supernatant, 9 days post transfection (Fig 5.5). Sanger sequencing confirmed that the M-SLE mutations were maintained, confirming that no reversion mutations had occurred within progeny ZIKV genomes within the capsid and prM coding sequences, although it is possible that mutations may have occurred in other regions of the genome.

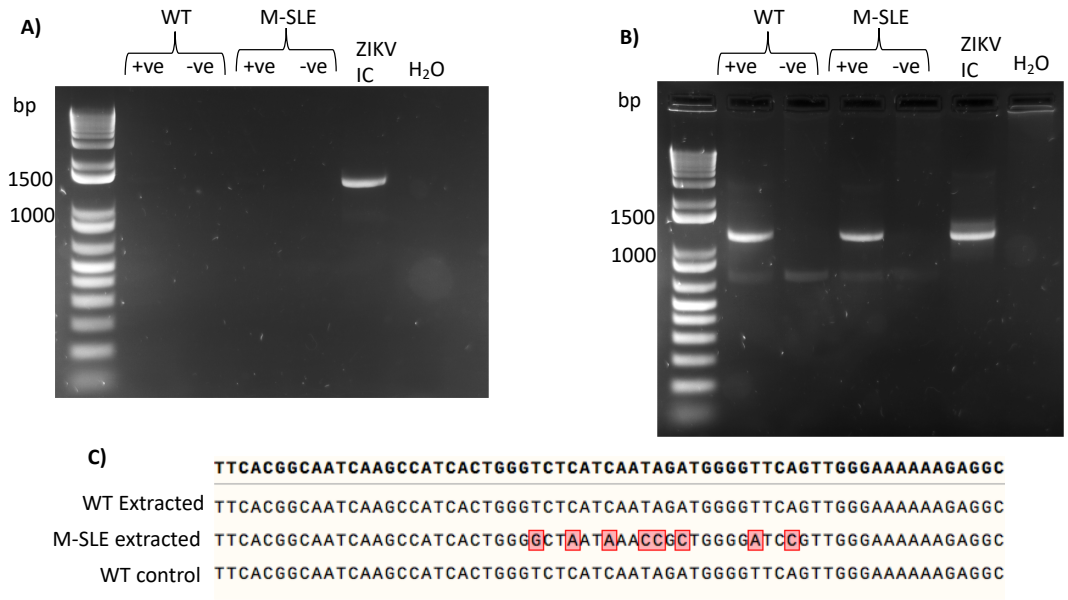


Figure 5.5 2 step PCR reactions for RNA recovered from WT or M-SLE ZIKV
Step 1 ***(A)*** *and step 2* ***(B)****. RNA extracted from ZIKV infected (+ve) and uninfected (-ve) cells was reverse transcribed and used as a template for 2-step PCR. ZIKV IC DNA was included as a positive control and the water used to set up the reactions was included as a negative control.* ***(C)*** *Sequencing data for WT ZIKV and M-SLE ZIKV extracted RNA. (n = 1).*

5.3.2 Effects of ZIKV Pseudoknot and SLD mutations on ZIKV replication

Both the pseudoknot interaction and SLD RNA elements have not previously been characterised. To investigate the potential roles of these elements in ZIKV replication, mutations were introduced to disrupt the RNA interactions. The ZIKV pseudoknot region was identified by sequence analysis during the SHAPE analysis discussed in chapter 4. Based on the full-length RNA structure models, which are the most biologically relevant, the pseudoknot mutations prevent the predicted interaction between the two pseudoknot regions (ZIKV M-PK). The RNA structure SLD has not previously been characterised, so mutations were designed to disrupt both the top loop lop and stem region of the RNA structure, in order to determine if the SLD functions during replication (ZIKV M-SLD).

Prior to analysis of ZIKV mutations, virus stocks were generated by transfection of 2.5 µg WT, M-PK or M-SLD ZIKV IC RNA into Vero cells (Fig 5.6). Cell supernatant was harvested at 5 days post infection and virus titre was determined by plaque assay on Vero cells. Compared with WT transfected cells, SLD RNA transfected cells produced around 30% fewer virus particles, whereas M-PK RNA transfected cells showed a greater reduction in virus titre, reducing viral titre by ~1 log compared with wild type transfected cells ($P \leq 0.005$). This indicates that the M-PK mutations impair ZIKV replication, suggesting that the pseudoknot region plays a role in the virus lifecycle.

Since the generation of ZIKV stocks relies upon transfection of infectious RNA, variations in transfection efficiency and RNA quality may produce variation in the viral titre. In order to further validate the results seen in the virus generation experiment, Vero cells were infected with either WT or mutant virus at an equal MOI of 0.005 (due to the low titre of the M-PK virus stocks, this was the highest MOI achievable in a 24 well plate). Since all cells are infected with the same MOI of ZIKV, this experiment examines how WT ZIKV replication compares with M-PK and M-SLD ZIKV when cells are infected with the same number of infectious virus particles.

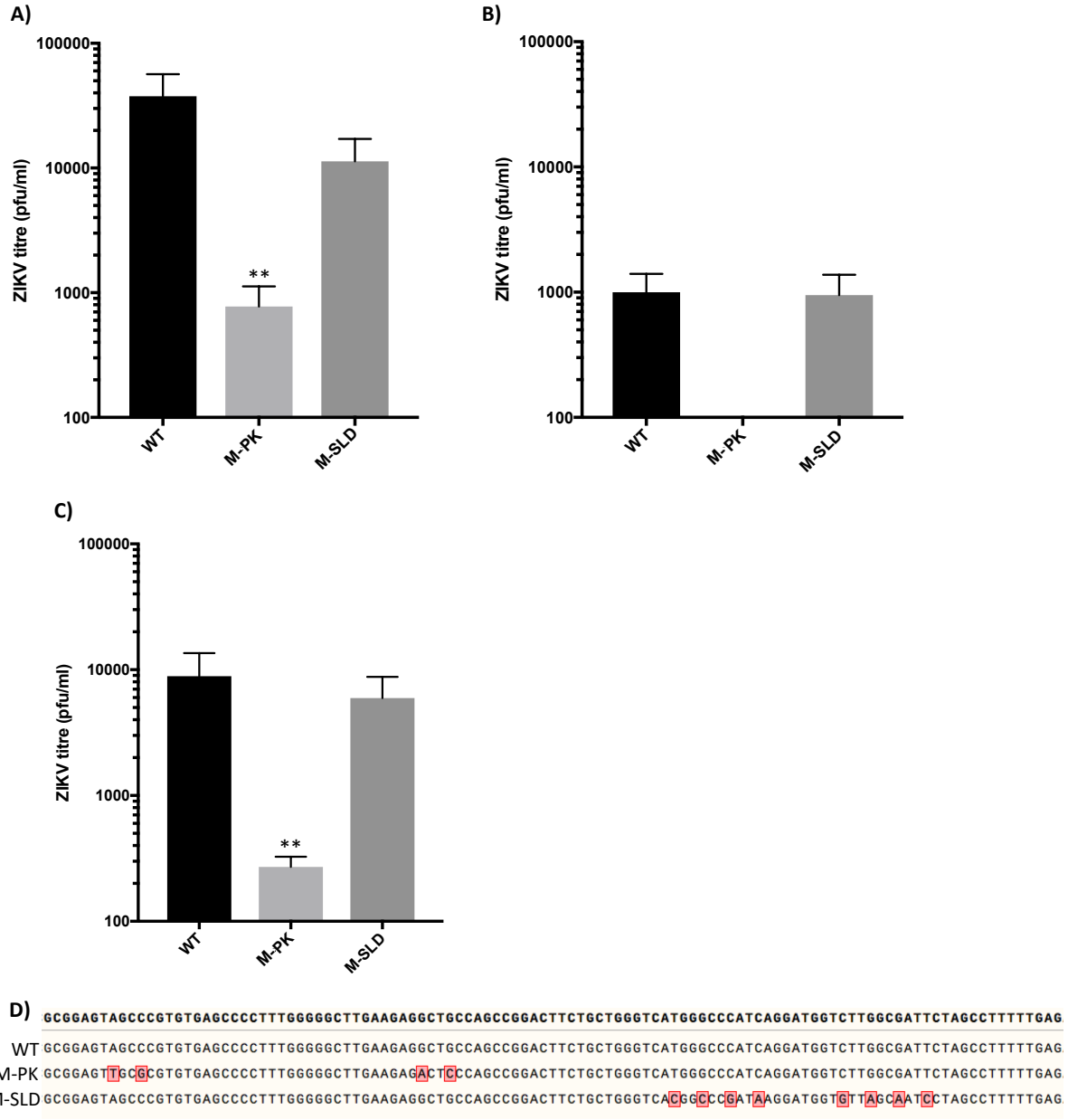


Figure 5.6 Analysis of ZIKV M-PK and M-SLD mutants

A) *Generation of ZIKV WT and mutant stocks. Vero cells were transfected with RNA generated by in vitro transcription and incubated for 5 days. Virus titre was determined by plaque assay on Vero cells. Transfection of Vero cells with M-PK mutant RNA resulted in production of significantly less ZIKV, compared with both WT and the M-SLD mutant. Vero cells were infected with WT, M-PK or M-SLD ZIKV (MOI = 0.005) and supernatant was harvested at 24* ***B)*** *and 48* ***C)*** *hours post infection. At 24 hours post infection, M-PK ZIKV was undetectable by plaque assay.* ***D)*** *Total RNA for WT and mutant ZIKV was extracted at 24 and 48 hours post infection. RNA was reverse transcribed and the region of interest was amplified using a two-step PCR reaction. ** = P ≤ 0.005 (Two-tailed T-test vs scrambled LNA) (n = 3). Error bars represent standard error of the mean.*

After the equal MOI infection, cell supernatant was harvested at 24 and 48 hours post infection and virus titre was determined by plaque assay. Viral titre from cells infected with M-SLD ZIKV had similarly levels to WT at both 24 and 48 hours post infection, indicating that the M-SLD mutations do not impair ZIKV replication and that the SLD does not function during the virus lifecycle in Vero cells. In contrast, the M-PK ZIKV mutations had a profound effect on ZIKV replication. At 24 hours post infection, no virus was detected in the viral supernatant, indicating that the virus was either not replicating or only replicating at very low levels. By 48 hours post infection, M-PK ZIKV was detectable by plaque assay at a low titre. Compared with WT ZIKV, the M-PK ZIKV showed a 97% reduction in viral titre ($P \leq 0.005$), indicating that the M-PK ZIKV maintained the ~1 log difference in viral titre seen in the virus generation experiment. This finding indicates that the pseudoknot region of the ZIKV genome plays a critical role in the virus lifecycle. The M-PK mutations also altered plaque morphology (Fig. 5.7), resulting in formation of small, pinpoint size plaques (~0.5 mm) compared with the medium size plaques (~2 mm) seen for the other ZIKV mutant and the WT ZIKV.

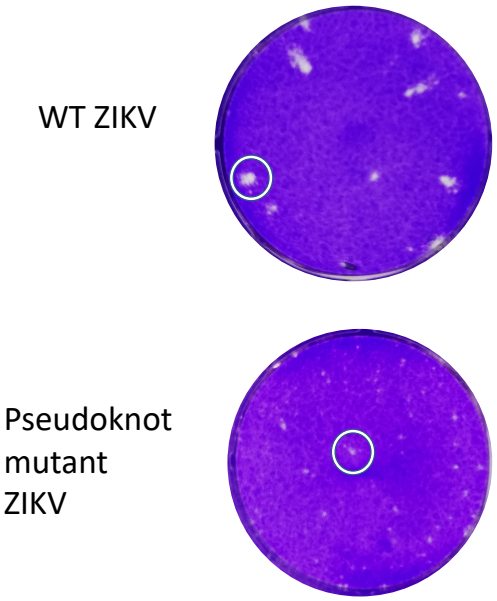


Figure 5.7 Difference in plaque sizes of WT and M-PK ZIKV
ZIKV plaque assays were performed in 12-well plates, with a 0.8% Methylcellulose (in DMEM) overlay and incubated for 5 days. ZIKV M-PK produced smaller, pinpoint sized, plaques compared with WT ZIKV. Reduction in plaque size is an indication of viral fitness and indicates that ZIKV M-PK is unable to efficiently lyse the Vero cell monolayer.

In order to investigate the possibility of reversion mutations within the ZIKV sequence, viral total RNA was extracted from the infected cells and analysed using Sanger sequencing. Both the M-PK and M-SLD sequence contained the introduced mutations, indicating that reversion had not taken place within the capsid and prM coding sequences at 24 and 48 hours post infection. Though again, it is possible that mutations had occurred in other regions of the genome.

5.4 ZIKV Antisense oligonucleotides

5.4.1 Initial antisense LNA oligonucleotides targeting ZIKV RREs

In order to further validate the RNA map produced in chapter 4 and to investigate the potential of ZIKV RREs to act as antiviral targets, a panel of antisense LNA oligonucleotides were designed to bind RNA elements in the ZIKV genome. Compared with DNA oligonucleotides, LNA oligonucleotides have improved binding specificity, nuclease resistance and do not induce RNase H activation. LNA oligonucleotides are deigned to bind to target sequences, thereby preventing RNA-RNA or RNA protein interactions required for efficient ZIKV replication. Since antisense oligonucleotides were designed based on the RNA model acquired by SHAPE mapping, the antiviral activity of antisense LNA oligonucleotides validates the RNA model.

The panel of LNA oligonucleotides described in following section is designated as ZIKV LNA panel 1 (ZLP1) (appendix Fig. 6). LNA-29 to targets the top loop and side stem of ZIKV SLA. Since SLA has been shown to play an important role in flavivirus replication [207], it was reasoned that SLA would represent a good target for antiviral agents and would provide a proof of principle for the use of antisense oligonucleotides against ZIKV RREs. Similarly, LNA-142 was designed to bind the region of the genome containing the 5' CS element, in order to block genome cyclisation. Additional LNA oligonucleotides were designed to target novel RNA interactions. Both LNA-158 and LNA-191 were designed to target the pseudoknot region. LNA-213 was designed to bind SLD, with the aim of blocking the top loop and out-competing base pairing within the stem region.

In order to determine if any ZLP1 LNA oligonucleotides were able to impair ZIKV replication, a Nano-luciferase expressing ZIKV was employed to improve efficiency. In this system ZIKV replication was quantified using luciferase expression, allowing ZIKV replication levels to be determined on the same day, compared with the 5 days required to obtain results using plaque assays. In the ZIKV N-luc IC construct, a nano-luciferase reporter is fused with a truncated form of the capsid protein. The truncated capsid was included to maintain the RNA structures within this region, which may be required for ZIKV replication. A foot and mouth disease virus (FMDV) 2A sequence was included upstream of the nano-luciferase sequence. The FMDV 2A sequence is commonly used to mediate co-translational protein cleavage [335]. Downstream of the FMDV 2A sequence a codon optimised full length ZIKV sequence is present, including a codon optimised full length ZIKV C gene to allow production of infectious virus particles (Fig. 5.8).

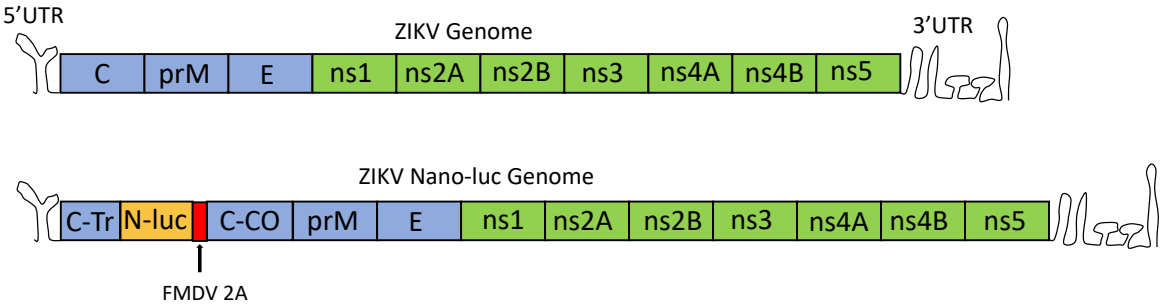


Figure 5.8 Genome schematics of ZIKV and N-luc ZIKV

A) *WT ZIKV genome and* ***B)*** *Nano-luciferase (N-luc) ZIKV genome. In the N-luc genome, the capsid protein is truncated (C-Tr) and fused with the N-luc gene. The capsid sequence is included downstream of N-luc in order to maintain RNA structures found within this region, which may be required for RNA replication. Inclusion of the FMDV 2A sequence mediates self-cleavage (at the C-terminus of the FMDV 2A sequence), releasing the N-luc protein and allowing expression of the viral polyprotein. A codon optimized capsid (C-CO) is expressed after FMDV 2A, in order to maintain the viral polyprotein.*

The ZIKV N-luc constructs therefore contains two copies of the capsid gene, a truncated capsid sequence found downstream of the 5' UTR, and a codon optimised capsid gene found after the FMDV 2A sequence. Mutations within the truncated capsid coding sequence can be introduced whilst maintaining the WT amino acid sequence using alternative codons however, changes in codon use may result in defects in replication due to codon usage bias. However, since ZIKV particles in this system are formed using the codon optimised capsid protein, codon changes in the truncated capsid protein should not affect ZIKV replication. Since nano-luciferase expression is dependent on replication of the ZIKV genome, levels of nano-luciferase in cell supernatant and in lysed cells can be used as a measure of ZIKV genome replication.

ZLP1 LNA oligonucleotides were transfected into ZIKV N-luc infected Vero cells (MOI 0.1) and luciferase expression was determined at 24 and 48 hours post infection. Compared with transfection of a scrambled LNA, transfection of the ZLP1 LNA oligonucleotides did not result in significant differences in luciferase expression, indicating that ZLP1 LNA oligonucleotides are not able to inhibit ZIKV replication (appendix Fig. 7-9).

5.4.2 Design of antisense oligonucleotides targeting ZIKV RREs

Since ZLP1 LNA oligonucleotides were not successful at inhibiting ZIKV replication, in order to identify potential binding sites for antisense oligonucleotides, an expanded number of target sequences were investigated. Due to the complex nature of RNA molecules, many sites within the 5' UTR and adjacent coding region may be in accessible for binding to *cis* acting agents. By designing a larger panel of antisense oligonucleotides, the number of available oligonucleotide binding sequences is expanded, therefore there is a higher likelihood of identifying accessible target sequences. The wider screen was performed using non-LNA modified DNA oligonucleotides.

As the target RNA sequences have highly folded complex structures, LNA oligonucleotides designed to target such regions can be susceptible to secondary structure formation and hybridisation issues, due to the improved binding properties of the LNA nucleotides. Since LNA oligonucleotides are locked in an ideal conformation for Watson-Crick base pairing, interactions between complimentary sequences in LNA oligonucleotides may bind with very high affinity, therefore preventing the LNA oligonucleotide from binding the target sequence. DNA oligonucleotides are less susceptible to these issues and are also significantly cheaper than LNA oligonucleotides, allowing a greater number of oligonucleotides to be tested, therefore maximising the number of potential target sequences screened. The screening pipeline of DNA oligonucleotides was therefore used to identify which RREs within the ZIKV genome are available for targeting.

This wider screen was performed using a ZIKV Nano-luciferase expressing sub-genomic replicon, as the during the CHIKV LNA experiments, the sub-genomic replicon proved to have a higher sensitivity to antisense oligonucleotides, likely as a result of the replicon DNA and oligonucleotide being premixed prior to transfection. The higher sensitively of the replicon system may negate the reduced half-life and RNA binding T_m properties of DNA oligonucleotides. The RNA target for each DNA oligonucleotide used in the screen is shown in table 5 and figure 5.9. “Connector oligos” are designed to bind two spatially separated region of the ZIKV RNA and consist of 5’ and 3’ antisense regions interspaced by a scrambled sequence.

Table 5 DNA oligonucleotide targets

	ZIKV Oligonucleotide	**Oligonucleotide target**
Group 1 DNA oligonucleotides	ZIKV Oligo-15	SLA Top loop
	ZIKV Oligo-25	Alternative SLA Top loop
	ZIKV Oligo-44	SLA side stem
	ZIKV Oligo-56	SLA base and UFS
	ZIKV Oligo-80	SLB stem
	ZIKV Oligo-91	SLB top loop
	ZIKV Oligo-96	SLB top loop and UFS
	ZIKV Oligo-123	cHP
	ZIKV Oligo-147	5’ CS
Group 2 DNA oligonucleotides	ZIKV Oligo-163	Pseudoknot 5’ region
	ZIKV Oligo-204	Pseudoknot 3’ region and SLD base
	ZIKV Oligo-229	SLD top loop
	ZIKV Oligo-236	SLD stem
	ZIKV Oligo-242	Alternative SLD stem
Group 3 DNA oligonucleotides	ZIKV connector Oligo 1	Connecting SLA top loop and SLA side stem
	ZIKV connector Oligo 2	Connecting SLA side stem and SLB top loop
	ZIKV connector Oligo 3	Connecting SLB top loop and cHP
	ZIKV connector Oligo 4	Connecting pseudoknot and SLD
	ZIKV connector Oligo 5	Alternative Connecting pseudoknot and SLD

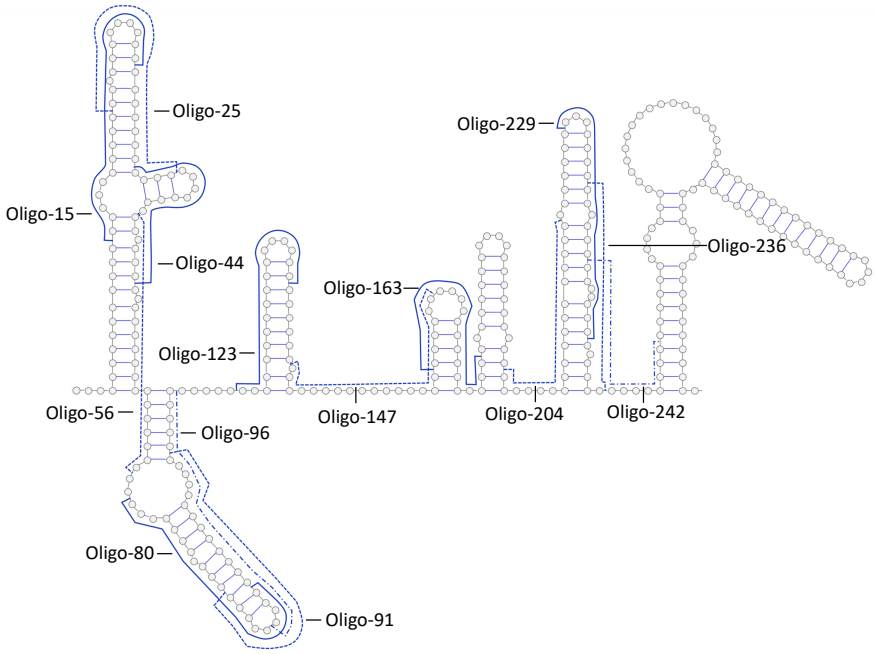


Figure 5.9 Binding positions for antisense oligonucleotides targeting RNA elements within the ZIKV 5 ’ UTR and adjacent coding region

Antisense oligonucleotides were designed to bind and block the function of key RNA elements required for ZIKV replication, therefore impairing ZIKV replication. In order to determine which sites within the ZIKV RNA were available for binding, a wide range target sequences were investigated.

5.4.3 Antisense DNA oligonucleotides inhibit ZIKV Nano-luc genome replication

In order to investigate the potential of ZIKV DNA oligonucleotides to impair ZIKV replication, the DNA oligonucleotides were co-transfected with ZIKV nano-luc replicon and luciferase levels in the cell supernatant was determined at 6, 24 and 48 hours post infection (Fig 5.10, appendix Fig. 10 and 11). In the same experiment, cells were also lysed at 48 hours post infection and luciferase expression was measured. Relative luciferase levels in cell lysate was highly consistent with the luciferase levels in the cell supernatant. DNA oligonucleotides were analysed in three groups. Group 1 oligonucleotides were designed to target RREs which have been previously characterised, such as SLA. Group 2 oligonucleotides were designed to target novel RNA structures which were identified using SHAPE mapping. Group 3 oligonucleotides contain 5’ and 3’ complimentary regions which bind to different RRE regions, linked by a scrambled sequence. For example, the 5’ end of ZIKV connector Oligo 1 binds to the SLA side stem and the 3’ end of the oligonucleotide binds to the top loop of SLA. By targeting multiple sites within the genome, Group 3 oligonucleotides are therefore designed to impair multiple RRE interactions simultaneously.

For the group 1 DNA oligonucleotides, co-transfection of oligonucleotides -15, -25, -56 and -96 led to inhibition of ZIKV genome replication. ZIKV oligo-15 caused the greatest reduction in replication, reducing luciferase expression by around 70% at 6 ($P \leq 0.005$) and 24 ($P \leq 0.001$) hours post infection and by 80% at 48 hours ($P \leq 0.05$). ZIKV oligo-25 was also highly effective, reducing luciferase expression by 60% at 6 hours ($P \leq 0.05$), 65% at 24 hours ($P \leq 0.001$) and

75% at 48 hours (P ≤ 0.05). Transfection of oligo-96 produced the next largest effect, resulting in a 50% reduction in luciferase expression at 6 (P ≤ 0.05) and 24 (P ≤ 0.005) hours and a 70% reduction at 48 hours (P ≤ 0.05). Transfection of oligo-56 resulted in a modest effect on replication, causing a 20% reduction at 6 hours, at 30% reduction at 24 hours (P ≤ 0.05) and a 45% reduction at 48 hours.

