# Structural and biochemical analysis of protein/RNA interactions during the initiation of dengue virus genome replication

Lauren Elizabeth Branfield

**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

May 2018

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

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

© 2018 The University of Leeds and Lauren Elizabeth Branfield

The right of Lauren Elizabeth Branfield to be identified as Author of this work has been asserted by her in accordance with the Copyright, Designs and Patents Act 1988.

## Acknowledgements

**Chapter 3:** Dr. Chi Trinh provided invaluable assistance, picked crystals, collected and refined datasets and aided in solving the structure of DENV-2 RdRp by molecular replacement. Thank you for teaching me about X-ray crystallography. Ed, thank you very much for your help and advice.

With thanks to Dr. Andrew Davidson (University of Bristol) for supplying the DENV-2 clones and for providing experimental protocols.

Thank you to the Medical Research Council for funding this project.

I would primarily like to thank Andy. Your invaluable help and guidance has got me to this point, from undergraduate until now… who would have thought it! Thank you for never panicking when I did and for being a constant source of encouragement.

Thank you to Mark and Anastasia for all your help and advice in the conduct of this project. Marietta, thank you for your priceless advice in the writing of this thesis.

I would like to thank everybody in Garstang 8.61 for the vital help and fun times throughout the years. You are the reason why 10.30 am necessitates coffee and cake. Emma, Chris, Joe and Andrew – thank you for your constant and continued friendship.

Ross, who took on a final year PhD student. You have been a continuous source of fun and care throughout this process, most importantly making me see the bigger picture. Thank you…….the list really does go on, I cannot wait!

Becky, there for me every day – a ray of sunshine to come home to. I will miss this so much. Thank you for being a constant voice of reason and support.

**To Mum, Dad, Ben and Matthew who will probably only read this page in the entire thesis…. Your blind faith and encouragement has never faltered. Words cannot express how much that means.**

## Abstract

The dengue virus (DENV) genome contains *cis*-acting RNA structures and undergoes structural rearrangement between alternative linear and cyclised conformations. Cyclisation is essential for the initiation of viral genome replication and involves remodeling of local RNA structures to transfer the RNA- dependent RNA polymerase (RdRp) from a 5´ promoter (stem-loop A (SLA)) to the 3´ promoter. The RNA/protein interactions involved in stabilising/initiating DENV genome conformational switching remain unclear. The 5´upstream of AUG flanking stem (UFS) at the base of stem-loop B (SLB) is suggested to be a regulatory element controlling the conformational change, yet its structure and influence remain uncertain in the context of the full-length viral genome and during active virus replication. The aim of this study was to analyse DENV genome/protein conformational changes associated with the initiation of viral genome replication.

To investigate the role of DENV-2 RdRp in initiating RNA conformational change and to interrogate RdRp structural changes associated with the initiation of replication, DENV-2 RdRp was expressed, purified and activity confirmed. X-ray crystallography was used to solve the first molecular structure of DENV-2 RdRp, presented here at 2.2 Å resolution, and RdRp/promoter RNA co-crystallization was attempted.

The influence of viral/host cell factors on initiation/stabilisation of RNA conformational changes within the DENV genome, associated with the initiation of genome replication, was investigated using Selective 2' Hydroxyl Acylation Analysed by Primer Extension (SHAPE) mapping. *In vitro*, we observed that RdRp and *trans*-activating factors destabilise local RNA structures, increasing the flexibility of adjacent unpaired regions. We hypothesise that this is the first step in genomic cyclisation - a critical stage in DENV negative-strand replication. Furthermore, we optimized a novel technique of intracellular SHAPE, mapping DENV RNA structure inside the cell during active viral replication. Intracellularly, we observed the linear genomic form to be prevalent in the host cell cytosol. In contrast, at the sites of DENV genome replication, within active viral replication complexes, the circular genomic conformation was observed.