For DNA oligonucleotide group 2, ZIKV oligo-204 produced the greatest effect on genome replication, especially at later time points. Oligo-204 co-transfection resulted in a 35% reduction in luciferase expression at 6 hours, a 60% reduction at 24 hours (P ≤ 0.05) and a 70% reduction at 48 hours (P ≤ 0.005).

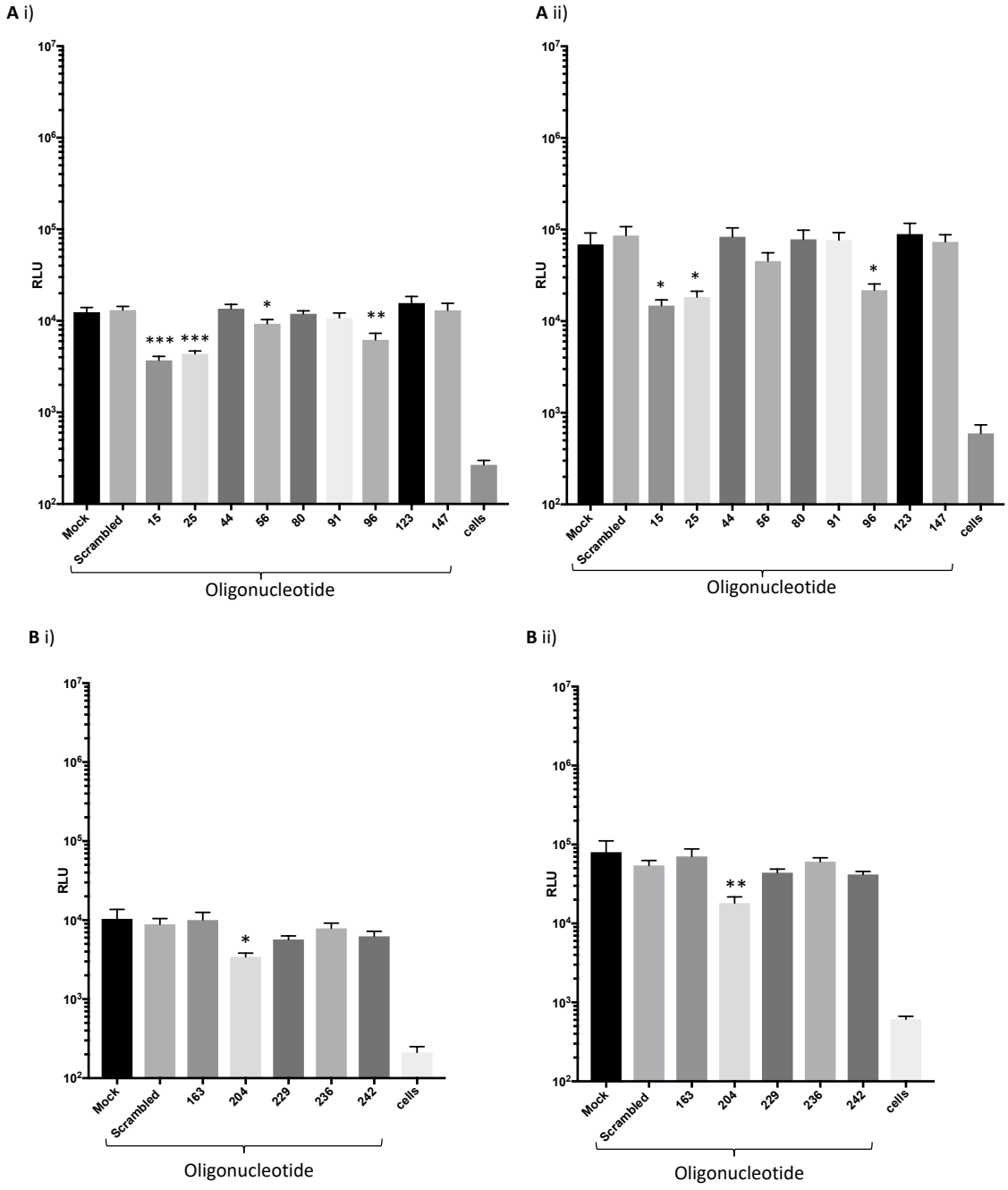


Figure 5.10 ZIKV DNA oligonucleotide screen

DNA oligonucleotides were co-transfected with ZIKV Nano-luc replicon and 100 nM DNA oligonucleotides. Nano-luciferase expression was determined at 24 (i) and 48 (ii) hours post transfection. ***A)*** *DNA oligonucleotides 15-147 (Group 1 oligonucleotides).* ***B)*** *DNA oligonucleotides 163-242 (Group 2 oligonucleotides). Data for 6 hours and lysed cells is shown in appendix figure 10 (Group 1) and 11 (Group 2). * = P ≤ 0.05 ** = P ≤ 0.005 *** = P ≤ 0.001 (Two-tailed T-test vs scrambled LNA) (n = 3). Error bars represent standard error of the mean.*

Group 3 ZIKV connector oligonucleotides did not significantly inhibit the ZIKV N-luc replicon. Since these oligos consisted of only short binding regions, the oligos may not have had sufficient binding affinity to bind to both targets (appendix figure 12 and 13).

5.4.4 Second antisense ZIKV LNA panel

Since a number of antisense DNA oligonucleotides were able to inhibit the ZIKV N-luc replicon, a second set of antisense ZIKV LNAs oligonucleotides (ZLP2) was designed, based on the DNA oligonucleotide sequences (Fig 5.11). Due to issues with self-hybridisation and secondary structure formation within the LNA oligonucleotide (as predicted by Qiagen oligo optimiser), some ZLP2 LNA oligonucleotides were shortened by several bases relative the DNA oligonucleotide counterpart. Like the DNA oligonucleotide counterparts, LNA-15 and -25 target SLA, LNA-56 targets the UFS and base of SLA, LNA-96 targets the UFS and LNA-204 targets the Pseudoknot region and base of SLD.

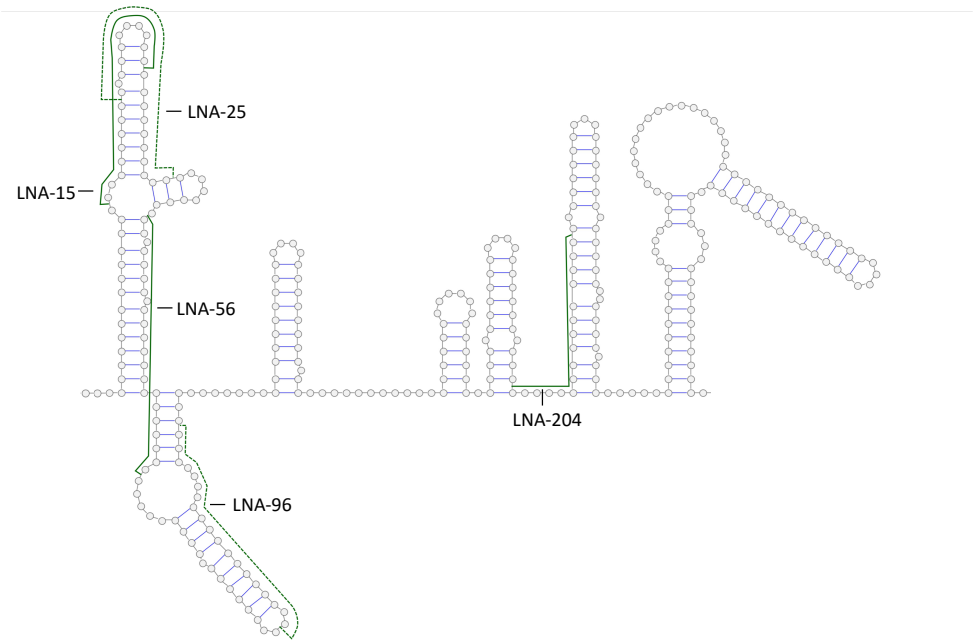


Figure 5.11 Binding sites for ZLP2 oligonucleotides
ZLP2 LNA oligonucleotides were designed based on DNA oligonucleotides, which had been shown to impair ZIKV genome replication.

To assess the effect of ZLP2 LNA oligonucleotides on ZIKV replication, the ZLP2 panel was transfected into ZIKV N-luc IC infected Vero cells (MOI 0.1) and luciferase expression was determined at 24 and 48 hours post infection (Fig 5.12). At 24 hours post infection, ZLP2 LNA oligonucleotides showed a moderate and non-significant reduction in luciferase expression compared with scrambled LNA transfected and mock transfected cells. At 48 hours post infection, LNAs -15, -25, -96 and -204 significantly inhibited ZIKV replication compared with mock transfected and scrambled-LNA transfected cells ($P \leq 0.05$). Cells were also lysed at 48 hours post transfection and luciferase expression was determined. Transfection of LNAs -25, -96 and -204 resulted in significant inhibition of ZIKV replication, reducing luciferase expression by 35-40% compared with cells transfected with scrambled-LNA. These results indicate that LNAs -25, -96 and -204 are able to successfully disrupt their intended RNA-RNA interactions within the ZIKV genome.

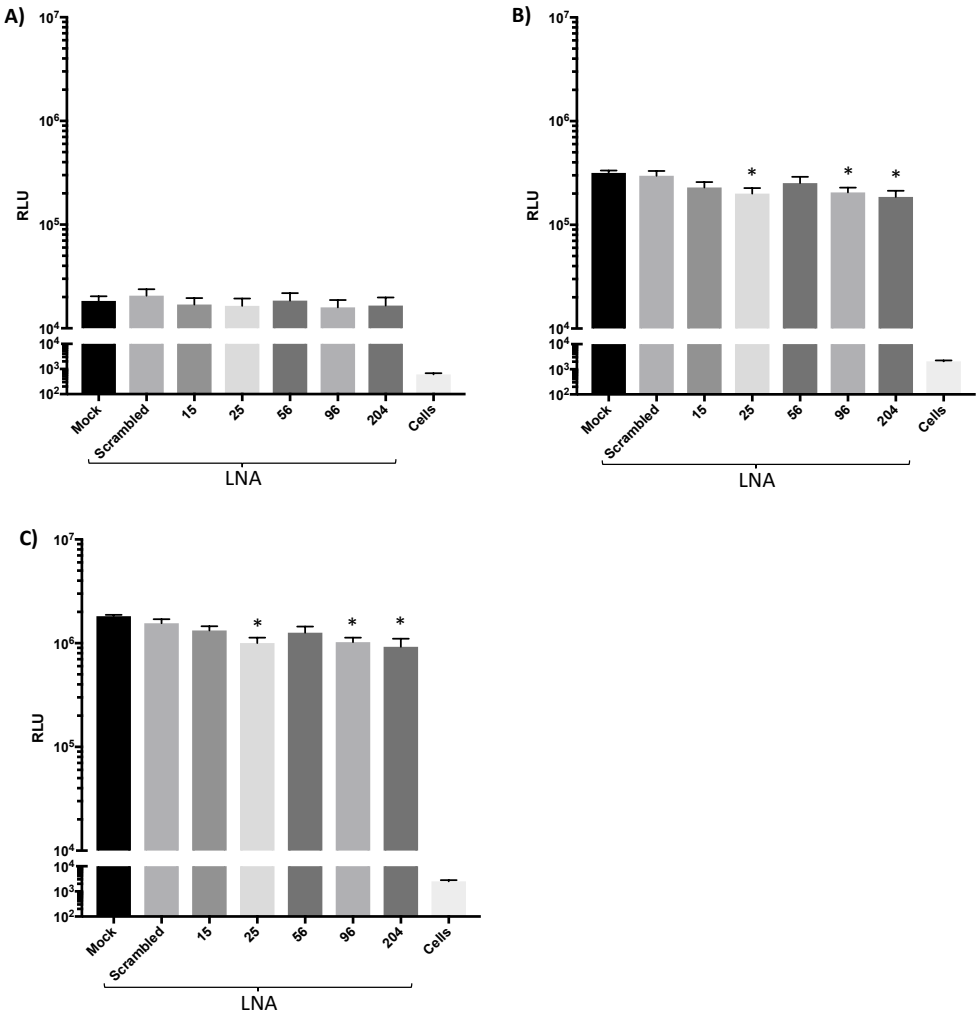


Figure 5.12 Transfection of ZLP2 oligonucleotides into ZIKV infected cells
Vero cells were infected with Nano-luc ZIKV (MOI = 0.1) and incubated for 3 hours, before transfection with 3 µM LNA oligonucleotides. Luciferase expression in cell supernatant was determined at 24 ***(A)*** *and 48* ***(B)*** *hours post infection and in lysed monolayers at 48 hours post infection* ***(C)****. Transfection of LNA-25, -96 and -204 resulted in significant inhibition of luciferase expression at 48 hours post infection, indicating that ZIKV replication has been impaired. * = P ≤ 0.05 (Two-tailed T-test vs scrambled LNA) (n = 3). Error bars represent standard error of the mean.*

5.4.5 ZIKV second LNA panel replicon assay

In order to further investigate ZLP2 oligonucleotides, the oligonucleotides were co-transfected with a ZIKV sub-genomic replicon, which expresses nano-luciferase in the same manner as the ZIKV N-luc IC construct, but lacks the full length ZIKV structural proteins and therefore is not capable of forming infectious virus particles. The LNA concentration used for this experiment was determined based on the CHIKV LNA replicon experiments (Fig 3.11), where 100 nM was the highest concentration used which did not adversely affect transfection efficiency. LNA oligonucleotides were pre-mixed with ZIKV sub-genomic replicon prior to transfection in order to maximise the potential for binding. Similar to the experiments

conducted using the nano-luciferase expressing infectious ZIKV, the LNA oligonucleotides had no significant effect on replication at 6 and 24 hours post transfection (appendix Fig 14). By 48 hours post transfection, LNA-25 transfection resulted in a 55% reduction in luciferase expression ($P \leq 0.05$), indicating inhibition of ZIKV genome replication (Fig 5.13). By contrast, the remining ZLP2 oligonucleotides had no significant effect on ZIKV genome replication.

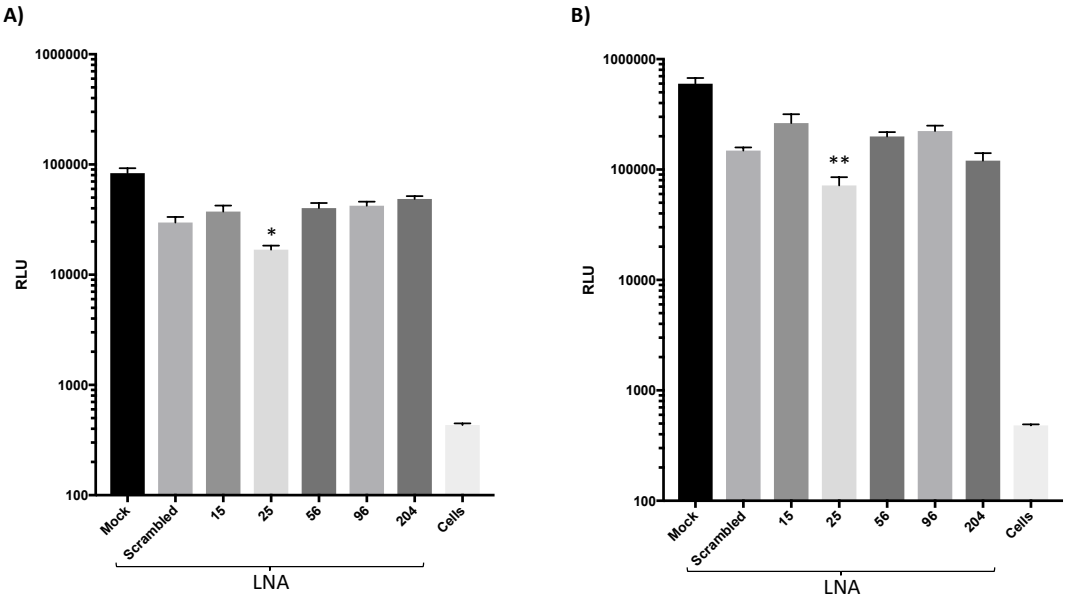


Figure 5.13 The effect of ZLP2 LNA transfection on ZIKV genome replication
*ZIKV nano-luc replicon RNA was co-transfected with 100nM ZLP2 LNA oligonucleotides and luciferase expression was measured at 48 hours post infection **(A)** and in the lysed monolayer at 48 hours post infection **(B)**. Co-transfection of LNA-25 resulted in inhibition of ZIKV genome replication. Data for 6 and 24 hours post infection is presented in appendix figure 14. * = $P \leq 0.05$, ** = $P \leq 0.005$ (Two-tailed T-test vs scrambled LNA) (n = 3). Error bars represent standard error of the mean.*

5.5 Discussion

RREs play essential roles in the lifecycles of many RNA viruses and represent a potential route towards virus specific antiviral therapies. In chapter 3 of this thesis, RNA structures within the 5' region of CHIKV were specifically targeted using antisense LNA oligonucleotides, resulting in inhibition of CHIKV genome replication. Having demonstrated the effectiveness of this approach using CHIKV, the viability this approach in flaviviruses, using ZIKV as a model flavivirus was investigated. Since at the beginning of this project, no RNA structural information for ZIKV had been published, a model of RNA structures within the ZIKV 5' UTR and adjacent capsid coding region was produced using SHAPE mapping.

The target region has previously been shown in other flaviviruses to contain key RREs which are essential for viral replication. For example, this region contains SLA, which is required for RNA synthesis [207] and sequences which are required for genome cyclisation, such as the 5' UAR [211]. It was therefore hypothesised that these RREs were likely to also be present in the ZIKV genome and that this region may also contain novel RREs which are required for ZIKV

replication. SHAPE mapping confirmed the presence of the conserved RNA structures and in addition, several novel RNA structures and a pseudoknot interaction were identified. The RNA model of the pseudoknot region differed between the full-length and fragment models. Since the full-length model include the full ZIKV RNA sequence, this model was considered to be more biologically relevant. Mutations were then introduced into the ZIKV coding sequence, maintaining the amino acid code whilst disrupting RNA folding, in order to determine if these novel RNA elements play a role in the ZIKV lifecycle. For the pseudoknot region, mutations either disrupt base pairing between the pseudoknot regions (full-length model) or disrupts base pairing in a stem region at the base of an RNA structure (fragment model).

Mutations within the pseudoknot region resulted in a reduction in ZIKV titre, indicating that this region plays a role in the ZIKV lifecycle. M-PK ZIKV was undetectable by plaque assay at 24 hours post infection and had a titre one log lower than WT ZIKV at 48 hours post infection. In addition, transfection of ZIKV M-PK RNA in Vero cells resulted in recovery of much less ZIKV compared with WT ZIKV. In addition, M-PK ZIKV produced alterations in plaque morphology, resulting in pinpoint sized plaques. Whilst variation in plaque morphology is present in WT ZIKV, pinpoint sized plaques only occurred in minority populations (around 10%), likely due to the inherent variation of RNA viruses due to the error rate if the viral polymerase NS5. By contrast, M-PK ZIKV formed pinpoint sized plaques exclusively, indicating that replication defects may be present. The size of virus plaques is associated with cytopathic effects, such as apoptosis and lysis, upon infection of the cell monolayer and is a common proxy measurement of viral fitness. Small plaque phenotype may arise as a result of slow virus replication rate or as a consequence of the virus being unable to evade antiviral responses [336]. Since ZIKV mutants were assessed in Vero cells, which are deficient in IFN production [337,338], impaired immune evasion of IFN-stimulated genes is unlikely to explain the pseudoknot mutation phenotype. Therefore, it is more likely that the pseudoknot mutant plaque phenotype is due to inefficient replication, leading to slow plaque growth. In addition, during live attenuated vaccine development, strains with plaque size is a selection criterion for further clinical development. Since reduced plaque size indicates less viral fitness, mutants with a small plaque size phenotype are more suitable vaccine candidates [339].

RNA pseudoknots have known functions in other RNA viruses, for example the non-coding region of HCV contains a conserved pseudoknot that is essential for the initiation of translation [340]. The ZIKV pseudoknot is found around 60 bp into the capsid proteins coding region, therefore may play a role in regulating RNA replication (or translation) by stalling the NS5 polymerase (or ribosome) at an early position on the ZIKV genome. The impact of the ZIKV pseudoknot mutations on translation could be assessed using a polymerase knockout system, whereby the ZIKV N-luc replicon is not capable of undergoing replication. By comparing WT polymerase knockout and M-PK polymerase knockout nano-luciferase signal, the effect of the ZIKV pseudoknot mutations on translation can be determined. The pseudoknot interaction may act as a molecule switch (similar to the UFS region) in order to regulate alternative genome conformations. Alternatively, the pseudoknot region may act as binding site for cellular or viral proteins. This hypothesis could be assessed using pull down experiments, comparing the ability of WT and M-PK ZIKV RNA to bind with proteins within ZIKV infected cell lysates. In addition, the pseudoknot interaction could be re-introduced using an alternative coding sequence, therefore investigating if the pseudoknot region functions in a sequence specific or structure specific manner. The data presented here acts a

basis for further exploration of this region, in order to more fully understand the mechanism by which the pseudoknot may function in the ZIKV lifecycle.

Mutations with SL293 did not produce any measurable phenotype in the infectious virus system. The presence of the inserted mutations at 9 days post transfection indicates that the virus was under no selection pressure to mutate these base pairs. Similarly, the M-SLD mutations were not sufficient to disrupt ZIKV replication, and maintained the inserted mutation after a total of 5 days post transfection, plus a single 48-hour passage. One explanation for the lack of phenotype is that the mutations introduced are not sufficient to disrupt the formation of the RNA structure, despite Mfold predictions of the region indicating that the mutations are sufficient to prevent RNA structure formation. It is also possible that reversion mutations occurred elsewhere in the genome, for example if the stem loops functions by forming long range interactions with another part of the genome (which was not included in the target PCR sequence), then reversions mutations may have occurred in the upstream region. It is also possible that M-SLD and SL293 do not have a function in the cell lines tested and that the mutations may show an effect on replication in other cells lines, for example in cells derived from the brain or placental tissue. M-SLD and SL293 may also simply have no function in the virus replication cycle and the presence of the RNA structure may be merely a non-functional consequence of the nucleic acid sequence being optimised for capsid protein expression.

Based on the reverse genetic analysis and published literature, the potential for RREs within ZIKV to be specifically targeted was investigated. Despite benefitting from LNA oligonucleotide design insights gained when designing the CHIKV LNA panel, the original ZIKV LNA oligonucleotide panel (ZLP1) failed to produce any oligonucleotides capable of inhibiting ZIKV replication. One explanation for these results could be that the RREs are inaccessible in the ZIKV replication complex, therefore the ZLP1 oligonucleotides cannot bind to the target RRE. In addition, although ZLP1 oligonucleotides were designed to minimise self-hybridisation and secondary structure formation, the ZLP1 oligonucleotides may be unable to bind target sequences due to self-interactions. In order to investigate a greater number of potential antisense oligonucleotide binding sites, a ZIKV sub-genomic replicon system was utilised. The ZIKV N-luc sub-genomic replicon contains the ZIKV non-structural proteins, which act to replicate the viral genomic RNA and therefore acts as a model of viral genome replication.

Despite concerns over the half-life of DNA oligonucleotides in cell culture, several antisense DNA oligonucleotides proved to be effective at impairing ZIKV genome replication. DNA oligonucleotide experiments indicate that the top loop region of SLA is available for binding, since both oligo-15 and -25 were able to inhibit ZIKV genome replication. Since SLA has been well characterised as an important factor in flavivirus RNA replication, disruption of the SLA top loop will likely have consequences for the function of this RRE. Since the viral polymerase is known to bind SLA, it is possible that these oligonucleotides act by preventing NS5 binding and activation to SLA, therefore preventing RNA replication for initiating. ZIKV oligo-44, which was designed to target the SLA side stem, did not impair ZIKV genome replication, indicating that the SLA side stem is not accessible for binding. Given that the SLA side stem is a small, tight hairpin, it may be possible that the oligonucleotide is unable to disrupt this structure. Additionally, cellular or virus proteins may be bound to the side stem, which may prevent the oligonucleotide from accessing the target sequence.

ZIKV oligo-96 was also able to successfully inhibit ZIKV genome replication, indicating that the SLB top loop and upstream UFS region are available for binding. ZIKV oligo-56, which targets the downstream region of the UFS was also shown to impair ZIKV genome replication, lending support to the hypothesis that the UFS region is an important factor in RNA replication. The UFS region has been proposed to act as a switch between circular and linear genome conformations [225]. By disrupting the UFS region, oligos-56 and -96 may prevent the switch between the two genome conformations, therefore preventing RNA synthesis from initiating. In addition, since oligo-96 overlaps AUG-1, it may impair RNA replication by preventing initiation of translation. Of the ZIKV oligos targeting novel RNA structures identified by SHAPE mapping, ZIKV oligo-204 was effective at impairing ZIKV genome replicon. Mutagenesis of the pseudoknot region impaired ZIKV replication, indicating that the regions plays a role in the virus lifecycle. By binding to the downstream pseudoknot sequence, oligo-204 may prevent the pseudoknot interaction from occurring, thereby impairing ZIKV genome replication. Since oligo-204 also binds to the base of SLD, it is also possible that oligo-204 functions by preventing formation of SLD, thereby preventing interactions between SLD and cellular/ viral proteins or other regions of the viral genome. ZIKV oligo-163, which binds the upstream pseudoknot region, had no effect on ZIKV genome replication. It is possible that the pseudoknot interaction itself could prevent binding of oligo-163.

RNase H hydrolyses RNA in DNA-RNA hybrids [285]. Since DNA oligonucleotide binding to ZIKV RNA targets results in DNA-RNA hybrids, it is possible that the inhibition seen in this experiment is due to RNase H activity, rather than steric hindrance. Wu et al. showed that antisense oligonucleotides showed greater potency in cells where RNase H1 was overexpressed. In addition, use of siRNAs to silence RNase H1 expression resulted in a reduction in antisense oligonucleotide potency [285]. Similar experiments could be used to determine if the ZIKV DNA oligonucleotides are dependent on RNase H activity. Regardless of the mode of action, the DNA oligonucleotide screening pipeline identified several target sequences which are available for binding within the ZIKV RNA replication complex.

Based on the results of the oligonucleotide experiments, successful oligonucleotide sequences were used as a basis for the design of antisense ZIKV LNA oligonucleotides. LNAs -25, -96 and -204 were effective at inhibiting ZIKV replication, indicating that SLA top loop; SLB top loop/UFS and the upstream pseudoknot region are all available for targeting within the infectious ZIKV lifecycle and that the LNA oligonucleotides are able to access the ZIKV RNA within the genomic RNA replication complex. This demonstrates the viability of targeting RNA structures within the ZIKV genome, providing a basis for the further development of antiviral agents targeting ZIKV RREs.

LNAs -15 and -56 functioned differently to their DNA oligonucleotide counterparts. Since the regions targeted by these LNA oligonucleotides are highly structured, designing functional antisense oligonucleotides proved to be technically challenging. The inclusion of LNA nucleotides into an oligonucleotide sequence was designed to improve the hybridisation efficiency and specificity of the oligonucleotides however, this also has the consequence of increasing the likelihood of both secondary structure formation and self-hybridisation. Formation of secondary structures within the LNA oligonucleotide will decrease the capacity of the LNA oligonucleotide to bind to the intended target, as the secondary structure will need

to unfold in order for the full oligonucleotide to interact with the target sequence. Similarly, LNA oligonucleotides may self-hybridise, causing similar issues. If secondary structure formation or self-hybridisation of the LNA oligonucleotides is more thermodynamically favourable than binding to the target sequence, this would prevent the LNA oligonucleotide from functioning as an antiviral agent. During the design phase, LNA oligonucleotides were designed to prevent both self-hybridisation and secondary structure formation using the Qiagen oligo optimiser software however, as a result of the high levels of secondary structure within the target sequences, the ZIKV LNA oligonucleotides were at the limit of the acceptable parameter scores. Therefore, is it possible that LNA-15 and -56 are unable to bind the intended target sequence due to self-hybridisation and secondary structure formation.

In order to investigate the effect of ZLP2 oligonucleotides on ZIKV sub-genomic replicon, LNA oligonucleotides were co-transfected with ZIKV RNA. The replicon contains the ZIKV replication proteins NS1-5 which form the RNA replication complex, allowing the replicon to function as a model of ZIKV genome replication. ZIKV LNA-25 effectively impaired ZIKV genome replication, indicating that LNA-25 is able to access the ZIKV RNA within the RNA replication complex. The LNA-25 target structure, SLA is known to play an essential role in genomic RNA replication, therefore LNA-25 is likely to function by preventing essential interactions between the SLA top loop and proteins, such as the NS5 polymerase, thereby reducing the efficiency of genome replication. The remaining ZLP2 LNA oligonucleotides were unable to inhibit the ZIKV sub-genomic replicon.

This result may be as a consequence of differences between the infectious ZIKV system and the sub-genomic replicon. For example, LNA-96 and LNA-204 may inhibit ZIKV by preventing long range RNA-RNA interactions between the target sequence and the structural protein coding region (which is absent in the replicon) or by preventing interactions with the ZIKV structural proteins. It is also possible that the LNA-96 and LNA-204 inhibit ZIKV by another mechanism, such as by inhibiting virus genome packaging, which is not present in the sub-genomic replicon system. This highlights the importance of validating findings from replicon experiments in an infectious virus system, as the ZIKV replicon acts only as a model of genome replication and does not take other lifecycle stages into account.

In summary, an RNA element within the ZIKV genome, identified by SHAPE mapping, was shown to play a functional role in the ZIKV lifecycle, further elucidating the complex requirements for ZIKV replication and providing new targets for antiviral agents. Using antisense DNA oligonucleotides, several ZIKV RREs were shown to be susceptible to specific targeting, resulting in inhibition of ZIKV genome replication. DNA oligonucleotides were used as a basis for the design of antisense LNA oligonucleotides and three of these LNA oligonucleotides were shown to inhibit ZIKV replication in an infectious virus system.

Chapter 6 Final Discussion

Functional RNA structures play important roles in the replication cycles of both CHIKV and ZIKV. Whilst the role of RREs in these viruses is beginning to be elucidated at the molecular level, the viability of such structures as targets for antiviral reagents requires further elucidation. Since there are no licensed specific antiviral reagents available to treat either pathogen, the development of novel reagents remains a priority.

6.1 RREs as antiviral targets in CHIKV

In chapter 3, the potential of RREs within the 5′ region of the CHIKV to act as antiviral targets was investigated. Using antisense LNA oligonucleotides, one RRE termed SL165, was identified as a viable target for antiviral agents. Specific targeting of SL165 resulted in inhibition of CHIKV replication in both infectious virus and replicon systems at the level of initiation of virus genome replication. This is a novel finding and demonstrates the viability of using CHIKV RREs in general as antiviral targets.

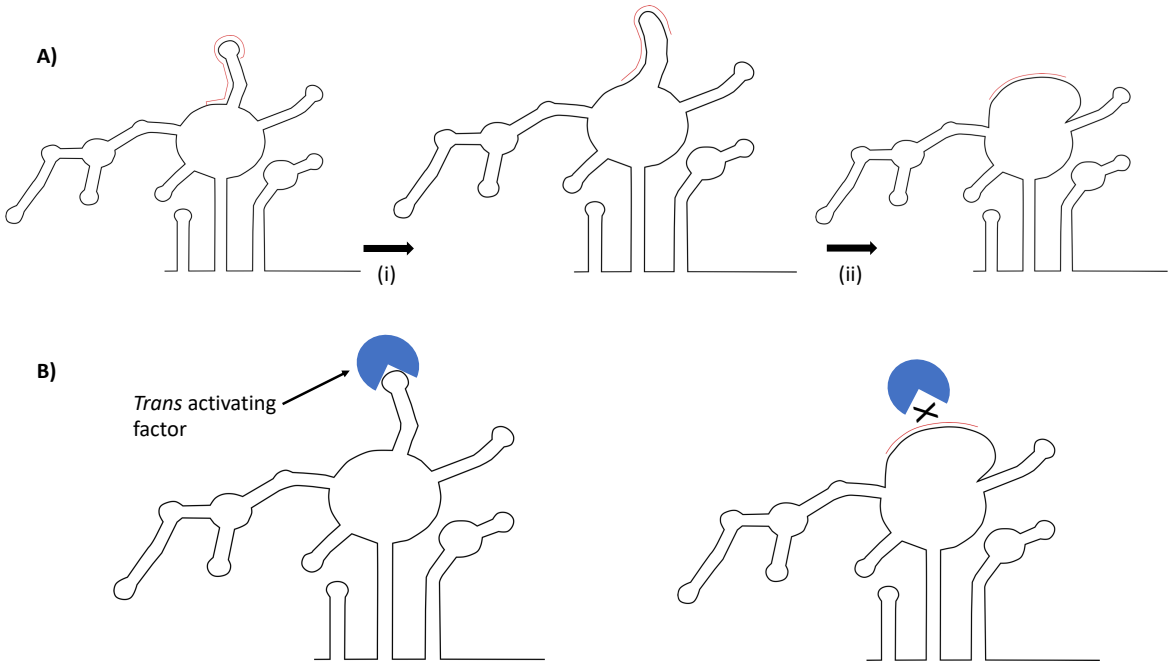


Figure 6.1 Model of LNA-162 binding to SL165
A) *(i) RNA nucleotides base pair with LNA-162, resulting in disruption of SL165 base pairing. (ii) nucleotides involved in SL165 formation are now bound to LNA-162.* ***B)*** *Trans activating factors can no longer interact with SL165, resulting in inhibition of virus replication.*

LNA oligonucleotides have previously been shown to have *in vivo* activity against HCV in clinical trials [299,300], demonstrating that LNA therapies have sufficient pharmacokinetic properties to be used clinically. In order to further validate the efficacy of LNA-162, an animal model can be employed. Newborn and 14-day-old mice show symptoms of CHIKV infection,

such as difficulty walking and lethargy, when infected sub-cutaneous with CHIKV [67]. The efficiency of LNA-162 could be determined by administering the oligonucleotide to CHIKV infected mice, to determine if LNA-162 is able to inhibit CHIKV in *in vivo* conditions. The findings of chapter 3 indicate that SL165 can be specifically targeted, therefore, SL165 can used to identify other antiviral molecules. For example, SL165 could be used as a target for the selection of canonical and non-canonical RNA binding aptamers.

Aptamers are produced through systematic evolution of ligands by exponential enrichment (SELEX). SELEX is an iterative process, whereby a large library of ligands is screened against the target molecule. In the case of CHIKV, SL165 could be selected as the target for SELEX, based on the results presented in chapter 3. The initial library is screened against the target molecule and partially bound or unbound RNAs are discarded. RNAs which bind are amplified using RT-PCR and the process in repeated 8-12 times until only a pool of RNAs with a high affinity for the target remains. The pool of RNAs is then sequenced and the highest affinity/ most effective aptamers identified [341]. Since the SELEX process screens a large number of potential ligands, this approach may lead to selection of aptamers which bind with higher affinity or have a greater effect on virus replication. Aptamers which bind with higher affinity may be more potent antiviral agents, as the aptamer will be less likely to be displaced by protein or RNA-RNA interactions. RNA aptamers may also bind to SL165 in a conformation which causes greater structural disruption, for example by forming a highly condensed structure upon binding, thereby rendering the target nucleotides inaccessible. RNA aptamers may also produce less cellular toxicity then LNA oligonucleotides, allowing higher concentrations to be administered.

Alternatively, SL165 could be used as a target for affimer selection [342]. Affimers (also known as adhirons) are non-antibody scaffold proteins, based on a plant-derived phytocystatin consensus sequence. The consensus sequence produces a small, monomeric, highly stable and highly soluble protein and contains two loops, which act as insertion sites. Amino acids within the insertion sites are replaced by 9 random amino acids, creating a library of potential ligands. Target molecules are then screened using a phage display library, where the Affimers are expressed on the surface of the phage coat protein, allowing the affimers to interact with the target sequence. Phages expressing Affimers with a high affinity for the target are then selected and used for subsequent rounds of selection, before being cloned into protein expression constructs for purification [343]. Similar to RNA aptamer selection, the affimer selection process may produce proteins with improved properties compared with LNA oligonucleotides. Affimers are also highly stable, due to the properties of the phytocystatin scaffold, and have high thermostability, which may result in a long half-life in a cellular environment. In addition, affimers can be labelled with fluorophores in site specific manner and can be used to identify the cellular localisation of target molecules [342,343]. Fluorescent affimers could therefore be used to identify the position of the CHIKV RNA within infected cells, which may elucidate features of the CHIKV lifecycle. For example, if fluorescence is only observed in certain cellular locations, for example in spherules, this would imply that the SL165 is only accessible for binding at specific stages of the virus lifecycle.

In addition, SL165 could be screened against small molecule libraries. For example, the 640-compound library from the French National Chemical Library was used to identify small molecules which bind to microRNA-372 precursor (pre-miR-372). Pre-miR-372 represses

expression of the tumour suppressor protein LATS2 (which is involved in cell cycle regulation), therefore overexpression of pre-miR-372 can result oncogenic effects. Using the small molecule screen, a polyamide compound was shown to decrease levels of pre-miR-372 expression in a human gastric carcinoma cell line, leading to a reduction in cell growth rate [344]. In addition, the small molecule risdiplam has been used in the treatment of spinal muscular atrophy and functions by modifying splicing of the SMN2 messenger RNA [345] (similar to the antisense oligonucleotide Nusinersen as discussed in section 3.1). Small molecule libraries could therefore be used to identify compounds which bind to and disrupt SL165. The advantage of such small molecule screens is that a large number of compounds can be tested, increasing the probability of identifying a compound with anti-CHIKV activity.

LNA-202 targets the CHIKV RRE SL194, an RNA structure known to be essential for replication. Therefore, it was anticipated that LNA-202 would be able to effectively inhibit CHIKV replication. However, despite showing an affinity for CHIKV RNA, LNA-220 was not observed to inhibit CHIKV replication. Similarly, LNA-225, which targets a conserved base paired region, was able to bind CHIKV RNA, but did not inhibit replication. These results suggest that both SL194 and the conserved base pair region are not available for binding within a host cell and that *trans* activating factors, such as viral or host proteins, or higher order RNA-RNA occlude the LNAs from accessing the RNA structures. This indicates that there are gaps in our understanding around the mechanisms by which CHIKV RREs carry out their functions and more research is needed to identify potential interacting partners or alternative RNA folding conformations at different stages of the virus lifecycle. For example, LNAs may bind to the CHIKV RNA at an earlier lifecycle stage but then become displaced by a *trans* activating factor during the genome replication stage.

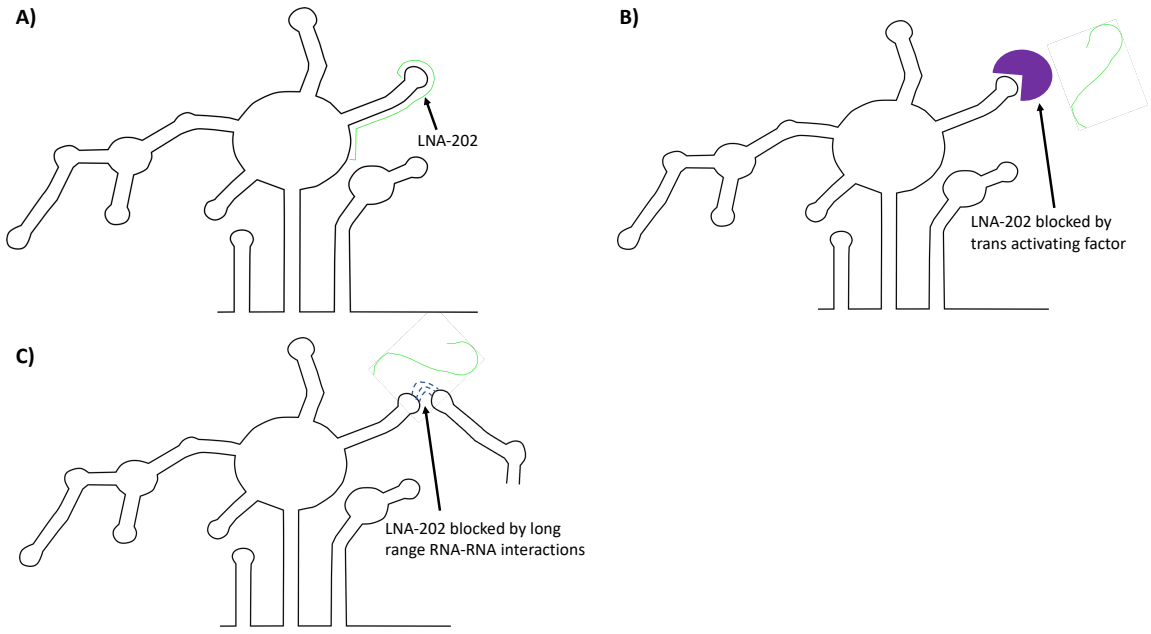


Figure 6.2 Model for LNA-202 interacting with CHIKV RNA

A) *LNA-202 binds to SL194 in EMSA experiments, in the absence of trans activating factors. In a cellular environment, trans activating factors* ***(B)*** *or long-range RNA-RNA interactions* ***(C)*** *may prevent LNA-202 binding to SL194.*

6.2 RNA structures within the 5 'UTR and adjacent coding region of ZIKV

Having demonstrated that RREs within CHIKV are potential targets for antiviral agents, the next aim was to demonstrate the viability for this approach in other viruses with RNA genomes, such as flaviviruses, using ZIKV as a model flavivirus. In order to target specific nucleotide sequences within the ZIKV genome, a single nucleotide resolution structural map is need to identify target RNA elements and the specific nucleotides which can act as the target sequence for antisense oligonucleotides.

Since at the outset of this project, no nucleotide resolution RNA structure map existed in the published literature for ZIKV, the first task was to identify the RNA elements within ZIKV by generating a thermodynamically constrained SHAPE model. The 5' UTR and adjacent capsid coding region were selected as the region of interest, as analysis published literature from other flaviviruses, such as DENV, indicated that RREs essential for virus genome replication were present in this region [182]. Purley thermodynamic models of RNA structures may not be biologically accurate, due to the effect of buffer conditions and the limitations on the sequence length which can be modelled. By investigating the biochemical reactivity of each nucleotide within an RNA molecule (based on nucleotides either being constrained or not constrained), SHAPE mapping allows the folding conformation of an RNA molecule to be

determined biochemically, in a biologically relevant buffer [316]. By constraining thermodynamic algorithms with SHAPE reactivity scores, the resultant model has a higher biological relevance and will therefore act as a better guide for further experiments.

Using *in vitro* SHAPE mapping, high resolution models were generated for ZIKV at both 28 °C and 37 °C, using either short 1-772 RNA or full-length ZIKV RNA. Since ZIKV infects both human and mosquito hosts, which have body temperatures of 37 °C and 28 °C respectively, it is possible that the viral genomic RNA has different folding conformations at these different temperatures. Temperature dependent folding conformations could imply that RREs change conformation in order to optimise interactions with the different host organisms. In addition, flavivirus genomes undergo cyclisation during the initiation of negative strand synthesis, therefore adopting a different folding conformation [182]. In order to investigate the effects of genome cyclisation, SHAPE analysis was performed on both truncated ZIKV RNA molecule (which lack the 3′ elements required for genome cyclisation) and full-length genomic RNA molecules.

In chapter 4, RNA structural models displayed a high level of consistency, indicating that the RNA genome adopts similar conformations at both 28 °C and 37 °C. SHAPE analysis of Full-length ZIKV genomic RNA indicated that the RNA was in a linear conformation, indicating that the linear genomic conformation is the most thermodynamically stable in a biologically relevant buffer *in vitro*, suggesting that additional factors, such as viral or cellular proteins, may be required for genome cyclisation to occur. Minimal differences were observed between the SHAPE maps based on full length template and the short 1-772 template, indicating that experiments conducted using the shortened RNA molecule (e.g. RNA binding assays) were biologically relevant.

During the course of the project, RNA SHAPE models were published in the literature, including an *in vitro* SHAPE model the 5′ region of ZIKV [225] and a full SHAPE map of the entire ZIKV genome, obtained using *in cellulo* SHAPE [212]. The published models are highly consistent with the RNA models presented in this thesis, indicating that the models presented in chapter 4 are of sufficient quality to be used as a basis for further investigation of ZIKV RNA elements. Comparing the published SHAPE models, the region containing the pseudoknot sequence has a different structural conformation. In the *in cellulo* model [212], the region contains a single RNA stem loop. Conversely, in the *in vitro* published model [225], the region contains two short RNA stem loops. These differences may reflect differences between the use of *in cellulo* vs *in vitro* SHAPE mapping. The models presented in chapter 4 show agreement with both of published models for this region. In the truncated ZIKV 1-772 RNA, the conformation is identical to the *in cellulo* published model, whereas the full-length models are identical to the published *in vitro* model. Since NMIA reactivity for this region was inconsistent between the full-length and 1-772 models at both temperatures and did not match well to the models generated by “RNAstructure”, it is difficult to determine which of the two potential conformations (two stem loops or one stem loop) is correct for this region. It is possible that this region forms unstable interactions, for example the pseudoknot interaction may form transiently, meaning that during NMIA treatment of the RNA, the nucleotides involved in the pseudoknot are based pair in some molecules and not in other molecules, resulting in variable reactivity for this region. Further work, including performing

SHAPE reactions on RNA derived from purified replication complexes, will be needed to definitively determine the secondary structure conformation of nucleotides 160-207.

Since the full-length model represents the most biologically relevant of the RNA SHAPE models presented in chapter 4, further experiments were performed using the full-length model as a basis for mutagenesis and antisense oligonucleotide design. In order to further investigate the folding conformation of the ZIKV RNA in mosquito hosts *in cellulo* SHAPE could be performed using mosquito cell lines, although SHAPE maps generated from reactions carried out at 28 °C suggest no changes in RNA secondary structure when compared with 37 °C, indicating that the ZIKV RNA forms the same RNA secondary structures in both mammalian and mosquito hosts.

6.3 RREs as antiviral targets in ZIKV

The SHAPE models produced in chapter 4 revealed that the 5' UTR and adjacent capsid coding region of ZIKV contains several RNA elements, including elements such as SLA, which has been previously shown to function in flavivirus genome replication [207]. In addition, the models showed that this region contained several novel RNA elements, including two RNA stem loops, termed SLD and SLE, and a pseudoknot interaction (based on the 37 °C full length model obtained in chapter 4). Since the RNA models are at single nucleotide resolution, it is possible to use these models as a guide to introduce mutations which disrupt RNA elements.

By comparing the replication efficiency of WT ZIKV to mutants in which these novel RNA elements are disrupted, it is possible to determine if the novel elements function during the ZIKV lifecycle. Mutations which disrupted the pseudoknot region resulted in highly significant impairment of ZIKV replication, resulting in the pseudoknot mutant virus being undetectable by plaque assay at 24 hours post infection and causing a reduction in plaque size. Together these results indicate the discovery of a novel RRE in ZIKV, which is essential for efficient ZIKV replication. This further validates the use of SHAPE mapping to identify novel antiviral targets, as an RNA element identified using SHAPE was shown to be functional.

The mutations introduced in chapter 5 were designed to fully impair the function of the target RRE, meaning that several RNA bases were modified. In order to further investigate the pseudoknot element, the role of individual RNA bases could be determined by generating additional mutant viruses which contain fewer mutations, for example, by only altering bases in either the upstream or downstream pseudoknot regions. This will allow further characterisation of the pseudoknot element, allowing the key nucleotides involved in the RRE to be fully determined. ZIKV mutations were designed such that compensatory mutations, which reform the RREs with an alternative sequence, could be introduced without altering the amino acid coding sequence. To further investigate the ZIKV RREs, these compensatory mutations could be introduced into the ZIKV reverse genetic system, in order to determine if the RREs function in a sequence-specific or structure-specific manner (or a combination of both).

We next investigated the potential of RREs with the ZIKV genome to be specifically targeted by antisense oligonucleotides. Using the RNA models shown in chapter 4, oligonucleotides were designed to bind specific nucleotide sequences, in order to disrupt ZIKV RREs. Several

DNA oligonucleotides were shown to inhibit ZIKV genome replication, using a sub-genomic replicon system. DNA oligonucleotides targeting SLA showed a strong inhibitory effect, highlighting the critical function of SLA in the ZIKV lifecycle and demonstrating that the top loop of SLA is accessible for binding during ZIKV RNA replication. DNA oligonucleotides targeting the UFS region also successfully impaired ZIKV genome replication, supporting the hypothesis that this genome region may be required for switching between linear and circular genome conformations. An oligonucleotide targeting the downstream pseudoknot region and base of SLD also proved to be effective at inhibiting ZIKV genome replication, confirming that RNA elements within this region are required for efficient RNA replication. Since the DNA oligonucleotides effectively inhibited replication in a model of genome replication, this implies that the target RREs function during an early stage of ZIKV replication, such as initiation of genome replication, as opposed to later lifecycle stages, such as genome packaging.

Building on this information, antisense LNA oligonucleotides were designed to target sequences identified in the oligonucleotide screen. Antisense LNA oligonucleotides were shown to impair ZIKV replication, as determined using a nano-luciferase expressing infectious ZIKV. This indicates that antisense oligonucleotides are able to access the genomic RNA within the RNA replication complex, in the context of the infectious ZIKV lifecycle. In additional, a ZIKV LNA oligonucleotides were shown to inhibit replication of a ZIKV replicon, indicating that the oligonucleotide specifically inhibits ZIKV genome replication.