# Table of Contents

**Acknowledgements** ..... **II**

**Abstract** ..... **III**

**Table of Figures** ..... **VIII**

**Table of Tables** ..... **XII**

**Abbreviations** ..... **XIII**

**Chapter 1 Introduction** ..... **1**

1.1 General introduction ..... 2

1.1.1 The *Flaviviridae* family, Flaviviruses and the emergence of DENV ..... 2

1.1.2 Epidemiology ..... 9

1.1.3 The pathogenesis of DENV infection ..... 12

1.1.4 Treatment ..... 16

1.1.5 The mosquito vector and control strategies ..... 17

1.1.6 Animal models of DENV infection ..... 20

1.2 DENV ..... 21

1.2.1 Virion structure ..... 21

1.2.2 The DENV life-cycle ..... 23

1.2.3 Genomic organisation ..... 26

1.3 DENV replication ..... 31

1.3.1 Non-structural protein 5 in detail ..... 31

1.3.2 RNA structure and DENV genomic replication ..... 42

1.4 Project aim and objectives ..... 60

**Chapter 2 Materials and methods** ..... **61**

2.1 Materials ..... 62

2.1.1 Plasmids and vectors ..... 62

2.1.2 Bacterial strains ..... 62

2.1.3 Continuous mammalian cell lines ..... 62

2.1.4 Continuous mosquito cell lines ..... 63

2.1.5 Primers ..... 63

2.1.6 Antibodies ..... 63

2.1.7 Buffer, media and stain preparation ..... 64

2.2 Molecular biological methods ..... 64

2.2.1 Agarose gel electrophoresis ..... 64

2.2.2 RNA quality analysis using denaturing formaldehyde MOPS agarose gel electrophoresis ..... 64
2.2.3 SDS-polyacrylamide gel electrophoresis (SDS-PAGE) ..... 65
2.2.4 Protein visualisation by Coomassie stain ..... 65
2.2.5 Western blot analysis ..... 65
2.2.6 Restriction endonuclease digestion ..... 66
2.2.7 The *in vitro* transcription of capped and uncapped RNA transcripts ..... 66
2.2.8 Polymerase Chain Reaction ..... 67
2.2.9 Bacterial transformation ..... 68
2.2.10 Ligations ..... 68
2.2.11 Plasmid DNA amplification ..... 68
2.2.12 Reverse transcription PCR (RT-PCR) ..... 68
2.2.13 Establishment of DENV-2 stable cell line ..... 69
2.2.14 Isolation of HEK293T DENV-2 Replicon Cell Cytosol (S100) protein extract ..... 69
2.2.15 RNA transfection and generation of infectious DENV-2 ..... 70
2.2.16 Titration of infectious DENV-2 and multiplicity of infection determination ..... 70
2.3 Protein expression and purification in *E.coli* ..... 71
2.3.1 Bacterial expression strain selection ..... 71
2.3.2 Optimisation of IPTG concentration and temperature for induction ..... 72
2.3.3 Bacterial induction and lysis for protein purification ..... 72
2.3.4 Nickel affinity chromatography ..... 73
2.3.5 Size exclusion chromatography ..... 73
2.3.6 Storage of purified DENV RdRp ..... 74
2.4 Determination of DENV-2 RdRp activity ..... 74
2.4.1 Circular Dichroism ..... 74
2.4.2 Polymerase activity assay ..... 74
2.4.3 Fluorescence Anisotropy ..... 75
2.4.4 Electromobility Shift assay (EMSA) ..... 75
2.5 DENV-2 RdRp crystallographic techniques ..... 77
2.5.1 Sparse matrix screening ..... 77
2.5.2 Crystal optimisation ..... 78
2.5.3 Crystal streak seeding ..... 78

2.5.4 Crystal harvesting and data collection ..... 78
2.5.5 Data processing and structure solution ..... 79
2.6 SHAPE mapping ..... 80
2.6.1 *In vitro* SHAPE mapping ..... 80
2.6.2 Intracellular SHAPE ..... 81
2.6.3 SHAPE data analysis ..... 83
**Chapter 3 Expression, purification, functional and structural analysis of DENV-2 RdRp ..... 85**
3.1 Introduction ..... 86
3.1.1 Introduction to the bacterial expression of DENV-2 RdRp .... 86
3.1.2 Introduction to the purification of DENV-2 RdRp ..... 87
3.1.3 Confirming the identity and activity of DENV-2 RdRp ..... 91
3.1.4 Determination of the crystal structure of DENV-2 RdRp and RdRp-RNA co-crystallisation studies ..... 93
3.2 Results ..... 103
3.2.1 DENV-2 RdRp expression, purification and activity ..... 103
3.2.2 X-ray crystallographic study of unbound DENV-2 RdRp .... 114
3.2.3 Co-crystallisation of DENV-2 RdRp with SLA RNA ..... 122
3.3 Discussion ..... 123
3.3.1 Expression and purification of DENV-2 RdRp ..... 123
3.3.2 The activity of DENV-2 RdRp ..... 123
3.3.3 The crystal structure of DENV-2 RdRp at 2.2 Å resolution and co-crystallisation studies ..... 124
**Chapter 4 *In vitro* SHAPE mapping of DENV-2 genomic RNA ..... 127**
4.1 Introduction ..... 128
4.2 Results ..... 135
4.2.1 *In vitro* SHAPE mapping in the absence/presence of *trans*-activating factors ..... 135
4.2.2 Identification of circular or linear structural genomic form .. 139
4.2.3 *In vitro* SHAPE mapping of DENV-2 RNA in the absence of *trans*-activating factors at 37 °C ..... 140
4.2.4 *In vitro* SHAPE mapping of DENV-2 RNA in the presence of DENV-2 RdRp at 37 °C ..... 144
4.2.5 *In vitro* SHAPE mapping of DENV-2 RNA in the presence of HEK293T replicon extract at 37 °C ..... 150
4.2.6 *In vitro* SHAPE mapping of DENV-2 RNA in the absence and presence of DENV-2 RdRp at 28 °C ..... 156

4.2.7 *In vitro* SHAPE mapping of DENV-2 genomic RNA results summary ..... 163

4.3 Discussion ..... 164

4.3.1 The linear DENV-2 genomic conformation is predominant in the absence of *trans*-activating factors at mammalian and insect physiological temperatures..... 164

4.3.2 The influence of DENV-RdRp binding on DENV RNA structure at mammalian and insect physiological temperatures..... 165

4.3.3 The influence of HEK293T DENV replicon protein extract on DENV RNA structure at mammalian physiological temperature 167

4.3.4 The limitations of *in vitro* SHAPE mapping and future directions 168

**Chapter 5 Intracellular SHAPE analysis of DENV-2 genomic RNA structure ..... 170**

5.1 Introduction..... 171

5.2 Results ..... 176

5.2.1 Optimisation of DENV-2 Intracellular SHAPE..... 176

5.2.2 Intracellular SHAPE conducted on stably expressing DENV HEK293T replicon cells ..... 186

5.2.3 Intracellular SHAPE conducted on DENV infected cells of mammalian origin ..... 190

5.2.4 Intracellular SHAPE conducted on DENV infected cells of insect origin ..... 193

5.2.5 Intracellular SHAPE mapping of fractionated stably expressing DENV HEK293T cells - enriched for active viral replication complexes ..... 197

5.2.6 Pan-serotype thermodynamic analysis of the DENV-UFS . 201

5.3 Discussion ..... 204

5.3.1 The linear genomic conformation and the DENV-2 UFS.... 204

5.3.2 Intracellular structural interrogation of DENV cHP ..... 205

5.3.3 Interrogation of DENV genomic RNA structure within the membrane fraction of stably expressing DENV HEK293T replicon cells ..... 206