The use of antisense oligonucleotides to specifically target and impair ZIKV RREs further validates the findings of the reverse genetic analysis, as both mutagenesis and specific targeting of the pseudoknot element resulted in inhibition of ZIKV replication. This demonstrates that antisense oligonucleotides are powerful tools for investigating the molecular biology of RNA viruses, allowing regions of the genome to be specifically targeted at the nucleotide level. Use of antisense oligonucleotides provides evidence for the accessibility of different regions of the genome, by demonstrating which regions of the genome are available for targeting within the context of the RNA replication complex. By investigating the effects of antisense oligonucleotides on systems such as sub-genomic replicons, the lifecycle stage at which the target RREs function can be elucidated.

6.4 Concluding statement

In summary, functional RNA elements within positive-sense RNA arboviruses genomes can be specifically targeted, resulting in inhibition of virus genome replication. These findings provide a basis for the development of specific antiviral agents which can be used to treat patients infected with CHIKV or ZIKV, as currently no licenced antiviral agents exist for either virus. Whilst characterisation of RREs within the genomes of arboviruses has typically been focussed on the extreme 5’ and 3’ ends of the genome, there are potentially many RNA structures/ interactions throughout the genome which act as replication elements. Further characterisation of RNA elements throughout the whole genome will likely elucidate more novel interactions and therefore lead to the discovery of novel, specific antiviral agents.

Chapter 7 Appendix

Appendix 1: Buffer compositions

0.5 X T.E buffer
5 mM Tris and 0.5 mM EDTA, pH 7.5

1 X TAE buffer
40 mM Tris 20 mM Acetic acid and 1 mM ethylene-diamine-tetraacetic acid (EDTA)

1 x PBS
137mM NaCl, 2.7mM KCl, 10mM $NAHPO_4$, 1.8mM KH_2PO_4

RNA folding buffer (prepared at 3.3 x concentration)
330 mM HEPES (pH 8.0), 20 mM $MgCl_2$ and 330 mM NaCl

10 X MOPS
0.4 M MOPS (pH 7.0), 0.1 M sodium acetate, 0.01 M EDTA (pH 8.0)

Denaturing sample buffer
60 mM Tris pH 6.8, 25% (v/v) glycerol, 2% (w/v) SDS, 5% β-mercaptoethanol, 0.01% (w/v) bromophenol blue

LB broth
LB Broth; 10 g tryptone, 10 g NaCl, 5 g yeast extract, autoclaved in 1 litre ddH_2O

Native RNA loading dye
30% Glycerol, 0.25% Xylene Cyanol, 0.25% Bromphenol Blue in DEPC H_2O

IP Lysis buffer
25 mM Tris-HCl pH 7.4, 150 mM NaCl, 1 mM EDTA, 1% NP-40 Alternative and 5% glycerol.

SDS running buffer
25 mM Tris, 192 mM glycine, 0.1% (w/v) SDS

Soy broth
10g Tryptic soy broth (Company needed), autoclaved in 1 litre ddH_2O

Towbin buffer
25 mM Tris, 192 mM glycine, 20% (v/v) methanol

Appendix 2: Additional figures

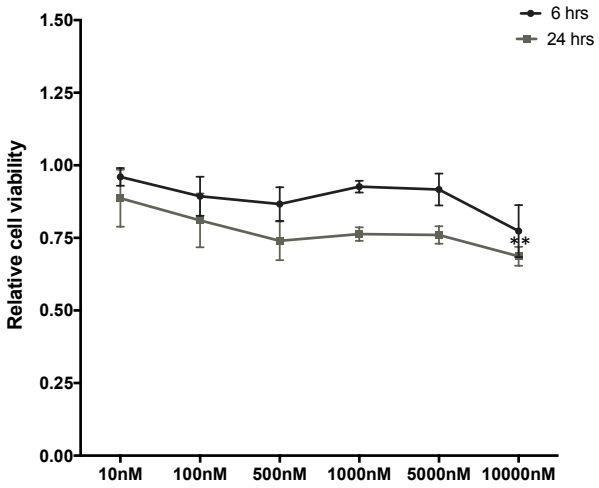


Figure 7.1 Appendix Figure 1

*Huh7 cells were incubated with increasing concentrations of LNA-162 and cells viability was determined by MTT assay after 6 hrs and 24 hrs treatment. At 24 hrs post transfection, transfection of 1000 nM of LNA-162 resulted in a 30% impairment in cell viability, compared with untreated/ lipofectamine only treated cells ** = P ≤ 0.01 (n = 3). Values are relative to lipofectamine only treated Huh7 cells i.e. cells where no LNA was transfected. Lipofectamine was used as a control to negate the effects of lipofectamine transfection on Huh7 cells, which may cause differences in cell viability relative to cells in the absence of lipofectamine. Error bars represent standard error of the mean.*

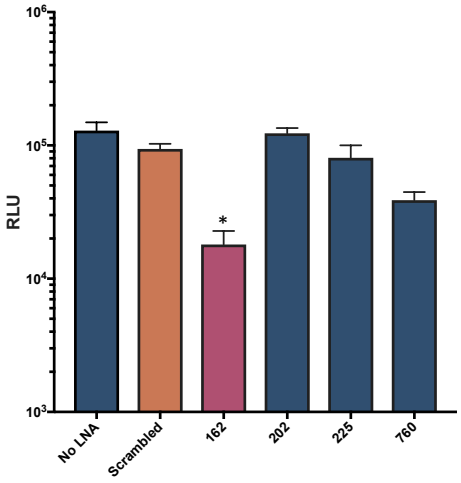


Figure 7.2 Appendix Figure 2

The effect of antisense LNA transfection on Fluc-SGR *replication. LNA oligonucleotides were co-transfection at 100nM with the CHIKV (mono-luc) replicon into Huh7 cells. Cells were harvested at 6 hours post infection. * = P ≤ 0.0001 (Two-tailed T-test vs scrambled LNA) (n = 3). Error bars represent standard error of the mean.*

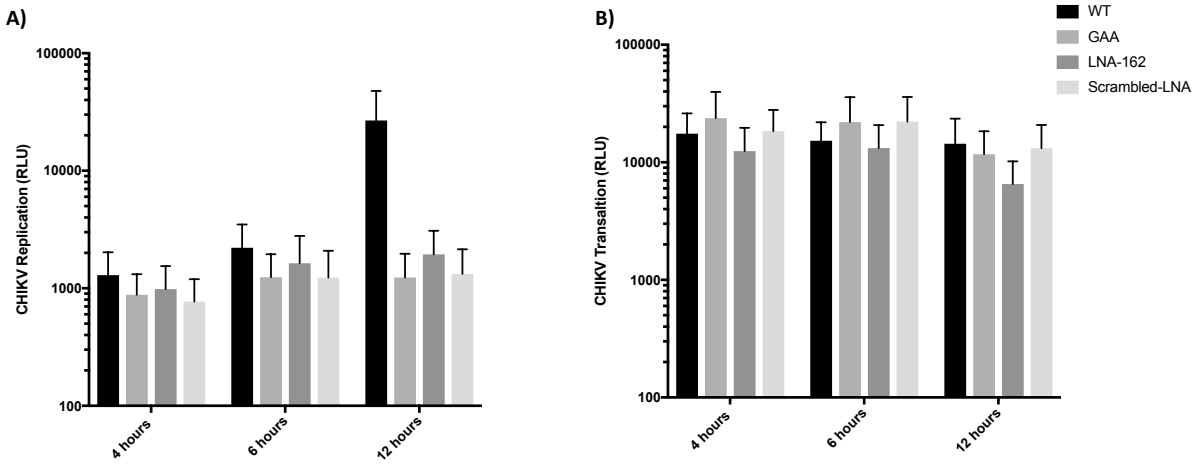


Figure 7.3 Appendix Figure 3

*The effect of LNA transfection on translation were determined using a replication deficient CHIKV replicon system (CHIKV_Rep(GDD>GAA)) in Huh7 cells. Firefly luciferase, representing genome replication (**A**), and Renilla luciferase, representing translation (**B**), were determined at 4, 6 and 12 hours post transfection. Data is shown as normalized to GAA signal for each time point (n = 3). Error bars represent standard error of the mean.*

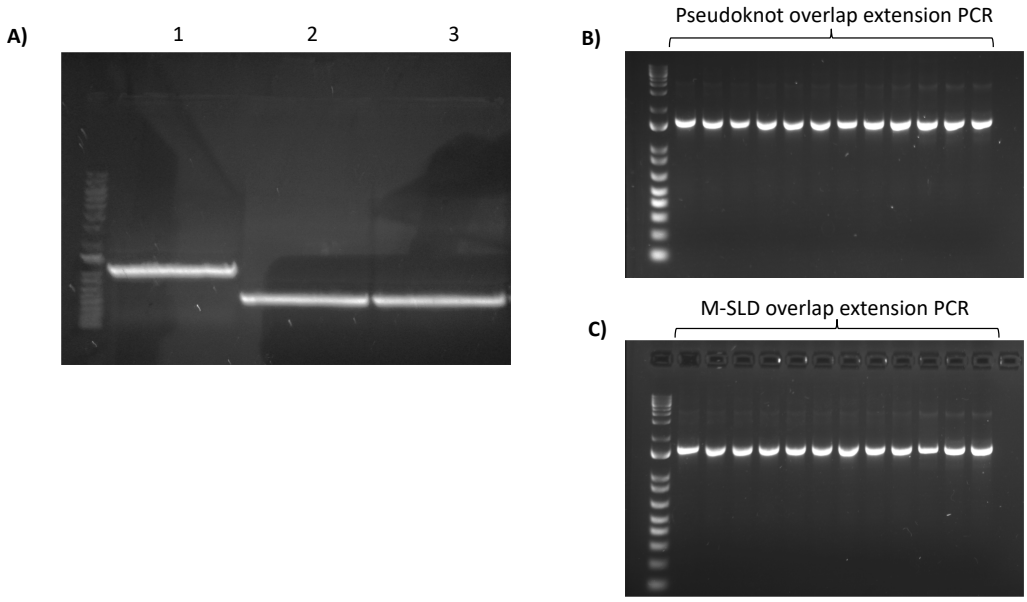


Figure 7.4 Appendix Figure 4

ZIKV IC overlap extension components. ***A)*** *ZIKV Overlap extension stage 1 components. 1) Overlap extension PCR precursor fragment 1 (1153 bp), encompassing the AVRII restriction site. 2-3) Overlap extension PCR precursor fragment 2 (491 bp), encompassing the EcoRI site, Pseudoknot mutations (3) and SLD mutations (4).* ***B-C)*** *Overlap extension PCR reactions. Precursor fragment 1 and 2 were added to the overlap extension reactions PCR reaction in an equimolar amount. PCR reactions were run using a gradient from 50-70 °C, using the pseudoknot precursor fragment 2 (**B**) and M-SLD precursor fragment 2 (**C**). All agarose gels were performed using 2% agarose and using 1Kb plus DNA ladder.*

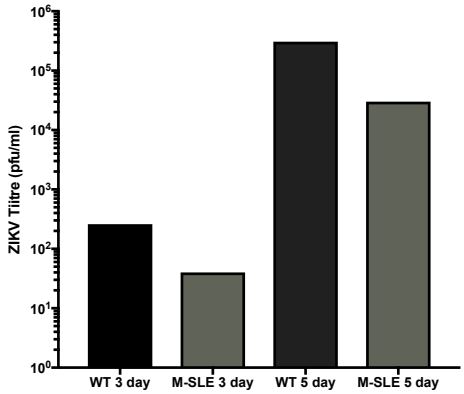


Figure 7.5 Appendix Figure 5

C6/36 cells were infected with WT or M-SLE ZIKV (MOI 1) (Huh 7 derived). Virus was harvested at 3- and 5-days post infection and virus titre was determined by plaque assay. The SLE mutations showed a small reduction in virus titre however, when combined with other data for SLE, SLE mutations was deemed to have no effect on ZIKV replication. (n = 1).

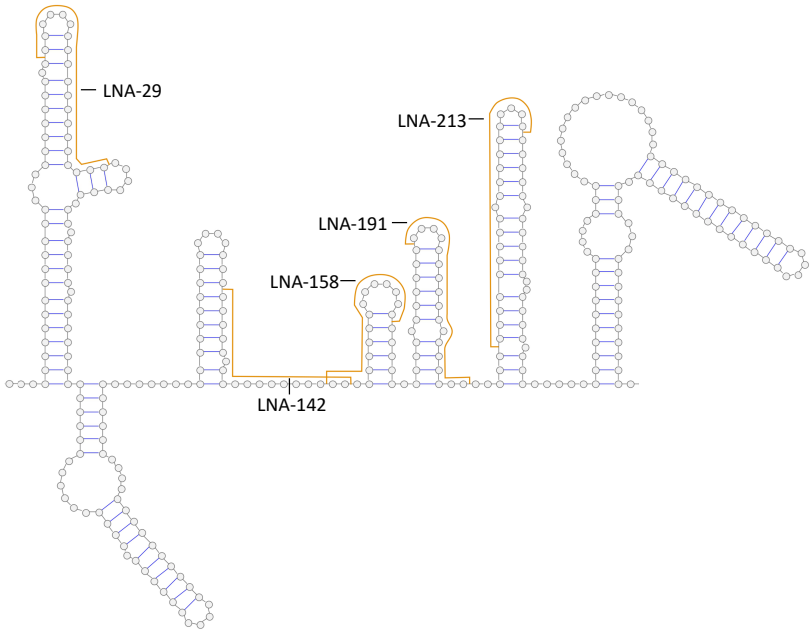


Figure 7.6 Appendix Figure 6

Binding positions of ZLP1 LNA oligonucleotides. ZLP1 oligonucleotides were designed to bind and impair the functional RNA elements within the ZIKV genome.

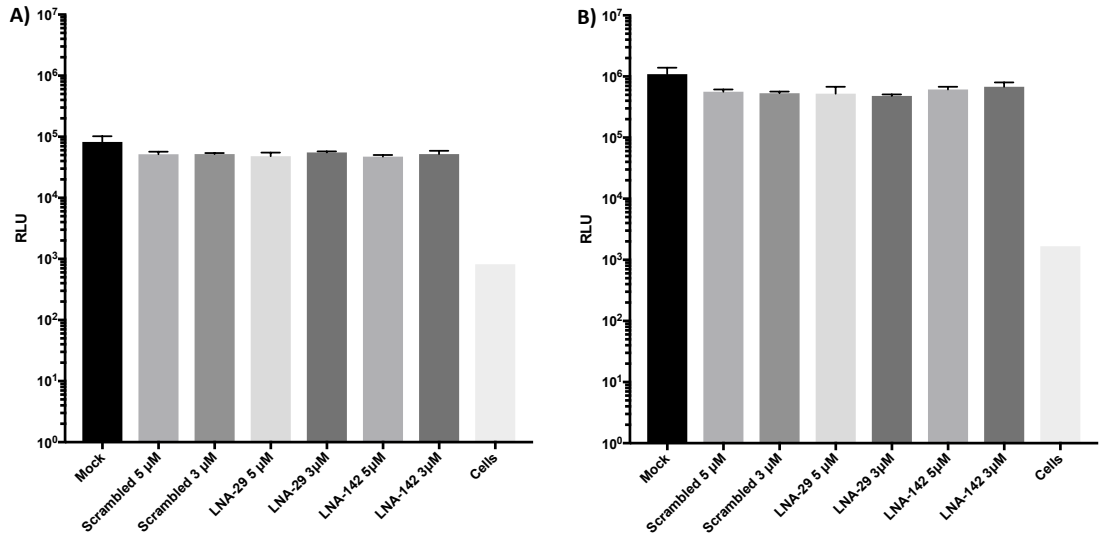


Figure 7.7 Appendix Figure 7

*The effect of ZLP1 LNA oligonucleotides LNA-29 and LNA-142 on ZIKV nano-luc IC replication. Vero cells were infected with ZIKV Nano-luc IC (MOI = 0.1) and incubated for 3 hours, before cells were transfected with either 5 µM or 3 µM ZLP1 LNA oligonucleotides. Luciferase expression was determined at **A)** 24 and **B)** 48 hours post infection. (n = 1) Error bars represent standard error of the mean.*

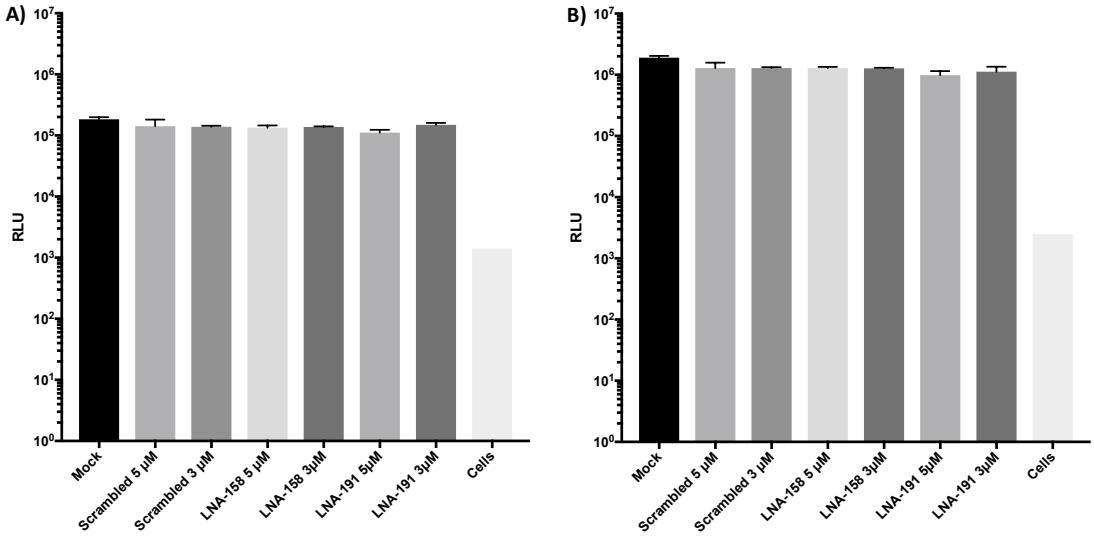


Figure 7.8 Appendix Figure 8

*The effect of ZLP1 LNA oligonucleotides LNA-158 and LNA-191 on ZIKV nano-luc IC replication. Vero cells were infected with ZIKV Nano-luc IC (MOI = 0.1) and incubated for 3 hours, before cells were transfected with either 5 µM or 3 µM ZLP1 panel oligonucleotides. Luciferase expression was determined at **A)** 24 and **B)** 48 hours post infection. (n = 1) Error bars represent standard error of the mean.*

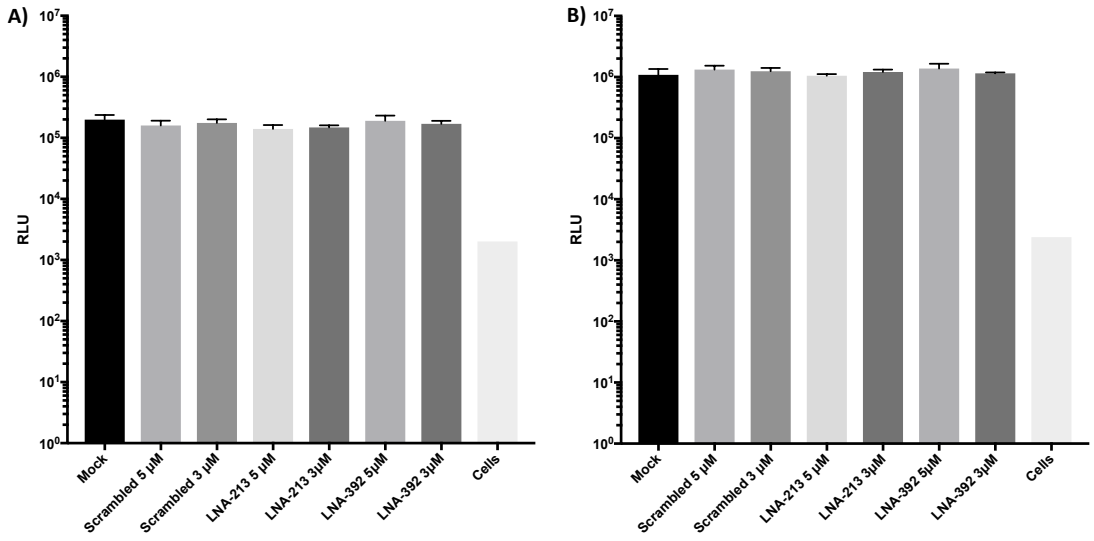


Figure 7.9 Appendix Figure 9

*The effect of ZLP1 LNA oligonucleotides LNA-213 and LNA-392 on ZIKV nano-luc IC replication. Vero cells were infected with ZIKV Nano-luc IC (MOI = 0.1) and incubated for 3 hours, before cells were transfected with either 5 µM or 3 µM ZLP1 LNAs. Luciferase expression was determined at **A)** 24 and **B)** 48 hours post infection. (n = 1) Error bars represent standard error of the mean.*

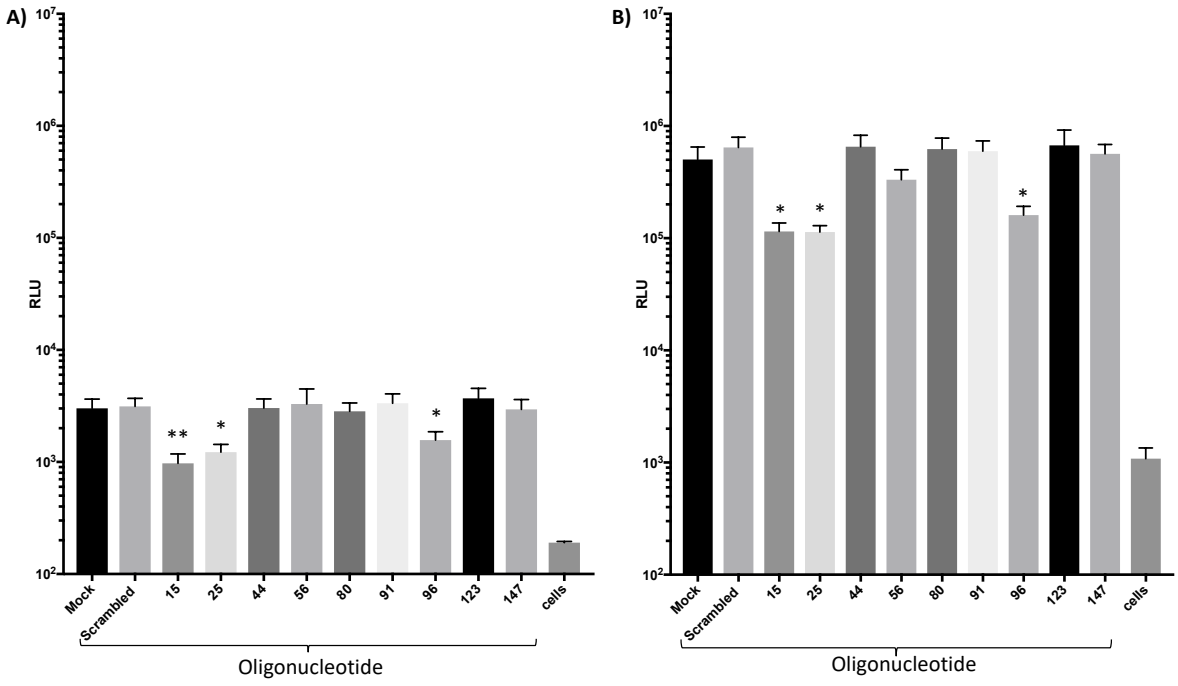


Figure 7.10 Appendix Figure 10

*ZIKV DNA oligonucleotide screen of group 1 oligonucleotides additional data. DNA oligonucleotides were co-transfected with ZIKV Nano-luc replicon and 100 nM DNA oligonucleotides**.** Nano-luciferase expression was determined at **A)** 6 hours post transfection and **B)** at 48 hours post transfection when cells were lysed. * = P ≤ 0.05, ** = P ≤ 0.005 *** = P ≤ 0.001 (Two-tailed T-test vs scrambled LNA) (n = 3). Error bars represent standard error of the mean.*

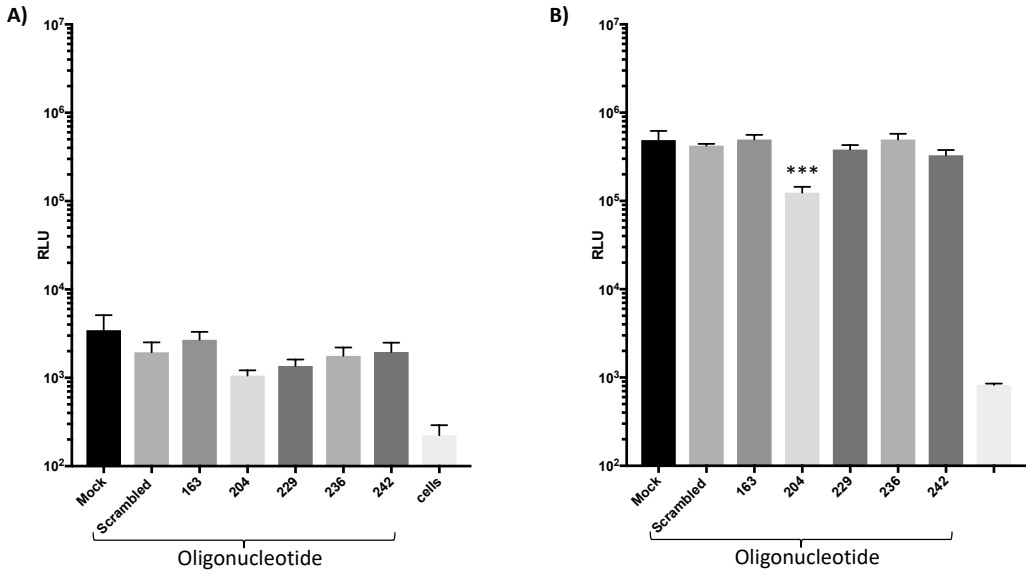


Figure 7.11 Appendix Figure 11

*ZIKV DNA oligonucleotide screen group 2 oligonucleotides additional data. DNA oligonucleotides were co-transfected with ZIKV Nano-luc replicon and 100 nM DNA oligonucleotides. Nano-luciferase expression was determined at **A)** 6 hours post transfection and **B)** at 48 hours post transfection when cells were lysed. * = P ≤ 0.05, ** = P ≤ 0.005 *** = P ≤ 0.001 (Two-tailed T-test vs scrambled LNA) (n = 3). Error bars represent standard error of the mean.*

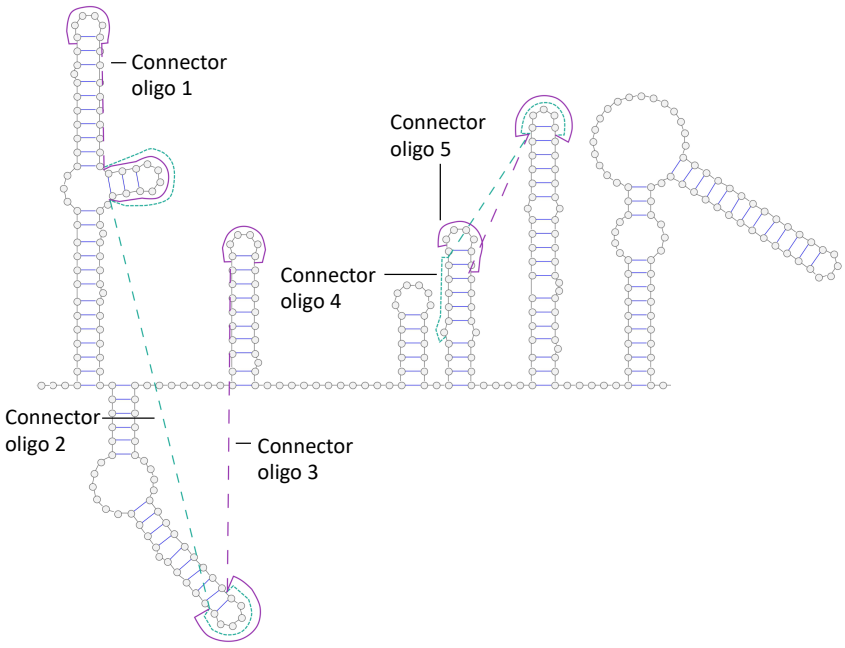


Figure 7.12 Appendix Figure 12

Binding positions of the group 3 DNA oligonucleotides. Group 3 oligonucleotides were designed to target multiple RNA elements simultaneously.