5.3.4 Limitations of Intracellular SHAPE mapping..... 207

**Chapter 6 General discussion ..... 209**

**Bibliography ..... 221**

**Appendices..... 248**

Appendix I ..... 248

Buffers 248
Stains 249
Media 249
Appendix II ..... 250
Example *in vitro*/ intracellular thermodynamic RNA structural predictions 250

## Table of Figures

**Figure 1 The *Flaviviridae* family:..... 3**
**Figure 2 Phylogenetic tree for the genus Flavivirus:..... 5**
**Figure 3 Map of Southeast Asia: ..... 6**
**Figure 4 The emergence and transmission cycle of DENV:..... 7**
**Figure 5 The global emergence of DENV and co-circulation of serotypes from 1943 – 2013:..... 8**
**Figure 6 Maximum likelihood phylogenetic tree of whole genome coding regions of 58 DENV strains representing all four DENV serotypes:10**
**Figure 7 The worldwide distribution of DENV in 2016:..... 11**
**Figure 8 Bleeding manifestations upon the clinical presentation of DENV infection:..... 13**
**Figure 9 Crystal structure of the dengue virus envelope protein:..... 14**
**Figure 10 DENV as a model for ADE: ..... 15**
**Figure 11 The vectors of DENV infection and replication within mosquito host:..... 18**
**Figure 12 Overview of the cryo-EM structure of the DENV virion and virion particle schematic:..... 22**
**Figure 13 The DENV life-cycle: ..... 25**
**Figure 14 DENV genomic organisation:..... 26**
**Figure 15 DENV NS5 in the context of the DENV genome: ..... 31**
**Figure 16 Type I cap formation by the MTase domain:..... 32**
**Figure 17 Crystal structure of DENV-3 NS5:..... 33**
**Figure 18 DENV genome replication catalysed by the DENV RdRp:.... 35**
**Figure 19 Structural position/ conservation of *Flaviviridae* RdRp catalytic motifs:..... 37**

**Figure 20 Annotated crystal structure of the DENV-3 RdRp domain: .. 39**

**Figure 21 Predicted DENV-RdRp conformational transition during replication..... 41**

**Figure 22 A schematic of RNA secondary structural elements:..... 43**

**Figure 23 A schematic representation of DENV genome cyclisation:.. 47**

**Figure 24 The binding site of DENV RdRp on SLA RNA: ..... 49**

**Figure 25 A schematic of DENV replication: ..... 50**

**Figure 26 The four potential biologically active forms of DENV genomic RNA:..... 52**

**Figure 27 Differing conformations of the 5´ terminus of DENV genomic RNA:..... 55**

**Figure 28 A schematic describing the function of the UFS switch model – a proposed mechanistic model of Flavivirus replication:..... 58**

**Figure 29 Expression of DENV-2 RdRp utilising the pET expression system: ..... 87**

**Figure 30 Schematic of DENV-2 full-length NS5 and DENV-2 RdRp constructs: ..... 88**

**Figure 31 The *E. coli* expressed DENV-2 RdRp construct and nickel affinity chromatography principle:..... 89**

**Figure 32 Size-exclusion chromatography:..... 90**

**Figure 33 Principle of fluorescence anisotropy: ..... 92**

**Figure 34 Techniques and principles behind crystal growth:..... 95**

**Figure 35 Constructive and destructive interference: ..... 98**

**Figure 36 Bragg's law:..... 99**

**Figure 37 A schematic representation of the crystal lattice and relationship with the unit cell and asymmetric unit:..... 99**

**Figure 38 The structure factor equation: ..... 100**

**Figure 39 *E.coli* expression strain testing:..... 104**

**Figure 40 Optimisation of IPTG induction conditions for protein expression:..... 105**

**Figure 41 Nickel affinity purification of his-SUMO-DENV-2 RdRp:..... 106**

**Figure 42 The use of size exclusion chromatography in the purification of DENV-2 RdRp:..... 107**

**Figure 43 The identification of purified DENV-2 RdRp by western blot:108**

**Figure 44 Analysis of the purity of DENV-2 RdRp:..... 108**

**Figure 45 Circular Dichroism spectrum of monomeric DENV-2 RdRp:109**

**Figure 46 DENV-2 RdRp Polymerase activity assay:..... 110**

**Figure 47 The RNA binding capability of DENV-2 RdRp analysed by FA: ..... 111**

**Figure 48 Analysis of DENV-2 RdRp-RNA binding by native Electromobility Shift assay: ..... 113**

**Figure 49 Crystals grown in mother liquor droplets during JCSG sparse matrix screening: ..... 115**

**Figure 50 DENV-2 RdRp streak seeding: ..... 116**

**Figure 51 Loop containing a rod-shaped DENV-2 RdRp crystal at the synchrotron: ..... 117**

**Figure 52 Diffraction pattern of RNA unbound DENV-2 RdRp: ..... 118**

**Figure 53 DENV-2 RdRp crystal structure at 2.2 Å resolution: ..... 120**

**Figure 54 DENV-2 RdRp superposed onto the crystal structure of DENV-3 RdRp: ..... 121**

**Figure 55 RNA modification by NMIA: ..... 130**

**Figure 56 Schematic representation of NMIA treatment, primer extension, reverse transcription and fluorescent capillary electrophoresis: 131**

**Figure 57 A schematic diagram of the *in vitro* SHAPE procedure in the presence of *trans*-activating factors: ..... 133**

**Figure 58 QuSHAPE data analysis pipeline and RNA structural prediction: ..... 136**

**Figure 59 Analysis of RNA quality as templates for *in vitro* SHAPE mapping: ..... 137**

**Figure 60 Schematic of 5´500 nt and full-length DENV genomic transcripts used throughout *in vitro* SHAPE experimentation: .. 138**

**Figure 61 Diagnosis of the linear of circular genomic form for data analysis: ..... 139**

**Figure 62 *In vitro* SHAPE mapping of DENV-2 5´500 nt and full-length genomic RNA transcripts: ..... 143**

**Figure 63 *In vitro* SHAPE mapping of DENV-2 5´500 nt transcript in the absence and presence of DENV-2 RdRp: ..... 146**

**Figure 64 *In vitro* SHAPE mapping of full-length DENV-2 genomic RNA in the absence and presence of DENV-2 RdRp: ..... 149**

**Figure 65 *In vitro* SHAPE mapping of 5´500 nt DENV-2 genomic RNA in the absence and presence of stably expressing DENV HEK293T replicon cellular protein extract: ..... 152**

**Figure 66 *In vitro* SHAPE mapping of full-length DENV-2 genomic RNA in the absence and presence of HEK293T DENV replicon S100 protein extract: ..... 155**

**Figure *67 In vitro SHAPE mapping* of 5´ 500 nt DENV-2 genomic RNA in the absence and presence of DENV-2 RdRp at insect physiological temperature: ..... 159**

**Figure 68 *In vitro* SHAPE mapping of full-length DENV-2 genomic RNA in the absence and presence of DENV-2 RdRp at insect physiological temperature ..... 162**

**Figure 69 Intracellular SHAPE modification by NAI: ..... 173**

**Figure 70 The activity of NAI: ..... 176**

**Figure 71 Confirmation of the stable expression of DENV-2 non-structural proteins: ..... 177**

**Figure 72 Quality analysis of total RNA extracted from stably expressing DENV HEK293T cells: ..... 178**

**Figure 73 DENV-2 plaque assay - *pfu* and MOI calculations: ..... 179**

**Figure 74 Western blot analysis of DENV-2 infection time-course for optimisation of Intracellular SHAPE: ..... 180**

**Figure 75 Quality analysis of total RNA extracted from DENV infected BHK-21 and C6/36 cells: ..... 181**

**Figure 76 Confirmation of DENV infection and the stable expression of DENV-2 non-structural proteins in mammalian cells via RT-PCR and western blot: ..... 182**