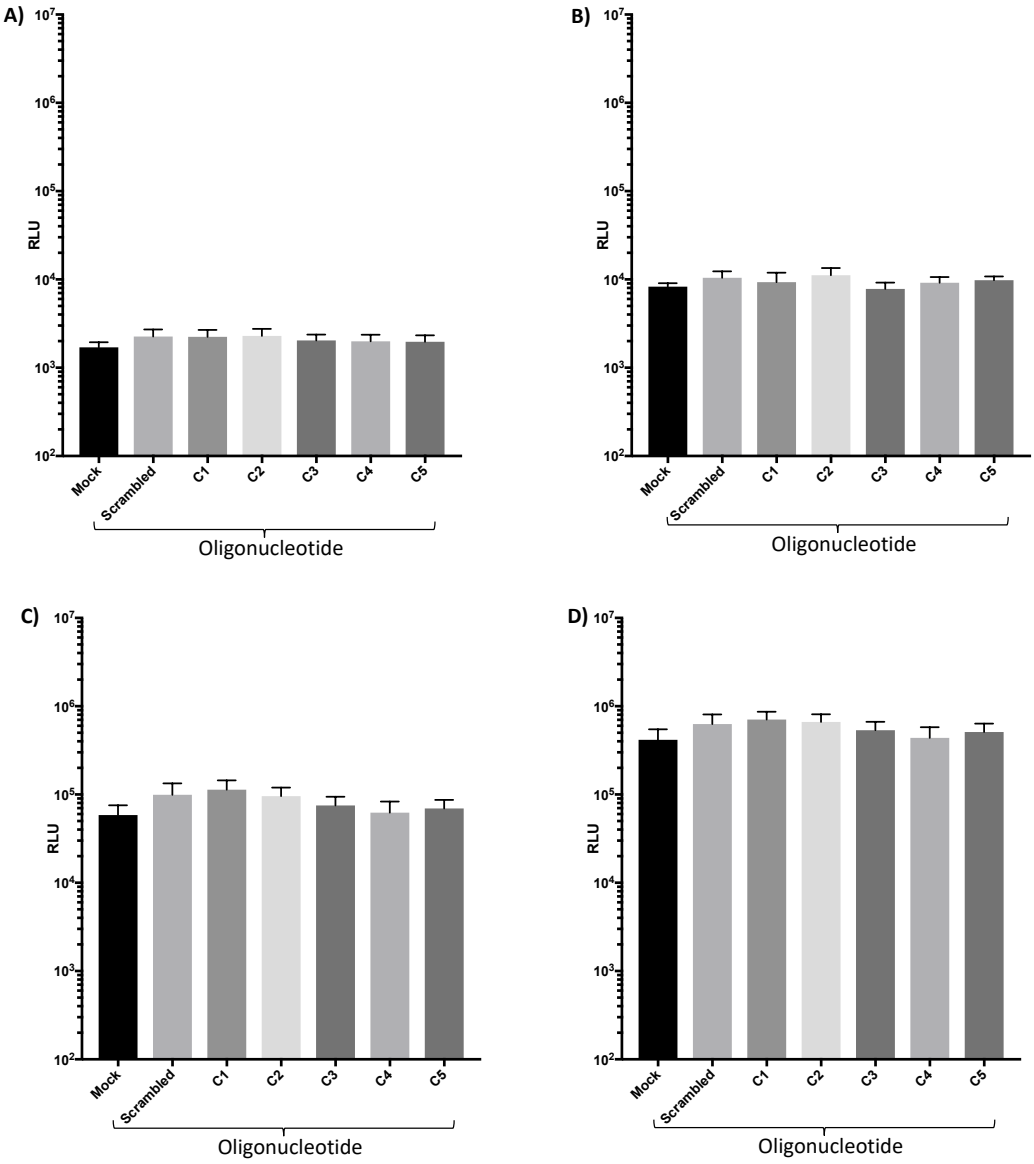


Figure 7.13 Appendix Figure 13

*Co-transfection of ZIKV Nano-luc replicon and connector DNA oligonucleotides, performed in Vero cells. Luciferase activity was determined at 6 (**A**), 24 (**B**) and 48 (**C**) hours post transfection and in lysed cells at 48 hours (**D**). Connector LNA oligonucleotides had no significant effect on ZIKV Nano-luc replicon replication. (n = 3) Error bars represent standard error of the mean.*

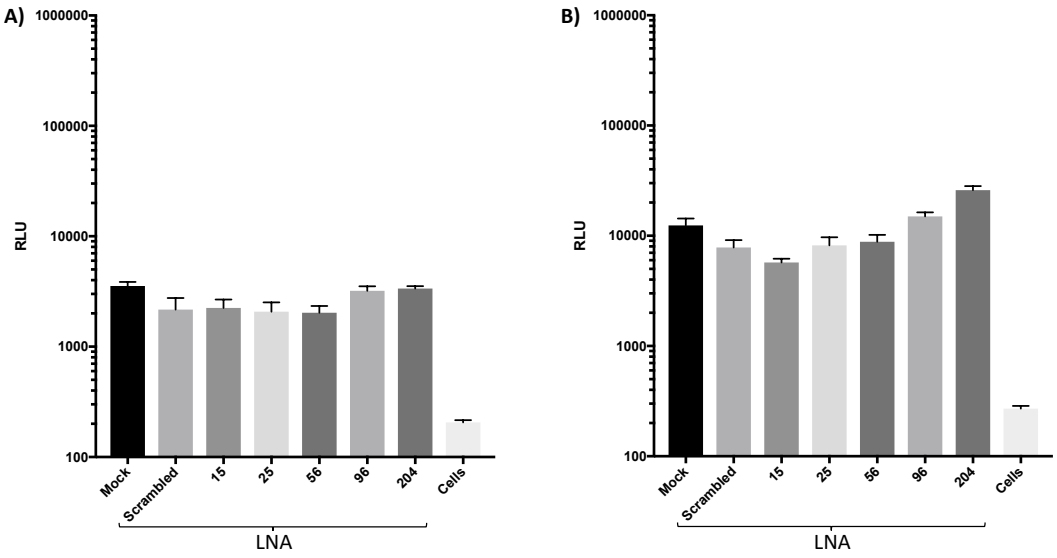


Figure 7.14 Appendix Figure 14

*The effect of ZLP2 LNA transfection on ZIKV genome replication. ZIKV nano-luc replicon RNA was co-transfected with ZLP2 100nM LNAs and luciferase expression was measured at 6 (**A**) and 24 (**B**) hours post infection. (n = 3) Error bars represent standard error of the mean.*

Chapter 8 References

[1] Kraemer MUG, Reiner RC, Brady OJ, Messina JP, Gilbert M, Pigott DM, et al. Past and future spread of the arbovirus vectors Aedes aegypti and Aedes albopictus. Nat Microbiol 2019;4:854–63. https://doi.org/10.1038/s41564-019-0376-y.

[2] 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 Res 2019. https://doi.org/10.1093/nar/gkz640.

[3] Pleij CWA, Rietveld K, Bosch L. A new principle of RNA folding based on pseudoknotting. Nucleic Acids Res 1985;13:1717–31. https://doi.org/10.1093/nar/13.5.1717.

[4] Peselis A, Serganov A. Structure and function of pseudoknots involved in gene expression control. Wiley Interdiscip Rev RNA 2014;5:803–22. https://doi.org/10.1002/wrna.1247.

[5] Brierley I, Pennell S, Gilbert RJC. Viral RNA pseudoknots: Versatile motifs in gene expression and replication. Nat Rev Microbiol 2007;5:598–610. https://doi.org/10.1038/nrmicro1704.

[6] Baranov P V., Henderson CM, Anderson CB, Gesteland RF, Atkins JF, Howard MT. Programmed ribosomal frameshifting in decoding the SARS-CoV genome. Virology 2005;332:498–510. https://doi.org/10.1016/j.virol.2004.11.038.

[7] Alam SL, Wills NM, Ingram JA, Atkins JF, Gesteland RF. Structural studies of the RNA pseudoknot required for readthrough of the gag-termination codon of murine leukemia virus. J Mol Biol 1999;288:837–52. https://doi.org/10.1006/jmbi.1999.2713.

[8] Villordo SM, Gamarnik A V. Genome cyclization as strategy for flavivirus RNA replication. Virus Res 2009;139:230–9. https://doi.org/10.1016/j.virusres.2008.07.016.

[9] Bujalowski PJ, Bujalowski W, Choi KH. Identification of the viral RNA promoter stem loop A (SLA)-binding site on Zika virus polymerase NS5. Sci Rep 2020;10:13306. https://doi.org/10.1038/s41598-020-70094-y.

[10] Dock-Bregeon AC, Chevrier B, Podjarny A, Johnson J, de Bear JS, Gough GR, et al. Crystallographic structure of an RNA helix: [U(UA)6A]2. J Mol Biol 1989;209:459–74. https://doi.org/10.1016/0022-2836(89)90010-7.

[11] Trikha J, Filman DJ, Hogle JM. Crystal structure of a 14 bp RNA duplex with non-symmetrical tandem G·U wobble base pairs. Nucleic Acids Res 1999;27:1728–39. https://doi.org/10.1093/nar/27.7.1728.

[12] Miyazaki Y, Irobalieva RN, Tolbert BS, Smalls-Mantey A, Iyalla K, Loeliger K, et al. Structure of a Conserved Retroviral RNA Packaging Element by NMR Spectroscopy and Cryo-Electron Tomography. J Mol Biol 2010;404:751–72. https://doi.org/10.1016/j.jmb.2010.09.009.

[13] Lee MK, Gal M, Frydman L, Varani G. Real-time multidimensional NMR follows RNA folding with second resolution. Proc Natl Acad Sci U S A 2010;107:9192–7. https://doi.org/10.1073/pnas.1001195107.

[14] Chen R, Mukhopadhyay S, Merits A, Bolling B, Nasar F, Coffey LL, et al. ICTV Virus Taxonomy Profile: Togaviridae ICTV VIRUS TAXONOMY PROFILES. J Gen Virol 2018;99:761–2. https://doi.org/10.1099/jgv.0.001072.

[15] 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). Arch Virol 2019;164:2417–29. https://doi.org/10.1007/s00705-019-04306-w.
[16] Perelygina L, Chen M hsin, Suppiah S, Adebayo A, Abernathy E, Dorsey M, et al. Infectious vaccine-derived rubella viruses emerge, persist, and evolve in cutaneous granulomas of children with primary immunodeficiencies. PLoS Pathog 2019;15:e1008080. https://doi.org/10.1371/journal.ppat.1008080.
[17] Kim DY, Reynaud JM, Rasalouskaya 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 Pathog 2016;12:e1005810. https://doi.org/10.1371/journal.ppat.1005810.
[18] Nasar F, Palacios G, Gorchakov R V., Guzman H, Travassos Da Rosa AP, Savji N, et al. Eilat virus, a unique alphavirus with host range restricted to insects by RNA replication. Proc Natl Acad Sci U S A 2012;109:14622–7. https://doi.org/10.1073/pnas.1204787109.
[19] Hermanns K, Zirkel F, Kopp A, Marklewitz M, Rwego IB, Estrada A, et al. Discovery of a novel alphavirus related to Eilat virus. J Gen Virol 2017;98:43–9. https://doi.org/10.1099/jgv.0.000694.
[20] Powers AM, Brault AC, Shirako Y, Strauss EG, Kang W, Strauss JH, et al. Evolutionary Relationships and Systematics of the Alphaviruses. J Virol 2001;75:10118–31. https://doi.org/10.1128/jvi.75.21.10118-10131.2001.
[21] Robinson MC. An epidemic of virus disease in Southern Province, Tanganyika territory, in 1952–1953. Trans R Soc Trop Med Hyg 1955;49:28–32. https://doi.org/10.1016/0035-9203(55)90080-8.
[22] Lumsden WHR. An epidemic of virus disease in Southern Province, Tanganyika territory, in 1952–1953 II. General description and epidemiology. Trans R Soc Trop Med Hyg 1955;49:33–57. https://doi.org/10.1016/0035-9203(55)90081-X.
[23] Ross RW. The newala epidemic: III. The virus: Isolation, pathogenic properties and relationship to the epidemic. J Hyg (Lond) 1956;54:177–91. https://doi.org/10.1017/S0022172400044442.
[24] Volk SM, Chen R, Tsetsarkin KA, Adams AP, Garcia TI, Sall AA, et al. Genome-Scale Phylogenetic Analyses of Chikungunya Virus Reveal Independent Emergences of Recent Epidemics and Various Evolutionary Rates. J Virol 2010;84:6497–504. https://doi.org/10.1128/jvi.01603-09.
[25] Powers AM, Brault AC, Tesh RB, Weaver SC. Re-emergence of chikungunya and o'nyong-nyong viruses: Evidence for distinct geographical lineages and distant evolutionary relationships. J Gen Virol 2000;81:471–9. https://doi.org/10.1099/0022-1317-81-2-471.
[26] Silva LA, Dermody TS. Chikungunya virus: epidemiology, replication, disease mechanisms, and prospective intervention strategies. J Clin Invest 2017;127:737–49. https://doi.org/10.1172/JCI84417.
[27] Humphrey JM, Cleton NB, Reusken CBEM, Glesby MJ, Koopmans MPG, Abu-Raddad LJ. Urban Chikungunya in the Middle East and North Africa: A systematic review. PLoS Negl Trop Dis 2017;11:e0005707. https://doi.org/10.1371/journal.pntd.0005707.
[28] Fahmy NT, Klena JD, Mohamed AS, Zayed A, Villinski JT. Complete genome sequence of chikungunya virus isolated from an Aedes aegypti mosquito during an outbreak in Yemen, 2011. Genome Announc 2015;3. https://doi.org/10.1128/genomeA.00789-15.

[29] Yactayo S, Staples JE, Millot V, Cibrelus L, Ramon-Pardo P. Epidemiology of Chikungunya in the Americas. J Infect Dis 2016;214:S441–5. https://doi.org/10.1093/infdis/jiw390.
[30] Mavalankar D, Shastri P, Raman P. Chikungunya epidemic in India: a major public-health disaster. Lancet Infect Dis 2007;7:306–7. https://doi.org/10.1016/S1473-3099(07)70091-9.
[31] Wimalasiri-Yapa BMCR, Stassen L, Huang X, Hafner LM, Hu W, Devine GJ, et al. Chikungunya virus in Asia–Pacific: a systematic review. Emerg Microbes Infect 2019;8:70–9. https://doi.org/10.1080/22221751.2018.1559708.
[32] Zeller H, Van Bortel W, Sudre B. Chikungunya: Its History in Africa and Asia and Its Spread to New Regions in 2013–2014. J Infect Dis 2016;214:S436–40. https://doi.org/10.1093/infdis/jiw391.
[33] 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. Am J Trop Med Hyg 2008;78:333–7. https://doi.org/10.4269/ajtmh.2008.78.333.
[34] Pialoux G, Gaüzère BA, Jauréguiberry S, Strobel M. Chikungunya, an epidemic arbovirosis. Lancet Infect Dis 2007;7:319–27. https://doi.org/10.1016/S1473-3099(07)70107-X.
[35] Renault P, Solet JL, Sissoko D, Balleydier E, Larrieu S, Filleul L, et al. A major epidemic of chikungunya virus infection on Réunion Island, France, 2005-2006. Am J Trop Med Hyg 2007;77:727–31. https://doi.org/10.4269/ajtmh.2007.77.727.
[36] Kraemer MU, Sinka ME, Duda KA, Mylne AQ, Shearer FM, Barker CM, et al. The global distribution of the arbovirus vectors Aedes aegypti and Ae. albopictus. Elife 2015;4. https://doi.org/10.7554/eLife.08347.
[37] An W, Ge N, Cao Y, Sun J, Jin X. Recent progress on chikungunya virus research. Virol Sin 2017;32:441–53. https://doi.org/10.1007/s12250-017-4072-x.
[38] Nunes MRT, Faria NR, de Vasconcelos JM, Golding N, Kraemer MU, de Oliveira LF, et al. Emergence and potential for spread of Chikungunya virus in Brazil. BMC Med 2015;13:102. https://doi.org/10.1186/s12916-015-0348-x.
[39] Gavotto A, Muanza B, Delion F, Dusacre JA, Amedro P. Chikungunya disease among infants in French West Indies during the 2014 outbreak. Arch Pediatr 2019;26:259–62. https://doi.org/10.1016/j.arcped.2019.05.014.
[40] Dorléans F, Hoen B, Najioullah F, Herrmann-Storck C, Schepers KM, Abel S, et al. Outbreak of chikungunya in the French caribbean islands of martinique and guadeloupe: Findings from a hospital-Based surveillance system (2013–2015). Am J Trop Med Hyg 2018;98:1819–25. https://doi.org/10.4269/ajtmh.16-0719.
[41] Sharp TM, Ryff KR, Alvarado L, Shieh W-J, Zaki SR, Margolis HS, et al. Surveillance for Chikungunya and Dengue During the First Year of Chikungunya Virus Circulation in Puerto Rico. J Infect Dis 2016;214:S475–81. https://doi.org/10.1093/infdis/jiw245.
[42] Madariaga M, Ticona E, Resurrecion C. Chikungunya: Bending over the Americas and the rest of the world. Brazilian J Infect Dis 2016;20:91–8. https://doi.org/10.1016/j.bjid.2015.10.004.
[43] Diallo D, Sall AA, Buenemann M, Chen R, Faye O, Diagne CT, et al. Landscape Ecology of Sylvatic Chikungunya Virus and Mosquito Vectors in Southeastern Senegal. PLoS Negl Trop Dis 2012;6:e1649. https://doi.org/10.1371/journal.pntd.0001649.
[44] Althouse BM, Guerbois M, Cummings DAT, Diop OM, Faye O, Faye A, et al. Role of monkeys in the sylvatic cycle of chikungunya virus in Senegal. Nat Commun 2018;9:1–

10. https://doi.org/10.1038/s41467-018-03332-7.
[45] Valentine MJ, Murdock CC, Kelly PJ. Sylvatic cycles of arboviruses in non-human primates. Parasites and Vectors 2019;12:463. https://doi.org/10.1186/s13071-019-3732-0.
[46] McCrae AWR, Henderson BE, Kirya BG, Sempala SDK. Chikungunya virus in the entebbe area of Uganda: Isolations and epidemiology. Trans R Soc Trop Med Hyg 1971;65:152–68. https://doi.org/10.1016/0035-9203(71)90212-4.
[47] Kading RC, Borland EM, Cranfield M, Powers AM. Prevalence of antibodies to alphaviruses and flaviviruses in free-ranging game animals and nonhuman primates in the greater Congo basin. J Wildl Dis 2013;49:587–99. https://doi.org/10.7589/2012-08-212.
[48] Eastwood G, Sang RC, Guerbois M, Taracha ELN, Weaver SC. Enzootic circulation of chikungunya virus in East Africa: Serological evidence in non-human Kenyan primates. Am J Trop Med Hyg 2017;97:1399–404. https://doi.org/10.4269/ajtmh.17-0126.
[49] Diallo M, Thonnon J, Traore-Lamizana M, Fontenille D. Vectors of Chikungunya virus in Senegal: Current data and transmission cycles. Am J Trop Med Hyg 1999;60:281–6. https://doi.org/10.4269/ajtmh.1999.60.281.
[50] Weaver SC. Evolutionary influences in arboviral disease. Curr Top Microbiol Immunol 2006;299:285–314. https://doi.org/10.1007/3-540-26397-7_10.
[51] Das T, Jaffar-Bandjee MC, Hoarau JJ, Krejbich Trotot P, Denizot M, Lee-Pat-Yuen G, et al. Chikungunya fever: CNS infection and pathologies of a re-emerging arbovirus. Prog Neurobiol 2010;91:121–9. https://doi.org/10.1016/j.pneurobio.2009.12.006.
[52] Lwande OW, Obanda V, Lindström A, Ahlm C, Evander M, Näslund J, et al. Globe-Trotting Aedes aegypti and Aedes albopictus: Risk Factors for Arbovirus Pandemics. Vector-Borne Zoonotic Dis 2020;20:71–81. https://doi.org/10.1089/vbz.2019.2486.
[53] Sam IC, Chua CL, Rovie-Ryan JJ, Fu JYL, Tong C, Sitam FT, et al. Chikungunya virus in macaques, Malaysia. Emerg Infect Dis 2015;21:1683–5. https://doi.org/10.3201/eid2109.150439.
[54] Apandi Y, Nazni WA, Azleen ZAN, Vythilingam I, Noorazian MY, Azahari AH, et al. The first isolation of chikungunya virus from non-human primates in Malaysia. J Gen Mol Virol 2009;1:35–039.
[55] Mavale M, Parashar D, Sudeep A, Gokhale M, Ghodke Y, Geevarghese G, et al. Venereal transmission of chikungunya virus by Aedes aegypti mosquitoes (Diptera: Culicidae). Am J Trop Med Hyg 2010;83:1242–4. https://doi.org/10.4269/ajtmh.2010.09-0577.
[56] Ganesan VK, Duan B, Reid SP. Chikungunya Virus: Pathophysiology, Mechanism, and Modeling. Viruses 2017;9. https://doi.org/10.3390/v9120368.
[57] Her Z, Malleret B, Chan M, Ong EKS, Wong S-C, Kwek DJC, et al. Active Infection of Human Blood Monocytes by Chikungunya Virus Triggers an Innate Immune Response. J Immunol 2010;184:5903–13. https://doi.org/10.4049/jimmunol.0904181.
[58] Dupuis-Maguiraga L, Noret M, Brun S, Le Grand R, Gras G, Roques P. Chikungunya disease: Infection-associated markers from the acute to the chronic phase of arbovirus-induced arthralgia. PLoS Negl Trop Dis 2012;6. https://doi.org/10.1371/journal.pntd.0001446.
[59] Manimunda SP, Sugunan AP, Rai SK, Vijayachari P, Shriram AN, Sharma S, et al. Short report: Outbreak of chikungunya fever, Dakshina Kannada District, South India, 2008. Am J Trop Med Hyg 2010;83:751–4. https://doi.org/10.4269/ajtmh.2010.09-0433.