**Figure 77 Confirmation of DENV infection of C6/36 cells via RT-PCR and western blot: ..... 183**

**Figure 78 Confirmation of the cellular fractionation of stably expressing DENV HEK293T replicon cells and analysis of total RNA extracted from the membrane fraction: ..... 184**

**Figure 79 The optimised Intracellular SHAPE protocol: ..... 185**

**Figure 80 *In vitro* vs intracellular SHAPE mapping normalised NMIA/NAI reactivities corresponding to full-length DENV replicon RNA at mammalian physiological temperature: ..... 189**

**Figure 81 *In vitro* vs intracellular SHAPE mapping normalised NMIA/NAI reactivities corresponding to full-length DENV RNA at mammalian physiological temperature: ..... 192**

**Figure 82 *In vitro vs* intracellular SHAPE mapping normalised NMIA/NAI reactivities corresponding to full-length DENV RNA at insect physiological temperature: ..... 196**

**Figure 83 Intracellular SHAPE mapping of total RNA extracted from stably expressing DENV HEK293T replicon cells vs Intracellular SHAPE mapping of total RNA extracted from the membrane fraction of stably expressing DENV HEK293T cells: ..... 200**

**Figure 84 Thermodynamic structural predictions of serotype specific DENV SLB at mammlian physiological temperature: ..... 202**

**Figure 85 Thermodynamic structural predictions of serotype specific DENV SLB at insect physiological temperature:..... 203**

**Figure 86 Proposed structural model of the transition of the DENV RdRp from the "closed" to the "open" conformation during the initiation of DENV replication: ..... 212**

**Figure 87 A schematic of *in vitro* SHAPE findings at both mammalian and insect physiological temperature and proposed RNA structural hypothesis:..... 216**

**Figure 88 Schematic representation of intracellular DENV RNA structure following total RNA extraction:..... 217**

**Figure 89 Schematic representation of DENV genomic RNA structure within the replication complex: ..... 218**

**Figure 90 Proposed DENV genomic conformation life-cycle model: . 219**

**Figure 91 *In vitro* SHAPE thermodynamically predicted RNA structure of full-length DENV-2 genomic RNA at mammalian physiological temperature in the absence of *trans*-activating factors ..... 250**

**Figure 92 Intracellular SHAPE thermodynamically predicted RNA structure of full-length DENV-2 genomic RNA extracted from the membrane fraction of stably expressing DENV-2 HEK293T cells at mammalian physiological temperature..... 250**

## Table of Tables

**Table 1 Oligonucleotides..... 63**

**Table 2 Primary and secondary antibodies for western blot ..... 64**

**Table 3 Data processing and scaling statistics of DENV-2 RdRp: ..... 119**

**Table 4 Summary of *in vitro* SHAPE mapping experimental results at mammalian physiological temperature..... 163**

**Table 5 Summary of *in vitro* SHAPE mapping experimental results at insect physiological temperature..... 163**

## Abbreviations

| | |
|---|---|
| ADE | Antibody-dependent enhancement |
| AFM | Atomic Force Microscopy |
| CD | Circular Dichroism |
| cHP | Capsid hair-pin |
| CS | Cyclisation sequence |
| Cyt | Cytoplasmic fraction |
| DAR | Downstream of AUG region |
| $ddH_20$ | Double distilled $H_20$ |
| DENV | Dengue Virus |
| DHF | Dengue haemorrhagic fever |
| DNA | Deoxyribonucleic acid |
| DSS | Dengue Shock Syndrome |
| DTT | Dithiothreitol |
| *E.coli* | Escherichia coli |
| EDTA | Ethylenediaminetetraacetic acid |
| EMSA | Electromobility Shift Assay |
| ER | Endoplasmic Reticulum |
| FA | Fluorescence Anisotropy |
| FBS | Foetal Bovine Serum |
| FL | Full-length |
| FLE | Fusion loop epitope |
| FPA | Fluorescence polarisation anisotropy |
| GFP | Green fluorescent protein |
| HCV | Hepatitis C Virus |
| IVT | *In vitro* transcription |
| JEV | Japanese encephalitis virus |
| MC | Methyl-cellulose |

| | |
|---|---|
| Min | Minutes |
| MOPS | 3-(n-Morpholino)Propanesulfonic Acid |
| MR | Molecular replacement |
| mRNA | Messenger RNA |
| NCR | Non-coding region |
| NEB | New England Biolabs |
| NS5 | Non-structural protein 5 |
| Nts | Nucleotides |
| Nuc | Nucleotide |
| OD | Optical density |
| ORF | Open reading frame |
| P/S | Penicillin-Streptomycin |
| PBS | Phosphate-Buffered Saline |
| PCR | Polymerase chain reaction |
| PDB | The Protein Data Bank |
| PLB | Passive lysis buffer |
| RNA | Ribonucleic acid |
| RT | Reverse transcription |
| RT-PCR | Reverse transcription – Polymerase chain reaction |
| SHAPE | Selective 2'-hydroxyl acylation analyzed by primer extension |
| SLA | Stem-loop A |
| SLB | Stem-loop B |
| SOC | Super Optimal broth with Catabolite repression |
| TBE | Tris-borate-EDTA |
| TCEP | Tris(2-carboxyethyl)phosphine |
| TEMED | Tetramethylethylenediamine |
| UAR | Upstream of AUG region |
| UFS | 5´ UAR flanking stem |

| | |
|---|---|
| UTR | un-translated region |
| WHO | World Health Organisation |
| WNV | West-Nile virus |
| YFV | Yellow-fever virus |

# Chapter 1 Introduction

## 1.1 General introduction

Flaviviruses such as dengue virus (DENV) represent an important public health concern whereby 40% of the global population are at risk of infection (Ye *et al.*, 2015). Without direct acting anti-virals or a widely accessible vaccine, the DENV life-cycle demands interrogation to broaden our knowledge at the molecular level - identifying critical anti-viral targets. For example, DENV replication provides one such target. However, critical protein/RNA interactions during this fundamental life-cycle stage remain understudied. Therefore, interrogating the replication cycle of DENV, focusing on protein and RNA structure during replication would increase our knowledge of viral replication across the Flavivirus genus and act to elucidate critical anti-viral targets.

### 1.1.1 The *Flaviviridae* family, Flaviviruses and the emergence of DENV

The *Flaviviridae* are a family of positive sense, single-stranded, enveloped RNA viruses with fundamental differences in host-range and transmissibility. The name *Flaviviridae* is derived from Yellow Fever virus (YFV). *Flavus* is the latin word for yellow and eludes to the jaundice pathology of YFV, the prototype virus of this family.

The *Flaviviridae* family includes four genera – Flavivirus, Pestivirus, Hepacivirus and Pegivirus (Fig. 1). Major diseases caused by the *Flaviviridae* family include dengue fever/ shock syndrome, Hepatitis C virus (HCV) associated complications such as liver disease and hepatocellular carcinoma, West Nile virus (WNV) associated encephalitis, Yellow fever and congenital Zika syndrome and therefore contribute a substantial risk to public health.