[60] Gérardin P, Guernier V, Perrau J, Fianu A, Le Roux K, Grivard P, et al. Estimating Chikungunya prevalence in La Réunion Island outbreak by serosurveys: Two methods for two critical times of the epidemic. BMC Infect Dis 2008;8:99. https://doi.org/10.1186/1471-2334-8-99.
[61] Ng KW, Chow A, Win MK, Dimatatac F, Neo HY, Lye DC, et al. Clinical features and epidemiology of chikungunya infection in Singapore. Orig Artic Singapore Med J 2009;50:785–90.
[62] Lum FM, Ng LFP. Cellular and molecular mechanisms of chikungunya pathogenesis. Antiviral Res 2015;120:165–74. https://doi.org/10.1016/j.antiviral.2015.06.009.
[63] 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. Clin Infect Dis 2007;44:1401–7. https://doi.org/10.1086/517537.
[64] Economopoulou A, Dominguez M, Helynck B, Sissoko D, Wichmann O, Quenel P, et al. Atypical Chikungunya virus infections: Clinical manifestations, mortality and risk factors for severe disease during the 2005-2006 outbreak on Réunion. Epidemiol Infect 2009;137:534–41. https://doi.org/10.1017/S0950268808001167.
[65] Chandak NH, Kashyap RS, Kabra D, Karandikar P, Saha SS, Morey SH, et al. Neurological complications of Chikungunya virus infection. Neurol India 2009;57:177–80. https://doi.org/10.4103/0028-3886.51289.
[66] Krejbich-Trotot P, Denizot M, Hoarau J, Jaffar-Bandjee M, Das T, Gasque P. Chikungunya virus mobilizes the apoptotic machinery to invade host cell defenses. FASEB J 2011;25:314–25. https://doi.org/10.1096/fj.10-164178.
[67] 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 Pathog 2008;4:e29. https://doi.org/10.1371/journal.ppat.0040029.
[68] Gérardin P, Fianu A, Michault A, Mussard C, Boussaïd K, Rollot O, et al. Predictors of Chikungunya rheumatism: A prognostic survey ancillary to the TELECHIK cohort study. Arthritis Res Ther 2013;15:R9. https://doi.org/10.1186/ar4137.
[69] Schilte C, Staikovsky F, Couderc T, Madec Y, Carpentier F, Kassab S, et al. Chikungunya Virus-associated Long-term Arthralgia: A 36-month Prospective Longitudinal Study. PLoS Negl Trop Dis 2013;7:e2137. https://doi.org/10.1371/journal.pntd.0002137.
[70] Couderc T, Lecuit M. Chikungunya virus pathogenesis: From bedside to bench. Antiviral Res 2015;121:120–31. https://doi.org/10.1016/j.antiviral.2015.07.002.
[71] Hoarau J-J, Jaffar Bandjee M-C, Krejbich Trotot P, 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. J Immunol 2010;184:5914–27. https://doi.org/10.4049/jimmunol.0900255.
[72] Couturier E, Guillemin F, Mura M, Léon L, Virion J-M, Letort M-J, et al. Impaired quality of life after chikungunya virus infection: a 2-year follow-up study. Rheumatology (Oxford) 2012;51:1315–22. https://doi.org/10.1093/rheumatology/kes015.
[73] Gérardin P, Fianu A, Malvy D, Mussard C, Boussaïd K, Rollot O, et al. Perceived morbidity and community burden after a Chikungunya outbreak: The TELECHIK survey, a population-based cohort study. BMC Med 2011;9:5. https://doi.org/10.1186/1741-7015-9-5.
[74] Sissoko D, Malvy D, Ezzedine K, Renault P, Moscetti F, Ledrans M, et al. Post-epidemic

Chikungunya disease on reunion island: Course of rheumatic manifestations and associated factors over a 15-month period. PLoS Negl Trop Dis 2009;3:389. https://doi.org/10.1371/journal.pntd.0000389.
[75] Li X-F, Jiang T, Deng Y-Q, Zhao H, Yu X-D, Ye Q, et al. Complete Genome Sequence of a Chikungunya Virus Isolated in Guangdong, China. J Virol 2012;86:8904–5. https://doi.org/10.1128/JVI.01289-12.
[76] Solignat M, Gay B, Higgs S, Briant L, Devaux C. Replication cycle of chikungunya: A re-emerging arbovirus. Virology 2009;393:183–97. https://doi.org/10.1016/j.virol.2009.07.024.
[77] Dubey SK, Shrinet J, Sunil S. Aedes aegypti microRNA, miR-2944b-5p interacts with 3'UTR of chikungunya virus and cellular target vps-13 to regulate viral replication. PLoS Negl Trop Dis 2019;13:e0007429. https://doi.org/10.1371/journal.pntd.0007429.
[78] Pfeffer M, Kinney RM, Kaaden OR. The Alphavirus 3'-nontranslated region: Size heterogeneity and arrangement of repeated sequence elements. Virology 1998;240:100–8. https://doi.org/10.1006/viro.1997.8907.
[79] Gebhart NN, Sokoloski KJ, Rupp JC, Hardy RW. Alphavirus RNA synthesis and non-structural protein functions. J Gen Virol 2015;96:2483–500. https://doi.org/10.1099/jgv.0.000249.
[80] Jones JE, Long KM, Whitmore AC, Sanders W, Thurlow LR, Brown JA, et al. Disruption of the opal stop codon attenuates chikungunya virus-induced arthritis and pathology. MBio 2017;8. https://doi.org/10.1128/mBio.01456-17.
[81] Sourisseau M, Schilte C, Casartelli N, Trouillet C, Guivel-Benhassine F, Rudnicka D, et al. Characterization of reemerging chikungunya virus. PLoS Pathog 2007;3:0804–17. https://doi.org/10.1371/journal.ppat.0030089.
[82] Labadie K, Larcher T, Joubert C, Mannioui A, Delache B, Brochard P, et al. Chikungunya disease in nonhuman primates involves long-term viral persistence in macrophages. J Clin Invest 2010;120:894–906. https://doi.org/10.1172/JCI40104.
[83] Zhang R, Kim AS, Fox JM, Nair S, Basore K, Klimstra WB, et al. Mxra8 is a receptor for multiple arthritogenic alphaviruses. Nature 2018;557:570–4. https://doi.org/10.1038/s41586-018-0121-3.
[84] van Duijl-Richter MKS, Hoornweg TE, Rodenhuis-Zybert IA, Smit JM. Early events in chikungunya virus infection—from virus cell binding to membrane fusion. Viruses 2015;7:3647–74. https://doi.org/10.3390/v7072792.
[85] Hoornweg TE, van Duijl-Richter MKS, Ayala Nuñez N V., Albulescu IC, van Hemert MJ, Smit JM. Dynamics of Chikungunya Virus Cell Entry Unraveled by Single-Virus Tracking in Living Cells. J Virol 2016;90:4745–56. https://doi.org/10.1128/jvi.03184-15.
[86] Voss JE, Vaney MC, Duquerroy S, Vonrhein C, Girard-Blanc C, Crublet E, et al. Glycoprotein organization of Chikungunya virus particles revealed by X-ray crystallography. Nature 2010;468:709–12. https://doi.org/10.1038/nature09555.
[87] Li L, Jose J, Xiang Y, Kuhn RJ, Rossmann MG. Structural changes of envelope proteins during alphavirus fusion. Nature 2010;468:705–8. https://doi.org/10.1038/nature09546.
[88] Kuo SC, Chen YJ, Wang YM, Tsui PY, Kuo M Der, Wu TY, et al. Cell-based analysis of Chikungunya virus E1 protein in membrane fusion. J Biomed Sci 2012;19:44. https://doi.org/10.1186/1423-0127-19-44.
[89] 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;67:5062–7. https://doi.org/10.1128/jvi.67.8.5062-5067.1993.
[90] Jones JE, Long KM, Whitmore AC, Sanders W, Thurlow LR, Brown JA, et al. Disruption of the opal stop codon attenuates chikungunya virus-induced arthritis and pathology. MBio 2017;8. https://doi.org/10.1128/mBio.01456-17.
[91] Firth AE, Wills NM, Gesteland RF, Atkins JF. Stimulation of stop codon readthrough: frequent presence of an extended 3' RNA structural element. Nucleic Acids Res 2011;39:6679–91. https://doi.org/10.1093/nar/gkr224.
[92] 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. vol. 68. 1994.
[93] Müller M, Slivinski N, Todd EJAA, Khalid H, Li R, Karwatka M, et al. Chikungunya virus requires cellular chloride channels for efficient genome replication. PLoS Negl Trop Dis 2019;13:e0007703. https://doi.org/10.1371/journal.pntd.0007703.
[94] Hardy RW, Rice CM. Requirements at the 3' End of the Sindbis Virus Genome for Efficient Synthesis of Minus-Strand RNA. J Virol 2005;79:4630–9. https://doi.org/10.1128/jvi.79.8.4630-4639.2005.
[95] Hahn CS, Strauss JH. Site-directed mutagenesis of the proposed catalytic amino acids of the Sindbis virus capsid protein autoprotease. J Virol 1990;64:3069–73. https://doi.org/10.1128/jvi.64.6.3069-3073.1990.
[96] Griffiths G, Simons K, Warren G, Tokuyasu KT. Immunoelectron microscopy using thin, frozen sections: Application to studies of the intracellular transport of Semliki Forest virus spike glycoproteins. Methods Enzymol 1983;96:466–85. https://doi.org/10.1016/S0076-6879(83)96041-X.
[97] Griffiths G, Quinn P, Warren G. Dissection of the Golgi complex. I. Monensin inhibits the transport of viral membrane proteins from medial to trans Golgi cisternae in baby hamster kidney cells infected with Semliki Forest virus. J Cell Biol 1983;96:835–50. https://doi.org/10.1083/jcb.96.3.835.
[98] Mollenhauer HH, Morré DJ, Rowe LD. Alteration of intracellular traffic by monensin; mechanism, specificity and relationship to toxicity. Biochim Biophys Acta - Rev Biomembr 1990;1031:225. https://doi.org/10.1016/0304-4157(90)90008-Z.
[99] Ozden S, Lucas-Hourani M, Ceccaldi PE, Basak A, Valentine M, Benjannet S, et al. Inhibition of Chikungunya virus infection in cultured human muscle cells by furin inhibitors: Impairment of the maturation of the E2 surface glycoprotein. J Biol Chem 2008;283:21899–908. https://doi.org/10.1074/jbc.M802444200.
[100] Soonsawad P, Xing L, Milla E, Espinoza JM, Kawano M, Marko M, et al. Structural Evidence of Glycoprotein Assembly in Cellular Membrane Compartments prior to Alphavirus Budding. J Virol 2010;84:11145–51. https://doi.org/10.1128/jvi.00036-10.
[101] Tellinghuisen TL, Hamburger AE, Fisher BR, Ostendorp R, Kuhn RJ. In vitro assembly of alphavirus cores by using nucleocapsid protein expressed in Escherichia coli. J Virol 1999;73:5309–19.
[102] Kim DY, Atasheva S, Frolova EI, Frolov I. Venezuelan Equine Encephalitis Virus nsP2 Protein Regulates Packaging of the Viral Genome into Infectious Virions. J Virol 2013;87:4202–13. https://doi.org/10.1128/jvi.03142-12.
[103] Jose J, Taylor AB, Kuhn RJ. Spatial and temporal analysis of alphavirus replication and assembly in mammalian and mosquito cells. MBio 2017;8. https://doi.org/10.1128/mBio.02294-16.

[104] Froshauer S, Kartenbeck J, Helenius A. Alphavirus RNA replicase is located on the cytoplasmic surface of endosomes and lysosomes. J Cell Biol 1988;107:2075–86. https://doi.org/10.1083/jcb.107.6.2075.
[105] Spuul P, Balistreri G, Hellstrom K, Golubtsov A V., Jokitalo E, Ahola T. Assembly of Alphavirus Replication Complexes from RNA and Protein Components in a Novel trans-Replication System in Mammalian Cells. J Virol 2011;85:4739–51. https://doi.org/10.1128/jvi.00085-11.
[106] 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;84:7543–57. https://doi.org/10.1128/jvi.00477-10.
[107] Thaa B, Biasiotto R, Eng K, Neuvonen M, Götte B, Rheinemann L, et al. Differential Phosphatidylinositol-3-Kinase-Akt-mTOR Activation by Semliki Forest and Chikungunya Viruses Is Dependent on nsP3 and Connected to Replication Complex Internalization. J Virol 2015;89:11420–37. https://doi.org/10.1128/jvi.01579-15.
[108] 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;7:1638–51. https://doi.org/10.1017/S135583820101010X.
[109] 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;78:4953–64. https://doi.org/10.1128/jvi.78.10.4953-4964.2004.
[110] Niesters HG, Strauss JH. Defined mutations in the 5' nontranslated sequence of Sindbis virus RNA. J Virol 1990;64:4162–8. https://doi.org/10.1128/jvi.64.9.4162-4168.1990.
[111] 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;83:8327–39. https://doi.org/10.1128/JVI.00586-09.
[112] Reynaud JM, Kim DY, Atasheva S, Rasalouskaya A, White JP, Diamond MS, et al. IFIT1 Differentially Interferes with Translation and Replication of Alphavirus Genomes and Promotes Induction of Type I Interferon. PLOS Pathog 2015;11:e1004863. https://doi.org/10.1371/journal.ppat.1004863.
[113] Wang HL, O'Rear J, Stollar V. Mutagenesis of the Sindbis virus nsP1 protein: Effects on methyltransferase activity and viral infectivity. Virology 1996;217:527–31. https://doi.org/10.1006/viro.1996.0147.
[114] 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. https://doi.org/10.1128/jvi.02183-19.
[115] Jones R, Bragagnolo G, Arranz R, Reguera J. Capping pores of alphavirus nsP1 gate membranous viral replication factories. Nat 2020 5897843 2020;589:615–9. https://doi.org/10.1038/s41586-020-3036-8.
[116] Law YS, Utt A, Tan YB, Zheng J, Wang S, Chen MW, et al. Structural insights into RNA recognition by the Chikungunya virus nsP2 helicase. Proc Natl Acad Sci U S A 2019;116:9558–67. https://doi.org/10.1073/pnas.1900656116.
[117] Vasiljeva L, Merits A, Auvinen P, Kääriäinen L. Identification of a novel function of the

Alphavirus capping apparatus. RNA 5'-triphosphatase activity of Nsp2. J Biol Chem 2000;275:17281–7. https://doi.org/10.1074/jbc.M910340199.
[118] Narwal M, Singh H, Pratap S, Malik A, Kuhn RJ, Kumar P, et al. Crystal structure of chikungunya virus nsP2 cysteine protease reveals a putative flexible loop blocking its active site. Int J Biol Macromol 2018;116:451–62. https://doi.org/10.1016/j.ijbiomac.2018.05.007.
[119] Rausalu K, Utt A, Quirin T, Varghese FS, Žusinaite E, Das PK, et al. Chikungunya virus infectivity, RNA replication and non-structural polyprotein processing depend on the nsP2 protease's active site cysteine residue. Sci Rep 2016;6:1–17. https://doi.org/10.1038/srep37124.
[120] 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;83:6534–45. https://doi.org/10.1128/jvi.00189-09.
[121] 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 Pathog 2019;15:e1007239. https://doi.org/10.1371/journal.ppat.1007239.
[122] Eckei L, Krieg S, Bütepage M, Lehmann A, Gross A, Lippok B, et al. The conserved macrodomains of the non-structural proteins of Chikungunya virus and other pathogenic positive strand RNA viruses function as mono-ADP-ribosylhydrolases. Sci Rep 2017;7. https://doi.org/10.1038/srep41746.
[123] 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. Proc Natl Acad Sci U S A 2017;114:1666–71. https://doi.org/10.1073/pnas.1621485114.
[124] Rupp JC, Jundt N, Hardy RW. Requirement for the Amino-Terminal Domain of Sindbis Virus nsP4 during Virus Infection. J Virol 2011;85:3449–60. https://doi.org/10.1128/jvi.02058-10.
[125] 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;384:201–8. https://doi.org/10.1016/j.virol.2008.10.030.
[126] Thal MA, Wasik BR, Posto J, Hardy RW. Template requirements for recognition and copying by Sindbis virus RNA-dependent RNA polymerase. Virology 2007;358:221–32. https://doi.org/10.1016/j.virol.2006.08.022.
[127] Tomar S, Hardy RW, Smith JL, Kuhn RJ. Catalytic Core of Alphavirus Nonstructural Protein nsP4 Possesses Terminal Adenylyltransferase Activity. J Virol 2006;80:9962–9. https://doi.org/10.1128/jvi.01067-06.
[128] Simmonds P, Becher P, Bukh J, Gould EA, Meyers G, Monath T, et al. ICTV virus taxonomy profile: Flaviviridae. J Gen Virol 2017;98:2–3. https://doi.org/10.1099/jgv.0.000672.
[129] Bhatt S, Gething PW, Brady OJ, Messina JP, Farlow AW, Moyes CL, et al. The global distribution and burden of dengue. Nature 2013;496:504–7. https://doi.org/10.1038/nature12060.
[130] França T, Medeiros W, Souza N, Longo E, Pereira S, França T, et al. Growth and Development of Children with Microcephaly Associated with Congenital Zika Virus Syndrome in Brazil. Int J Environ Res Public Health 2018;15:1990.

https://doi.org/10.3390/ijerph15091990.
[131] Chang HH, Huber RG, Bond PJ, Grad YH, Camerini D, Maurer-Stroh S, et al. Analyse systématique des similarités protéiques entre le virus Zika et d'autres virus transmis par des arthropodes. Bull World Health Organ 2017;95:517–25. https://doi.org/10.2471/BLT.16.182105.
[132] Tautz N, Tews BA, Meyers G. The Molecular Biology of Pestiviruses. Adv. Virus Res., vol. 93, Academic Press Inc.; 2015, p. 47–160. https://doi.org/10.1016/bs.aivir.2015.03.002.
[133] Petrelli E, Manzin A, Paolucci S, Cioppi A, Brugia M, Muretto P, et al. Chronic liver disease and active hepatitis C virus infection in patients with antibodies to this virus. J Clin Pathol 1994;47:148–51. https://doi.org/10.1136/jcp.47.2.148.
[134] Scheel TKH, Simmonds P, Kapoor A. Surveying the global virome: Identification and characterization of HCV-related animal hepaciviruses. Antiviral Res 2015;115:83–93. https://doi.org/10.1016/j.antiviral.2014.12.014.
[135] Stapleton JT, Foung S, Muerhoff AS, Bukh J, Simmonds P. The GB viruses: A review and proposed classification of GBV-A, GBV-C (HGV), and GBV-D in genus Pegivirus within the family Flaviviridae. J Gen Virol 2011;92:233–46. https://doi.org/10.1099/vir.0.027490-0.
[136] Dick GWA, Kitchen SF, Haddow AJ. Zika Virus (I). Isolations and serological specificity. Trans R Soc Trop Med Hyg 1952;46:509–20. https://doi.org/10.1016/0035-9203(52)90042-4.
[137] Dick GWA. Epidemiological notes on some viruses isolated in Uganda (Yellow fever, Rift Valles fever, Bwamba fever, West Nile, Mengo, Semliki forest, Bunyamwera, Ntaya, Uganda S and Zika Virus). Ordinary Meet R Soc Trop Med Hyg 1952;26.
[138] Song BH, Yun SI, Woolley M, Lee YM. Zika virus: History, epidemiology, transmission, and clinical presentation. J Neuroimmunol 2017;308:50–64. https://doi.org/10.1016/j.jneuroim.2017.03.001.
[139] MacNamara FN. Zika virus : A report on three cases of human infection during an epidemic of jaundice in Nigeria. Trans R Soc Trop Med Hyg 1954;48:139–45. https://doi.org/10.1016/0035-9203(54)90006-1.
[140] Duffy MR, Chen T-H, Hancock WT, Powers AM, Kool JL, Lanciotti RS, et al. Zika Virus Outbreak on Yap Island, Federated States of Micronesia. N Engl J Med 2009;360:2536–43. https://doi.org/10.1056/NEJMoa0805715.
[141] Haddow AD, Schuh AJ, Yasuda CY, Kasper MR, Heang V, Huy R, et al. Genetic Characterization of Zika Virus Strains: Geographic Expansion of the Asian Lineage. PLoS Negl Trop Dis 2012;6:e1477. https://doi.org/10.1371/journal.pntd.0001477.
[142] Cao-Lormeau VM, Roche C, Teissier A, Robin E, Berry AL, Mallet HP, et al. Zika virus, French Polynesia, South Pacific, 2013. Emerg Infect Dis 2014;20:1085–6. https://doi.org/10.3201/eid2006.140138.
[143] Oehler E, Watrin L, Larre P, Leparc-Goffart I, Lastere S, Valour F, et al. Zika virus infection complicated by Guillain-Barre syndrome - case report, French Polynesia, December 2013. Euro Surveill 2014;19:20720. https://doi.org/10.2807/1560-7917.es2014.19.9.20720.
[144] Zanluca C, De Melo VCA, Mosimann ALP, Dos Santos GIV, dos Santos CND, Luz K. First report of autochthonous transmission of Zika virus in Brazil. Mem Inst Oswaldo Cruz 2015;110:569–72. https://doi.org/10.1590/0074-02760150192.
[145] Schuler-Faccini L, Ribeiro EM, Feitosa IML, Horovitz DDG, Cavalcanti DP, Pessoa A, et

al. Possible Association Between Zika Virus Infection and Microcephaly — Brazil, 2015. MMWR Morb Mortal Wkly Rep 2016;65:59–62. https://doi.org/10.15585/mmwr.mm6503e2.
[146] Cauchemez S, Besnard M, Bompard P, Dub T, Guillemette-Artur P, Eyrolle-Guignot D, et al. Association between Zika virus and microcephaly in French Polynesia, 2013-15: A retrospective study. Lancet 2016;387:2125–32. https://doi.org/10.1016/S0140-6736(16)00651-6.
[147] Beaver JT, Lelutiu N, Habib R, Skountzou I. Evolution of two major Zika virus lineages: Implications for pathology, immune response, and vaccine development. Front Immunol 2018;9. https://doi.org/10.3389/fimmu.2018.01640.
[148] Pettersson JHO, Eldholm V, Seligman SJ, Lundkvist Å, Falconar AK, Gaunt MW, et al. How did zika virus emerge in the Pacific Islands and Latin America? MBio 2016;7. https://doi.org/10.1128/mBio.01239-16.
[149] Enfissi A, Codrington J, Roosblad J, Kazanji M, Rousset D. Zika virus genome from the Americas. Lancet 2016;387:227–8. https://doi.org/10.1016/S0140-6736(16)00003-9.
[150] Anfasa F, Siegers JY, van der Kroeg M, Mumtaz N, Stalin Raj V, de Vrij FMS, et al. Phenotypic Differences between Asian and African Lineage Zika Viruses in Human Neural Progenitor Cells. MSphere 2017;2. https://doi.org/10.1128/msphere.00292-17.
[151] Bhatnagar J, Rabeneck DB, Martines RB, Reagan-Steiner S, Ermias Y, Estetter LBC, et al. Zika virus RNA replication and persistence in brain and placental tissue. Emerg Infect Dis 2017;23:405–14. https://doi.org/10.3201/eid2303.161499.
[152] Villordo SM, Carballeda JM, Filomatori C V., Gamarnik A V. RNA Structure Duplications and Flavivirus Host Adaptation. Trends Microbiol 2016;24:270–83. https://doi.org/10.1016/j.tim.2016.01.002.
[153] Ye Q, Liu ZY, Han JF, Jiang T, Li XF, Qin CF. Genomic characterization and phylogenetic analysis of Zika virus circulating in the Americas. Infect Genet Evol 2016;43:43–9. https://doi.org/10.1016/j.meegid.2016.05.004.
[154] Gregory CJ, Oduyebo T, Brault AC, Brooks JT, Chung K-W, Hills S, et al. Modes of Transmission of Zika Virus The Journal of Infectious Diseases ® 2017;216(S10):S875-83. Zika Transm Modes • JID n.d.;2017:216. https://doi.org/10.1093/infdis/jix396.
[155] Diallo D, Sall AA, Diagne CT, Faye O, Faye O, Ba Y, et al. Zika virus emergence in mosquitoes in Southeastern Senegal, 2011. PLoS One 2014;9. https://doi.org/10.1371/journal.pone.0109442.
[156] Grard G, Caron M, Mombo IM, Nkoghe D, Mboui Ondo S, Jiolle D, et al. Zika Virus in Gabon (Central Africa) – 2007: A New Threat from Aedes albopictus? PLoS Negl Trop Dis 2014;8:e2681. https://doi.org/10.1371/journal.pntd.0002681.
[157] Chouin-Carneiro T, Vega-Rua A, Vazeille M, Yebakima A, Girod R, Goindin D, et al. Differential Susceptibilities of Aedes aegypti and Aedes albopictus from the Americas to Zika Virus. PLoS Negl Trop Dis 2016;10:e0004543. https://doi.org/10.1371/journal.pntd.0004543.
[158] Musso D, Nilles EJ, Cao-Lormeau VM. Rapid spread of emerging Zika virus in the Pacific area. Clin Microbiol Infect 2014;20:O595–6. https://doi.org/10.1111/1469-0691.12707.
[159] Ledermann JP, Guillaumot L, Yug L, Saweyog SC, Tided M, Machieng P, et al. Aedes hensilli as a Potential Vector of Chikungunya and Zika Viruses. PLoS Negl Trop Dis 2014;8:e3188. https://doi.org/10.1371/journal.pntd.0003188.

[160] Reiter P, Lathrop S, Bunning M, Biggerstaff B, Singer D, Tiwari T, et al. Texas lifestyle limits transmission of dengue virus. Emerg Infect Dis 2003;9:86–9. https://doi.org/10.3201/eid0901.020220.
[161] Foy BD, Kobylinski KC, Foy JLC, Blitvich BJ, da Rosa AT, Haddow AD, et al. Probable Non-Vector-borne Transmission of Zika Virus, Colorado, USA. Emerg Infect Dis 2011;17:880–2. https://doi.org/10.3201/eid1705.101939.
[162] Nicastri E, Castilletti C, Liuzzi G, Iannetta M, Capobianchi MR, Ippolito G. Persistent detection of Zika virus RNA in semen for six months after symptom onset in a traveller returning from Haiti to Italy, February 2016. Eurosurveillance 2016;21. https://doi.org/10.2807/1560-7917.ES.2016.21.32.30314.
[163] Paz-Bailey G, Rosenberg ES, Doyle K, Munoz-Jordan J, Santiago GA, Klein L, et al. Persistence of Zika virus in body fluids-Final report. N Engl J Med 2018;379:1234–43. https://doi.org/10.1056/NEJMoa1613108.
[164] Moreira J, Peixoto TM, Siqueira AM, Lamas CC. Sexually acquired Zika virus: a systematic review. Clin Microbiol Infect 2017;23:296–305. https://doi.org/10.1016/j.cmi.2016.12.027.
[165] Musso D, Gould E, Lanteri MC. Documentation of transfusion-transmitted arbovirus infections in endemic areas. Transfusion 2016;56:3143–4. https://doi.org/10.1111/trf.13861.
[166] Musso D, Nhan T, Robin E, Roche C, Bierlaire D, Zisou K, et al. Potential for Zika virus transmission through blood transfusion demonstrated during an outbreak in French Polynesia, November 2013 to February 2014. Eurosurveillance 2014;19. https://doi.org/10.2807/1560-7917.ES2014.19.14.20761.
[167] Calvet G, Aguiar RS, Melo ASO, Sampaio SA, de Filippis I, Fabri A, et al. Detection and sequencing of Zika virus from amniotic fluid of fetuses with microcephaly in Brazil: a case study. Lancet Infect Dis 2016;16:653–60. https://doi.org/10.1016/S1473-3099(16)00095-5.
[168] Mlakar J, Korva M, Tul N, Popović M, Poljšak-Prijatelj M, Mraz J, et al. Zika Virus Associated with Microcephaly. N Engl J Med 2016;374:951–8. https://doi.org/10.1056/NEJMoa1600651.
[169] Kurscheidt FA, Mesquita CSS, Damke GMZF, Damke E, Carvalho ARB d. A, Suehiro TT, et al. Persistence and clinical relevance of Zika virus in the male genital tract. Nat Rev Urol 2019;16:211–30. https://doi.org/10.1038/s41585-019-0149-7.
[170] Heang V, Yasuda CY, Sovann L, Haddow AD, da Rosa APT, Tesh RB, et al. Zika virus infection, Cambodia, 2010. Emerg Infect Dis 2012;18:349–51. https://doi.org/10.3201/eid1802.111224.
[171] Krow-Lucal ER, Biggerstaff BJ, Staples JE. Estimated incubation period for zika virus disease. Emerg Infect Dis 2017;23:841–4. https://doi.org/10.3201/eid2305.161715.
[172] Lazear HM, Diamond MS. Zika Virus: New Clinical Syndromes and Its Emergence in the Western Hemisphere. J Virol 2016;90:4864–75. https://doi.org/10.1128/jvi.00252-16.
[173] Barbi L, Coelho AVC, Alencar LCA de, Crovella S. Prevalence of Guillain-Barré syndrome among Zika virus infected cases: a systematic review and meta-analysis. Brazilian J Infect Dis 2018;22:137–41. https://doi.org/10.1016/j.bjid.2018.02.005.
[174] Blázquez A-B, Saiz J-C. Neurological manifestations of Zika virus infection. World J Virol 2016;5:135. https://doi.org/10.5501/wjv.v5.i4.135.
[175] Alter M. The epidemiology of Guillain-Barré syndrome. Ann Neurol 1990;27:S7–12.

https://doi.org/10.1002/ana.410270704.
[176] von der Hagen M, Pivarcsi M, Liebe J, von Bernuth H, Didonato N, Hennermann JB, et al. Diagnostic approach to microcephaly in childhood: a two-center study and review of the literature. Dev Med Child Neurol 2014;56:732–41. https://doi.org/10.1111/dmcn.12425.
[177] Woods CG. Human microcephaly. Curr Opin Neurobiol 2004;14:112–7. https://doi.org/10.1016/j.conb.2004.01.003.
[178] Oliveira Melo AS, Malinger G, Ximenes R, Szejnfeld PO, Alves Sampaio S, Bispo de Filippis AM. Zika virus intrauterine infection causes fetal brain abnormality and microcephaly: tip of the iceberg? Ultrasound Obstet Gynecol 2016;47:6–7. https://doi.org/10.1002/uog.15831.
[179] Fernandez-Garcia M-D, Mazzon M, Jacobs M, Amara A. Pathogenesis of Flavivirus Infections: Using and Abusing the Host Cell. Cell Host Microbe 2009;5:318–28. https://doi.org/10.1016/j.chom.2009.04.001.
[180] Coutard B, Barral K, Lichière J, Selisko B, Martin B, Aouadi W, et al. Zika Virus Methyltransferase: Structure and Functions for Drug Design Perspectives. J Virol 2017;91. https://doi.org/10.1128/JVI.02202-16.
[181] Polacek C, Friebe P, Harris E. Poly(A)-binding protein binds to the non-polyadenylated 3′ untranslated region of dengue virus and modulates translation efficiency. J Gen Virol 2009;90:687–92. https://doi.org/10.1099/vir.0.007021-0.
[182] Friebe P, Harris E. Interplay of RNA Elements in the Dengue Virus 5′ and 3′ Ends Required for Viral RNA Replication. J Virol 2010;84:6103–18. https://doi.org/10.1128/jvi.02042-09.
[183] Villordo SM, Gamarnik A V. Genome cyclization as strategy for flavivirus RNA replication. Virus Res 2009;139:230–9. https://doi.org/10.1016/j.virusres.2008.07.016.
[184] Miner JJ, Cao B, Govero J, Smith AM, Fernandez E, Cabrera OH, et al. Zika Virus Infection during Pregnancy in Mice Causes Placental Damage and Fetal Demise. Cell 2016;165:1081–91. https://doi.org/10.1016/j.cell.2016.05.008.
[185] Bayer A, Lennemann NJ, Ouyang Y, Bramley JC, Morosky S, Marques ETDA, et al. Type III Interferons Produced by Human Placental Trophoblasts Confer Protection against Zika Virus Infection. Cell Host Microbe 2016;19:705–12. https://doi.org/10.1016/j.chom.2016.03.008.
[186] Dudley DM, Aliota MT, Mohr EL, Weiler AM, Lehrer-Brey G, Weisgrau KL, et al. A rhesus macaque model of Asian-lineage Zika virus infection. Nat Commun 2016;7. https://doi.org/10.1038/ncomms12204.
[187] Adams Waldorf KM, Stencel-Baerenwald JE, Kapur RP, Studholme C, Boldenow E, Vornhagen J, et al. Fetal brain lesions after subcutaneous inoculation of Zika virus in a pregnant nonhuman primate. Nat Med 2016;22:1256–9. https://doi.org/10.1038/nm.4193.
[188] Govero J, Esakky P, Scheaffer SM, Fernandez E, Drury A, Platt DJ, et al. Zika virus infection damages the testes in mice. Nature 2016;540:438–42. https://doi.org/10.1038/nature20556.
[189] Sirohi D, Chen Z, Sun L, Klose T, Pierson TC, Rossmann MG, et al. The 3.8 A resolution cryo-EM structure of Zika virus. Science 2016;352:467–70. https://doi.org/10.1126/science.aaf5316.
[190] Kim SY, Zhao J, Liu X, Fraser K, Lin L, Zhang X, et al. Interaction of Zika Virus Envelope

Protein with Glycosaminoglycans. Biochemistry 2017;56:1151–62. https://doi.org/10.1021/acs.biochem.6b01056.
[191] Tan CW, Huan Hor CH, Kwek S Sen, Tee HK, Sam IC, Goh ELK, et al. Cell surface α2,3-linked sialic acid facilitates Zika virus internalization. Emerg Microbes Infect 2019;8:426–37. https://doi.org/10.1080/22221751.2019.1590130.
[192] Strange DP, Jiyarom B, Zarandi NP, Xie X, Baker C, Sadri-Ardekani H, et al. Axl promotes zika virus entry and modulates the antiviral state of human sertoli cells. MBio 2019;10. https://doi.org/10.1128/mBio.01372-19.
[193] Richard AS, Shim BS, Kwon YC, Zhang R, Otsuka Y, Schmitt K, et al. AXL-dependent infection of human fetal endothelial cells distinguishes Zika virus from other pathogenic flaviviruses. Proc Natl Acad Sci U S A 2017;114:2024–9. https://doi.org/10.1073/pnas.1620558114.
[194] Li M, Zhang D, Li C, Zheng Z, Fu M, Ni F, et al. Characterization of Zika Virus Endocytic Pathways in Human Glioblastoma Cells. Front Microbiol 2020;11. https://doi.org/10.3389/fmicb.2020.00242.
[195] Modis Y, Ogata S, Clements D, Harrison SC. Structure of the dengue virus envelope protein after membrane fusion. Nature 2004;427:313–9. https://doi.org/10.1038/nature02165.
[196] Moesker B, Rodenhuis-Zybert IA, Meijerhof T, Wilschut J, Smit JM. Characterization of the functional requirements of West Nile virus membrane fusion. J Gen Virol 2010;91:389–93. https://doi.org/10.1099/vir.0.015255-0.
[197] Falgout B, Markoff L. Evidence that flavivirus NS1-NS2A cleavage is mediated by a membrane-bound host protease in the endoplasmic reticulum. vol. 69. 1995.
[198] Yusof R, Clum S, Wetzel M, Murthy HMK, Padmanabhan R. Purified NS2B/NS3 serine protease of dengue virus type 2 exhibits cofactor NS2B dependence for cleavage of substrates with dibasic amino acids in vitro. J Biol Chem 2000;275:9963–9. https://doi.org/10.1074/jbc.275.14.9963.
[199] Miller S, Kastner S, Krijnse-Locker J, Bühler S, Bartenschlager R. The non-structural protein 4A of dengue virus is an integral membrane protein inducing membrane alterations in a 2K-regulated manner. J Biol Chem 2007;282:8873–82. https://doi.org/10.1074/jbc.M609919200.
[200] Welsch S, Miller S, Romero-Brey I, Merz A, Bleck CKE, Walther P, et al. Composition and Three-Dimensional Architecture of the Dengue Virus Replication and Assembly Sites. Cell Host Microbe 2009;5:365–75. https://doi.org/10.1016/j.chom.2009.03.007.
[201] Lorenz IC, Kartenbeck J, Mezzacasa A, Allison SL, Heinz FX, Helenius A. Intracellular Assembly and Secretion of Recombinant Subviral Particles from Tick-Borne Encephalitis Virus. J Virol 2003;77:4370–82. https://doi.org/10.1128/jvi.77.7.4370-4382.2003.
[202] Stadler K, Allison SL, Schalich J, Heinz FX. Proteolytic activation of tick-borne encephalitis virus by furin. J Virol 1997;71:8475–81. https://doi.org/10.1128/jvi.71.11.8475-8481.1997.
[203] Burlaud-Gaillard J, Sellin C, Georgeault S, Uzbekov R, Lebos C, Guillaume JM, et al. Correlative scanning-transmission electron microscopy reveals that a chimeric flavivirus is released as individual particles in secretory vesicles. PLoS One 2014;9:93573. https://doi.org/10.1371/journal.pone.0093573.
[204] Huber RG, Lim XN, Ng WC, Sim AYL, Poh HX, Shen Y, et al. Structure mapping of dengue and Zika viruses reveals functional long-range interactions. Nat Commun

2019;10:1408. https://doi.org/10.1038/s41467-019-09391-8.
[205] Clyde K, Barrera J, Harris E. The capsid-coding region hairpin element (cHP) is a critical determinant of dengue virus and West Nile virus RNA synthesis. Virology 2008;379:314–23. https://doi.org/10.1016/j.virol.2008.06.034.
[206] Liu Z-Y, Yu J-Y, Huang X-Y, Fan H, Li X-F, Deng Y-Q, et al. Characterization of cis-Acting RNA Elements of Zika Virus by Using a Self-Splicing Ribozyme-Dependent Infectious Clone. J Virol 2017;91. https://doi.org/10.1128/jvi.00484-17.
[207] Lodeiro MF, Filomatori C V., Gamarnik A V. Structural and Functional Studies of the Promoter Element for Dengue Virus RNA Replication. J Virol 2009;83:993–1008. https://doi.org/10.1128/jvi.01647-08.
[208] Cahour A, Pletnev A, Vazeille-Falcoz M, Rosen L, Lai CJ. Growth-Restricted Dengue Virus Mutants Containing Deletions in the 5′ Noncoding Region of the RNA Genome. Virology 1995;207:68–76. https://doi.org/10.1006/viro.1995.1052.
[209] Ng WC, Soto-Acosta R, Bradrick SS, Garcia-Blanco MA, Ooi EE. The 5′ and 3′ untranslated regions of the flaviviral genome. Viruses 2017;9:137. https://doi.org/10.3390/v9060137.
[210] Li XF, Jiang T, Yu XD, Deng YQ, Zhao H, Zhu QY, et al. RNA elements within the 5′ untranslated region of the West Nile virus genome are critical for RNA synthesis and virus replication. J Gen Virol 2010;91:1218–23. https://doi.org/10.1099/vir.0.013854-0.
[211] Alvarez DE, Lodeiro MF, Ludueña SJ, Pietrasanta LI, Gamarnik A V. Long-Range RNA-RNA Interactions Circularize the Dengue Virus Genome. J Virol 2005;79:6631–43. https://doi.org/10.1128/jvi.79.11.6631-6643.2005.
[212] Li P, Wei Y, Mei M, Tang L, Sun L, Huang W, et al. Integrative Analysis of Zika Virus Genome RNA Structure Reveals Critical Determinants of Viral Infectivity. Cell Host Microbe 2018;24:875-886.e5. https://doi.org/10.1016/J.CHOM.2018.10.011.
[213] Akiyama BM, Laurence HM, Massey AR, Costantino DA, Xie X, Yang Y, et al. Zika virus produces noncoding RNAs using a multi-pseudoknot structure that confounds a cellular exonuclease. Science 2016;354:1148–52. https://doi.org/10.1126/science.aah3963.
[214] Pijlman GP, Funk A, Kondratieva N, Leung J, Torres S, van der Aa L, et al. A Highly Structured, Nuclease-Resistant, Noncoding RNA Produced by Flaviviruses Is Required for Pathogenicity. Cell Host Microbe 2008;4:579–91. https://doi.org/10.1016/j.chom.2008.10.007.
[215] Chapman EG, Costantino DA, Rabe JL, Moon SL, Wilusz J, Nix JC, et al. The Structural Basis of Pathogenic Subgenomic Flavivirus RNA (sfRNA) Production. Science 2014;344:307–10. https://doi.org/10.1126/science.1250897.
[216] Schuessler A, Funk A, Lazear HM, Cooper DA, Torres S, Daffis S, et al. West Nile Virus Noncoding Subgenomic RNA Contributes to Viral Evasion of the Type I Interferon-Mediated Antiviral Response. J Virol 2012;86:5708–18. https://doi.org/10.1128/jvi.00207-12.
[217] Göertz GP, van Bree JWM, Hiralal A, Fernhout BM, Steffens C, Boeren S, et al. Subgenomic flavivirus RNA binds the mosquito DEAD/H-box helicase ME31B and determines Zika virus transmission by Aedes aegypti. Proc Natl Acad Sci U S A 2019;116:19136–44. https://doi.org/10.1073/pnas.1905617116.
[218] Sztuba-Solinska J, Teramoto T, Rausch JW, Shapiro BA, Padmanabhan R, Le Grice SFJ. Structural complexity of Dengue virus untranslated regions: cis -acting RNA motifs

and pseudoknot interactions modulating functionality of the viral genome . Nucleic Acids Res 2013;41:5075–89. https://doi.org/10.1093/nar/gkt203.
[219] 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. J Biol Chem 2011;286:22521–34. https://doi.org/10.1074/jbc.M111.234302.
[220] Villordo SM, Alvarez DE, Gamarnik A V. A balance between circular and linear forms of the dengue virus genome is crucial for viral replication. RNA 2010;16:2325–35. https://doi.org/10.1261/rna.2120410.
[221] Zeng L, Falgout B, Markoff L. Identification of Specific Nucleotide Sequences within the Conserved 3'-SL in the Dengue Type 2 Virus Genome Required for Replication. J Virol 1998;72:7510–22. https://doi.org/10.1128/jvi.72.9.7510-7522.1998.
[222] Davis WG, Basu M, Elrod EJ, Germann MW, Brinton MA. Identification of cis-Acting Nucleotides and a Structural Feature in West Nile Virus 3'-Terminus RNA That Facilitate Viral Minus Strand RNA Synthesis. J Virol 2013;87:7622–36. https://doi.org/10.1128/jvi.00212-13.
[223] Blackwell JL, Brinton MA. Translation elongation factor-1 alpha interacts with the 3' stem-loop region of West Nile virus genomic RNA. J Virol 1997;71:6433–44. https://doi.org/10.1128/jvi.71.9.6433-6444.1997.
[224] Filomatori C V., Lodeiro MF, Alvarez DE, Samsa MM, Pietrasanta L, Gamarnik A V. A 5' RNA element promotes dengue virus RNA synthesis on a circular genome. Genes Dev 2006;20:2238–49. https://doi.org/10.1101/gad.1444206.
[225] 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. https://doi.org/10.7554/eLife.17636.
[226] Filomatori C V., Iglesias NG, Villordo SM, Alvarez DE, Gamarnik A V. RNA sequences and structures required for the recruitment and activity of the dengue virus polymerase. J Biol Chem 2011;286:6929–39. https://doi.org/10.1074/jbc.M110.162289.
[227] Clyde K, Harris E. RNA Secondary Structure in the Coding Region of Dengue Virus Type 2 Directs Translation Start Codon Selection and Is Required for Viral Replication. J Virol 2006;80:2170–82. https://doi.org/10.1128/jvi.80.5.2170-2182.2006.
[228] Pestova T V., Kolupaeva VG, Lomakin IB, Pilipenko E V., Shatsky IN, Agol VI, et al. Molecular mechanisms of translation initiation in eukaryotes. Proc Natl Acad Sci U S A 2001;98:7029–36. https://doi.org/10.1073/pnas.111145798.
[229] Kozak M. Regulation of translation via mRNA structure in prokaryotes and eukaryotes. Gene 2005;361:13–37. https://doi.org/10.1016/j.gene.2005.06.037.
[230] Kozak M. Downstream secondary structure facilitates recognition of initiator codons by eukaryotic ribosomes. Proc Natl Acad Sci U S A 1990;87:8301–5. https://doi.org/10.1073/pnas.87.21.8301.
[231] Kozak M. Structural features in eukaryotic mRNAs that modulate the initiation of translation. J Biol Chem 1991;266:19867–70.
[232] Filomatori C V., Lodeiro MF, Alvarez DE, Samsa MM, Pietrasanta L, Gamarnik A V. A 5' RNA element promotes dengue virus RNA synthesis on a circular genome. Genes Dev 2006;20:2238–49. https://doi.org/10.1101/gad.1444206.
[233] Saeedi BJ, Geiss BJ. Regulation of flavivirus RNA synthesis and capping. Wiley Interdiscip Rev RNA 2013;4:723–35. https://doi.org/10.1002/wrna.1191.

[234] Hahn CS, Hahn YS, Rice CM, Lee E, Dalgarno L, Strauss EG, et al. Conserved elements in the 3' untranslated region of flavivirus RNAs and potential cyclization sequences. J Mol Biol 1987;198:33–41. https://doi.org/10.1016/0022-2836(87)90455-4.
[235] Lo MK, Tilgner M, Bernard KA, Shi P-Y. Functional Analysis of Mosquito-Borne Flavivirus Conserved Sequence Elements within 3' Untranslated Region of West Nile Virus by Use of a Reporting Replicon That Differentiates between Viral Translation and RNA Replication. J Virol 2003;77:10004–14. https://doi.org/10.1128/jvi.77.18.10004-10014.2003.
[236] Alvarez DE, De Lella Ezcurra AL, Fucito S, Gamarnik A V. Role of RNA structures present at the 3'UTR of dengue virus on translation, RNA synthesis, and viral replication. Virology 2005;339:200–12. https://doi.org/10.1016/j.virol.2005.06.009.
[237] Zhang B, Dong H, Stein DA, Iversen PL, Shi PY. West Nile virus genome cyclization and RNA replication require two pairs of long-distance RNA interactions. Virology 2008;373:1–13. https://doi.org/10.1016/j.virol.2008.01.016.
[238] Alvarez DE, Filomatori C V., Gamarnik A V. Functional analysis of dengue virus cyclization sequences located at the 5' and 3'UTRs. Virology 2008;375:223–35. https://doi.org/10.1016/j.virol.2008.01.014.
[239] Sanford TJ, Mears H V, Fajardo T, Locker N, Sweeney TR. Circularization of flavivirus genomic RNA inhibits de novo translation initiation. Nucleic Acids Res 2019. https://doi.org/10.1093/nar/gkz686.
[240] Brown EA, Zhang H, Ping LH, Lemon SM. Secondary structure of the 5' nontranslated regions of hepatitis C virus and pestivirus genomic RNAs. Nucleic Acids Res 1992;20:5041. https://doi.org/10.1093/NAR/20.19.5041.
[241] Tsukiyama-Kohara K, Iizuka N, Kohara M, Nomoto A. Internal ribosome entry site within hepatitis C virus RNA. J Virol 1992;66:1476–83. https://doi.org/10.1128/jvi.66.3.1476-1483.1992.
[242] Kieft JS, Zhou K, Jubin R, Murray MG, Lau JYN, Doudna JA. The Hepatitis C Virus Internal Ribosome Entry Site Adopts an Ion-dependent Tertiary Fold. J Mol Biol 1999;292:513–29. https://doi.org/10.1006/JMBI.1999.3095.
[243] Kieft JS, Zhou K, Grech A, Jubin R, Doudna JA. Crystal structure of an RNA tertiary domain essential to HCV IRES-mediated translation initiation. Nat Struct Biol 2002;9:370–4. https://doi.org/10.1038/nsb781.
[244] Lukavsky PJ, Kim I, Otto GA, Puglisi JD. Structure of HCV IRES domain II determined by NMR. Nat Struct Mol Biol 2003;10:1033–8. https://doi.org/10.1038/nsb1004.
[245] Fletcher SP, Jackson RJ. Pestivirus Internal Ribosome Entry Site (IRES) Structure and Function: Elements in the 5' Untranslated Region Important for IRES Function. J Virol 2002;76:5024. https://doi.org/10.1128/JVI.76.10.5024-5033.2002.
[246] R D, KV B. 5' and 3' untranslated regions of pestivirus genome: primary and secondary structure analyses. Nucleic Acids Res 1993;21:1949–57. https://doi.org/10.1093/NAR/21.8.1949.
[247] Pestova T V., Shatsky IN, Fletcher SP, Jackson RJ, Hellen CUT. A prokaryotic-like mode of cytoplasmic eukaryotic ribosome binding to the initiation codon during internal translation initation of hepatitis C and classical swine fever virus RNAs. Genes Dev 1998;12:67. https://doi.org/10.1101/GAD.12.1.67.
[248] Simons JN, Desai SM, Schultz DE, Lemon SM, Mushahwar IK. Translation initiation in GB viruses A and C: evidence for internal ribosome entry and implications for genome organization. J Virol 1996;70:6126.

[249] You S, Stump DD, Branch AD, Rice CM. A cis-Acting Replication Element in the Sequence Encoding the NS5B RNA-Dependent RNA Polymerase Is Required for Hepatitis C Virus RNA Replication. J Virol 2004;78:1352. https://doi.org/10.1128/JVI.78.3.1352-1366.2004.
[250] Friebe P, Bartenschlager R. Genetic Analysis of Sequences in the 3' Nontranslated Region of Hepatitis C Virus That Are Important for RNA Replication. J Virol 2002;76:5326. https://doi.org/10.1128/JVI.76.11.5326-5338.2002.
[251] KJ B, CM R. Secondary structure determination of the conserved 98-base sequence at the 3' terminus of hepatitis C virus genome RNA. J Virol 1997;71:7345–52. https://doi.org/10.1128/JVI.71.10.7345-7352.1997.
[252] Akey DL, Brown WC, Jose J, Kuhn RJ, Smith JL. Structure-guided insights on the role of NS1 in flavivirus infection. BioEssays 2015;37:489–94. https://doi.org/10.1002/bies.201400182.
[253] Young PR, Hilditch PA, Bletchly C, Halloran W. An antigen capture enzyme-linked immunosorbent assay reveals high levels of the dengue virus protein NS1 in the sera of infected patients. J Clin Microbiol 2000;38:1053–7. https://doi.org/10.1128/JCM.38.3.1053-1057.2000.
[254] Gutsche I, Coulibaly F, Voss JE, Salmon J, D'Alayer J, Ermonval M, et al. Secreted dengue virus nonstructural protein NS1 is an atypical barrel-shaped high-density lipoprotein. Proc Natl Acad Sci U S A 2011;108:8003–8. https://doi.org/10.1073/pnas.1017338108.
[255] Akey DL, Brown WC, Dutta S, Konwerski J, Jose J, Jurkiw TJ, et al. Flavivirus NS1 Structures Reveal Surfaces for Associations with Membranes and the Immune System. Science 2014;343:881–5. https://doi.org/10.1126/science.1247749.
[256] Muylaert IR, Galler R, Rice CM. Genetic analysis of the yellow fever virus NS1 protein: identification of a temperature-sensitive mutation which blocks RNA accumulation. J Virol 1997;71:291–8. https://doi.org/10.1128/jvi.71.1.291-298.1997.
[257] Mackenzie JM, Jones MK, Young PR. Immunolocalization of the Dengue virus nonstructural glycoprotein NS1 suggests a role in viral RNA replication. Virology 1996;220:232–40. https://doi.org/10.1006/viro.1996.0307.
[258] Westaway EG, Mackenzie JM, Kenney MT, Jones MK, Khromykh AA. Ultrastructure of Kunjin virus-infected cells: colocalization of NS1 and NS3 with double-stranded RNA, and of NS2B with NS3, in virus-induced membrane structures. J Virol 1997;71:6650–61. https://doi.org/10.1128/jvi.71.9.6650-6661.1997.
[259] Kyung MC, Liszewski MK, Nybakken G, Davis AE, Townsend RR, Fremont DH, et al. West Nile virus nonstructural protein NS1 inhibits complement activation by binding the regulatory protein factor H. Proc Natl Acad Sci U S A 2006;103:19111–6. https://doi.org/10.1073/pnas.0605668103.
[260] Zhang X, Xie X, Zou J, Xia H, Shan C, Chen X, et al. Genetic and biochemical characterizations of Zika virus NS2A protein. Emerg Microbes Infect 2019;8:585–602. https://doi.org/10.1080/22221751.2019.1598291.
[261] MacKenzie JM, Khromykh AA, Jones MK, Westaway EG. Subcellular localization and some biochemical properties of the flavivirus Kunjin nonstructural proteins NS2A and NS4A. Virology 1998;245:203–15. https://doi.org/10.1006/viro.1998.9156.
[262] Xie X, Zou J, Puttikhunt C, Yuan Z, Shi P-Y. Two Distinct Sets of NS2A Molecules Are Responsible for Dengue Virus RNA Synthesis and Virion Assembly. J Virol 2015;89:1298–313. https://doi.org/10.1128/jvi.02882-14.

[263] Wu R-H, Tsai M-H, Chao D-Y, Yueh A. Scanning Mutagenesis Studies Reveal a Potential Intramolecular Interaction within the C-Terminal Half of Dengue Virus NS2A Involved in Viral RNA Replication and Virus Assembly and Secretion. J Virol 2015;89:4281–95. https://doi.org/10.1128/JVI.03011-14.
[264] Muñoz-Jordán JL, Sánchez-Burgos GG, Laurent-Rolle M, García-Sastre A. Inhibition of interferon signaling by dengue virus. Proc Natl Acad Sci U S A 2003;100:14333–8. https://doi.org/10.1073/pnas.2335168100.
[265] Phoo WW, Li Y, Zhang Z, Lee MY, Loh YR, Tan YB, et al. Structure of the NS2B-NS3 protease from Zika virus after self-cleavage. Nat Commun 2016;7:1–8. https://doi.org/10.1038/ncomms13410.
[266] Noble CG, Seh CC, Chao AT, Shi PY. Ligand-Bound Structures of the Dengue Virus Protease Reveal the Active Conformation. J Virol 2012;86:438–46. https://doi.org/10.1128/jvi.06225-11.
[267] Luo D, Xu T, Hunke C, Grüber G, Vasudevan SG, Lescar J. Crystal Structure of the NS3 Protease-Helicase from Dengue Virus. J Virol 2008;82:173–83. https://doi.org/10.1128/jvi.01788-07.
[268] Xu S, Ci Y, Wang L, Yang Y, Zhang L, Xu C, et al. Zika virus NS3 is a canonical RNA helicase stimulated by NS5 RNA polymerase. Nucleic Acids Res 2019;47:8693–707. https://doi.org/10.1093/nar/gkz650.
[269] Issur M, Geiss BJ, Bougie I, Picard-Jean F, Despins S, Mayette J, et al. The flavivirus NS5 protein is a true RNA guanylyltransferase that catalyzes a two-step reaction to form the RNA cap structure. RNA 2009;15:2340–50. https://doi.org/10.1261/rna.1609709.
[270] Coutard B, Barral K, Lichière J, Selisko B, Martin B, Aouadi W, et al. Zika Virus Methyltransferase: Structure and Functions for Drug Design Perspectives. J Virol 2017;91. https://doi.org/10.1128/jvi.02202-16.
[271] Lin C, Amberg SM, Chambers TJ, Rice CM. Cleavage at a novel site in the NS4A region by the yellow fever virus NS2B-3 proteinase is a prerequisite for processing at the downstream 4A/4B signalase site. J Virol 1993;67:2327.
[272] Mertens E, Kajaste-Rudnitski A, Torres S, Funk A, Frenkiel MP, Iteman I, et al. Viral determinants in the NS3 helicase and 2K peptide that promote West Nile virus resistance to antiviral action of 2',5'-oligoadenylate synthetase 1b. Virology 2010;399:176–85. https://doi.org/10.1016/J.VIROL.2009.12.036.
[273] Stern O, Hung Y-F, Valdau O, Yaffe Y, Harris E, Hoffmann S, et al. An N-Terminal Amphipathic Helix in Dengue Virus Nonstructural Protein 4A Mediates Oligomerization and Is Essential for Replication. J Virol 2013;87:4080–5. https://doi.org/10.1128/jvi.01900-12.
[274] Miller S, Sparacio S, Bartenschlager R. Subcellular localization and membrane topology of the dengue virus type 2 non-structural protein 4B. J Biol Chem 2006;281:8854–63. https://doi.org/10.1074/jbc.M512697200.
[275] Liang Q, Luo Z, Zeng J, Chen W, Foo SS, Lee SA, et al. Zika Virus NS4A and NS4B Proteins Deregulate Akt-mTOR Signaling in Human Fetal Neural Stem Cells to Inhibit Neurogenesis and Induce Autophagy. Cell Stem Cell 2016;19:663–71. https://doi.org/10.1016/j.stem.2016.07.019.
[276] Mizushima N, Yoshimori T. How to interpret LC3 immunoblotting. Autophagy 2007;3:542–5. https://doi.org/10.4161/auto.4600.
[277] Grant A, Ponia SS, Tripathi S, Balasubramaniam V, Miorin L, Sourisseau M, et al. Zika

Virus Targets Human STAT2 to Inhibit Type i Interferon Signaling. Cell Host Microbe 2016;19:882–90. https://doi.org/10.1016/j.chom.2016.05.009.
[278] Zhao B, Yi G, Du F, Chuang YC, Vaughan RC, Sankaran B, et al. Structure and function of the Zika virus full-length NS5 protein. Nat Commun 2017;8. https://doi.org/10.1038/ncomms14762.
[279] Egloff M-P, Benarroch D, Selisko B, Romette J-L, Canard B. An RNA cap (nucleoside-2'-O-)-methyltransferase in the flavivirus RNA polymerase NS5: crystal structure and functional characterization. EMBO J 2002;21:2757–68. https://doi.org/10.1093/emboj/21.11.2757.
[280] Zamecnik PC, Stephenson ML. Inhibition of Rous sarcoma virus replication and cell transformation by a specific oligodeoxynucleotide. Proc Natl Acad Sci U S A 1978;75:280–4. https://doi.org/10.1073/pnas.75.1.280.
[281] Xu Y, Villa A, Nilsson L. The free energy of locking a ring: Changing a deoxyribonucleoside to a locked nucleic acid. J Comput Chem 2017;38:1147–57. https://doi.org/10.1002/jcc.24692.
[282] Grünweller A, Hartmann RK. Locked Nucleic Acid Oligonucleotides. BioDrugs 2007;21:235–43. https://doi.org/10.2165/00063030-200721040-00004.
[283] Suresh G, Priyakumar UD. Structures, Dynamics, and Stabilities of Fully Modified Locked Nucleic Acid (β- d -LNA and α- l -LNA) Duplexes in Comparison to Pure DNA and RNA Duplexes. J Phys Chem B 2013;117:5556–64. https://doi.org/10.1021/jp4016068.
[284] Crooke ST. Molecular mechanisms of action of antisense drugs. Biochim Biophys Acta - Gene Struct Expr 1999;1489:31–43. https://doi.org/10.1016/S0167-4781(99)00148-7.
[285] Wu H, Lima WF, Zhang H, Fan A, Sun H, Crooke ST. Determination of the Role of the Human RNase H1 in the Pharmacology of DNA-like Antisense Drugs. J Biol Chem 2004;279:17181–9. https://doi.org/10.1074/jbc.M311683200.
[286] Braasch DA, Corey DR. Locked nucleic acid (LNA): Fine-tuning the recognition of DNA and RNA. Chem Biol 2001;8:1–7. https://doi.org/10.1016/S1074-5521(00)00058-2.
[287] Kurreck J, Wyszko E, Gillen C, Erdmann VA. Design of antisense oligonucleotides stabilized by locked nucleic acids. Nucleic Acids Res 2002;30:1911–8. https://doi.org/10.1093/nar/30.9.1911.
[288] Hagedorn PH, Persson R, Funder ED, Albæk N, Diemer SL, Hansen DJ, et al. Locked nucleic acid: modality, diversity, and drug discovery. Drug Discov Today 2018;23:101–14. https://doi.org/10.1016/j.drudis.2017.09.018.
[289] 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:e28923. https://doi.org/10.1371/journal.pone.0028923.
[290] Mutso M, Saul S, Rausalu K, Susova O, Žusinaite E, Mahalingam S, et al. Reverse genetic system, genetically stable reporter viruses and packaged subgenomic replicon based on a Brazilian Zika virus isolate. J Gen Virol 2017;98:2712–24. https://doi.org/10.1099/JGV.0.000938.
[291] Karabiber F, McGinnis JL, Favorov O V., Weeks KM. QuShape: Rapid, accurate, and best-practices quantification of nucleic acid probing information, resolved by capillary electrophoresis. RNA 2013;19:63–73. https://doi.org/10.1261/rna.036327.112.
[292] Reuter JS, Mathews DH. RNAstructure: Software for RNA secondary structure

prediction and analysis. BMC Bioinformatics 2010;11:129. https://doi.org/10.1186/1471-2105-11-129.
[293] Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 2003;31:3406–15. https://doi.org/10.1093/nar/gkg595.
[294] 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 Res 2015;43:2914–26. https://doi.org/10.1093/nar/gkv142.
[295] Elmén J, Zhang H-Y, Zuber B, Ljungberg K, Wahren B, Wahlestedt C, et al. Locked nucleic acid containing antisense oligonucleotides enhance inhibition of HIV-1 genome dimerization and inhibit virus replication. FEBS Lett 2004;578:285–90. https://doi.org/10.1016/j.febslet.2004.11.015.
[296] Ducongé F, Toulmé JJ. In vitro selection identifies key determinants for loop-loop interactions: RNA aptamers selective for the TAR RNA element of HIV-1. RNA 1999;5:1605–14. https://doi.org/10.1017/S1355838299991318.
[297] Darfeuille F, Hansen JB, Orum H, Primo C Di, Toulmé J. LNA/DNA chimeric oligomers mimic RNA aptamers targeted to the TAR RNA element of HIV-1. Nucleic Acids Res 2004;32:3101–7. https://doi.org/10.1093/nar/gkh636.
[298] Chery J, Petri A, Wagschal A, Lim S-Y, Cunningham J, Vasudevan S, et al. Development of Locked Nucleic Acid Antisense Oligonucleotides Targeting Ebola Viral Proteins and Host Factor Niemann-Pick C1. Nucleic Acid Ther 2018;28:273–84. https://doi.org/10.1089/nat.2018.0722.
[299] Gebert LFR, Rebhan MAE, Crivelli SEM, Denzler R, Stoffel M, Hall J. Miravirsen (SPC3649) can inhibit the biogenesis of miR-122. Nucleic Acids Res 2014;42:609–21. https://doi.org/10.1093/nar/gkt852.
[300] Titze-de-Almeida R, David C, Titze-de-Almeida SS. The Race of 10 Synthetic RNAi-Based Drugs to the Pharmaceutical Market. Pharm Res 2017;34:1339–63. https://doi.org/10.1007/s11095-017-2134-2.
[301] 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;7:1638–51. https://doi.org/10.1017/s135583820101010x.
[302] Pietilä MK, van Hemert MJ, Ahola T. Purification of Highly Active Alphavirus Replication Complexes Demonstrates Altered Fractionation of Multiple Cellular Membranes. J Virol 2018;92. https://doi.org/10.1128/jvi.01852-17.
[303] Pietilä MK, Albulescu IC, Hemert MJ van, Ahola T. Polyprotein Processing as a Determinant for in Vitro Activity of Semliki Forest Virus Replicase. Viruses 2017;9:292. https://doi.org/10.3390/v9100292.
[304] Clewley JP, Kennedy SIT. Purification and Polypeptide Composition of Semliki Forest Virus RNA Polymerase. J Gen Virol 1976;32:395–411. https://doi.org/10.1099/0022-1317-32-3-395.
[305] Frolova EI, Gorchakov R, Pereboeva L, Atasheva S, Frolov I. Functional Sindbis Virus Replicative Complexes Are Formed at the Plasma Membrane. J Virol 2010;84:11679–95. https://doi.org/10.1128/jvi.01441-10.
[306] Jose J, Taylor AB, Kuhn RJ. Spatial and temporal analysis of alphavirus replication and assembly in mammalian and mosquito cells. MBio 2017;8. https://doi.org/10.1128/mBio.02294-16.
[307] Hua Y, Vickers TA, Okunola HL, Bennett CF, Krainer AR. Antisense Masking of an hnRNP A1/A2 Intronic Splicing Silencer Corrects SMN2 Splicing in Transgenic Mice.

Am J Hum Genet 2008;82:834–48. https://doi.org/10.1016/j.ajhg.2008.01.014.
[308] Bennett CF. Therapeutic Antisense Oligonucleotides Are Coming of Age. Annu Rev Med 2019;70:307–21. https://doi.org/10.1146/annurev-med-041217-010829.
[309] D'Ambrosio R, Degasperi E, Colombo M, Aghemo A. Direct-acting antivirals: the endgame for hepatitis C? Curr Opin Virol 2017;24:31–7. https://doi.org/10.1016/j.coviro.2017.03.017.
[310] Connelly CM, Abulwerdi FA, Schneekloth JS. Discovery of RNA Binding Small Molecules Using Small Molecule Microarrays. Methods Mol. Biol., vol. 1518, Humana Press Inc.; 2017, p. 157–75. https://doi.org/10.1007/978-1-4939-6584-7_11.
[311] Fernández-Sanlés A, Ríos-Marco P, Romero-López C, Berzal-Herranz A. Functional information stored in the conserved structural RNA domains of flavivirus genomes. Front Microbiol 2017;8. https://doi.org/10.3389/fmicb.2017.00546.
[312] Nilsen TW. RNA structure determination using nuclease digestion. Cold Spring Harb Protoc 2013;8:379–82. https://doi.org/10.1101/pdb.prot072330.
[313] Ziehler WA, Engelke DR. Probing RNA Structure with Chemical Reagents and Enzymes. Curr Protoc Nucleic Acid Chem 2000;00:6.1.1-6.1.21. https://doi.org/10.1002/0471142700.nc0601s00.
[314] Merino EJ, Wilkinson KA, Coughlan JL, Weeks KM. RNA structure analysis at single nucleotide resolution by Selective 2'-Hydroxyl Acylation and Primer Extension (SHAPE). J Am Chem Soc 2005;127:4223–31. https://doi.org/10.1021/ja043822v.
[315] Weeks KM, Mauger DM. Exploring RNA structural codes with SHAPE chemistry. Acc Chem Res 2011;44:1280–91. https://doi.org/10.1021/ar200051h.
[316] Wilkinson KA, Merino EJ, Weeks KM. Selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE): quantitative RNA structure analysis at single nucleotide resolution. Nat Protoc 2006;1:1610–6. https://doi.org/10.1038/nprot.2006.249.
[317] McGinnis JL, Duncan CDS, Weeks KM. High-throughput SHAPE and hydroxyl radical analysis of RNA structure and ribonucleoprotein assembly. Methods Enzymol 2009;468:67–89. https://doi.org/10.1016/s0076-6879(09)68004-6.
[318] Spasic A, Assmann SM, Bevilacqua PC, Mathews DH. Modeling RNA secondary structure folding ensembles using SHAPE mapping data. Nucleic Acids Res 2017;46:314–23. https://doi.org/10.1093/nar/gkx1057.
[319] Deigan KE, Li TW, Mathews DH, Weeks KM. Accurate SHAPE-directed RNA structure determination. Proc Natl Acad Sci U S A 2009;106:97–102. https://doi.org/10.1073/pnas.0806929106.
[320] Hajdin CE, Bellaousov S, Huggins W, Leonard CW, Mathews DH, Weeks KM. Accurate SHAPE-directed RNA secondary structure modeling, including pseudoknots. Proc Natl Acad Sci U S A 2013;110:5498–503. https://doi.org/10.1073/pnas.1219988110.
[321] Rice GM, Leonard CW, Weeks KM. RNA secondary structure modeling at consistent high accuracy using differential SHAPE. RNA 2014;20:846–54. https://doi.org/10.1261/rna.043323.113.
[322] Kenyon J, Prestwood L, Lever A. Current perspectives on RNA secondary structure probing. Biochem. Soc. Trans., vol. 42, Portland Press Ltd; 2014, p. 1251–5. https://doi.org/10.1042/BST20140084.
[323] Alvarez DE, Lodeiro MF, Ludueña SJ, Pietrasanta LI, Gamarnik A V. Long-Range RNA-RNA Interactions Circularize the Dengue Virus Genome. J Virol 2005;79:6631–43. https://doi.org/10.1128/jvi.79.11.6631-6643.2005.
[324] Cui T, Sugrue RJ, Xu Q, Lee AKW, Chan YC, Fu J. Recombinant dengue virus type 1 NS3

protein exhibits specific viral RNA binding and NTPase activity regulated by the NS5 protein. Virology 1998;246:409–17. https://doi.org/10.1006/viro.1998.9213.
[325] Bujalowski PJ, Bujalowski W, Choi KH. Interactions between the Dengue Virus Polymerase NS5 and Stem-Loop A. J Virol 2017;91:47–64. https://doi.org/10.1128/jvi.00047-17.
[326] Kenyon JC, Prestwood LJ, Le Grice SFJ, Lever AML. In-gel probing of individual RNA conformers within a mixed population reveals a dimerization structural switch in the HIV-1 leader. Nucleic Acids Res 2013;41:e174–e174. https://doi.org/10.1093/nar/gkt690.
[327] Thurner C, Witwer C, Hofacker IL, Stadler PF. Conserved RNA secondary structures in Flaviviridae genomes. J Gen Virol 2004;85:1113–24. https://doi.org/10.1099/vir.0.19462-0.
[328] Friebe P, Boudet J, Simorre J-P, Bartenschlager R. Kissing-Loop Interaction in the 3' End of the Hepatitis C Virus Genome Essential for RNA Replication. J Virol 2005;79:380–92. https://doi.org/10.1128/jvi.79.1.380-392.2005.
[329] Raviprakash K, Liu K, Matteucci M, Wagner R, Riffenburgh R, Carl M. Inhibition of dengue virus by novel, modified antisense oligonucleotides. J Virol 1995;69:69–74. https://doi.org/10.1128/jvi.69.1.69-74.1995.
[330] Deas TS, Binduga-Gajewska I, Tilgner M, Ren P, Stein DA, Moulton HM, et al. Inhibition of Flavivirus Infections by Antisense Oligomers Specifically Suppressing Viral Translation and RNA Replication. J Virol 2005;79:4599–609. https://doi.org/10.1128/jvi.79.8.4599-4609.2005.
[331] Summerton J. Morpholino antisense oligomers: The case for an RNase H-independent structural type. Biochim Biophys Acta - Gene Struct Expr 1999;1489:141–58. https://doi.org/10.1016/S0167-4781(99)00150-5.
[332] Summerton J, Weller D. Morpholino antisense oligomers: Design, preparation, and properties. Antisense Nucleic Acid Drug Dev 1997;7:187–95. https://doi.org/10.1089/oli.1.1997.7.187.
[333] Nakase I, Akita H, Kogure K, Gräslund A, Langel Ü, Harashima H, et al. Efficient intracellular delivery of nucleic acid pharmaceuticals using cell-penetrating peptides. Acc Chem Res 2012;45:1132–9. https://doi.org/10.1021/ar200256e.
[334] Nakase I, Noguchi K, Aoki A, Takatani-Nakase T, Fujii I, Futaki S. Arginine-rich cell-penetrating peptide-modified extracellular vesicles for active macropinocytosis induction and efficient intracellular delivery. Sci Rep 2017;7:1–11. https://doi.org/10.1038/s41598-017-02014-6.
[335] Minskaia E, Nicholson J, Ryan MD. Optimisation of the foot-and-mouth disease virus 2A co-expression system for biomedical applications. BMC Biotechnol 2013;13:67. https://doi.org/10.1186/1472-6750-13-67.
[336] Goh KCM, Tang CK, Norton DC, Gan ES, Tan HC, Sun B, et al. Molecular determinants of plaque size as an indicator of dengue virus attenuation. Sci Rep 2016;6:1–11. https://doi.org/10.1038/srep26100.
[337] Young DF, Andrejeva L, Livingstone A, Goodbourn S, Lamb RA, Collins PL, et al. Virus Replication in Engineered Human Cells That Do Not Respond to Interferons. J Virol 2003;77:2174–81. https://doi.org/10.1128/jvi.77.3.2174-2181.2003.
[338] Desmyter J, Melnick JL, Rawls WE. Defectiveness of Interferon Production and of Rubella Virus Interference in a Line of African Green Monkey Kidney Cells (Vero). J Virol 1968;2:955–61. https://doi.org/10.1128/jvi.2.10.955-961.1968.

[339] Butrapet S, Huang CY-H, Pierro DJ, Bhamarapravati N, Gubler DJ, Kinney RM. Attenuation Markers of a Candidate Dengue Type 2 Vaccine Virus, Strain 16681 (PDK-53), Are Defined by Mutations in the 5' Noncoding Region and Nonstructural Proteins 1 and 3. J Virol 2000;74:3011–9. https://doi.org/10.1128/jvi.74.7.3011-3019.2000.
[340] Wang C, Le SY, Ali N, Siddiqui A. An RNA pseudoknot is an essential structural element of the internal ribosome entry site located within the hepatitis C virus 5' noncoding region. RNA 1995;1:526–37.
[341] Nimjee SM, White RR, Becker RC, Sullenger BA. Aptamers as Therapeutics. Annu Rev Pharmacol Toxicol 2017;57:61–79. https://doi.org/10.1146/annurev-pharmtox-010716-104558.
[342] Tiede C, Bedford R, Heseltine SJ, Smith G, Wijetunga I, Ross R, et al. Affimer proteins are versatile and renewable affinity reagents. Elife 2017;6:24903. https://doi.org/10.7554/eLife.24903.
[343] Tiede C, Tang AAS, Deacon SE, Mandal U, Nettleship JE, Owen RL, et al. Adhiron: A stable and versatile peptide display scaffold for molecular recognition applications. Protein Eng Des Sel 2014;27:145–55. https://doi.org/10.1093/protein/gzu007.
[344] Staedel C, Tran TPA, Giraud J, Darfeuille F, Di Giorgio A, Tourasse NJ, et al. Modulation of oncogenic miRNA biogenesis using functionalized polyamines. Sci Rep 2018;8:1667. https://doi.org/10.1038/s41598-018-20053-5.
[345] Ratni H, Ebeling M, Baird J, Bendels S, Bylund J, Chen KS, et al. Discovery of Risdiplam, a Selective Survival of Motor Neuron-2 (SMN2) Gene Splicing Modifier for the Treatment of Spinal Muscular Atrophy (SMA). J Med Chem 2018;61:6501–17. https://doi.org/10.1021/acs.jmedchem.8b00741.