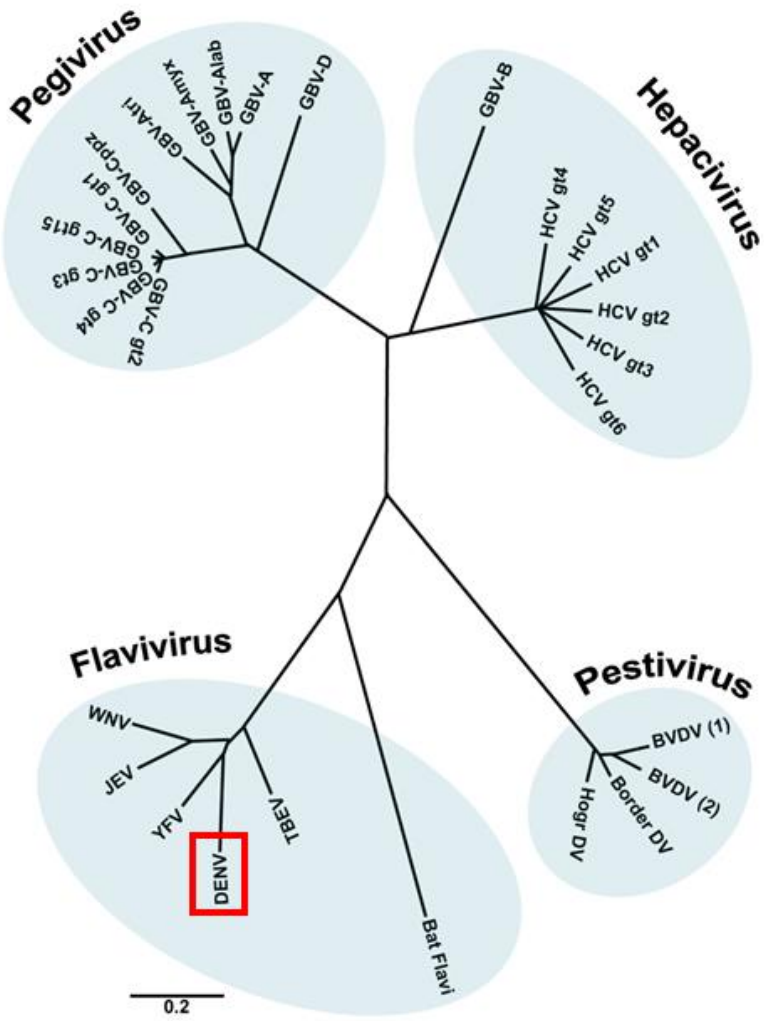



**Figure 1 The *Flaviviridae* family:**

*Phylogentic tree based on aligned amino acid conserved motifs of the RdRp. Distance scale shown corresponding to amino acid substitutions per position. The Flavivirus DENV, the subject of this thesis, is highlighted in red. Adapted from* (Romero-Brey and Bartenschlager, 2014)

The genus Flavivirus comprises over 70 viruses with members, such as DENV and Japanese Encephalitis virus (JEV), representing important human pathogens. Despite similar genomic organisation the Flaviviruses possess fundamental differences in their host range and transmissibility. Most Flaviviruses, such as DENV, are transmitted horizontally between hematophagous arthropods and vertebrate hosts and are therefore dual-host viruses. Dual host viruses can be further divided into mosquito/vertebrate and tick/vertebrate viruses. DENV is one such mosquito/vertebrate virus. Tick/vertebrate Flaviviruses associated with serious human disease include Tick-borne encephalitis virus and Powassan virus (Blitvich and Firth, 2015) (Fig. 2).

In contrast, not all Flaviviruses cycle between arthropods and vertebrates. Vertebrate specific Flaviviruses have been identified termed No Known Vector (NKV) Flaviviruses. NKV viruses can be further divided into two groups: those isolated exclusively from rodents, for example Modoc virus, and those isolated from bats, for example Rio Bravo virus, which are maintained in nature by horizontal transmission among hosts (Fairbrother and Yuill, 1987). Additionally, Insect-specific Flaviviruses (ISFs) can be divided into two groups: classical ISFs (cISFs) and dual-host affiliated ISFs (dISFs). cISFs named due to the fact that they were discovered first and dISFs named due to phylogenetic affiliation with mosquito/vertebrate Flaviviruses. Historically, ISFs have generated limited interest due to their inability to infect vertebrate cells. However, invertebrate infection with ISFs has been shown to enhance or suppress the replication of medically important Flaviviruses in co-infected mosquito cells. For example, the cISF Palm-creek virus modulates WNV infection and transmission by mosquitoes (Hall-Mendelin *et al.*, 2016). Consequently, there has been a dramatic increase in the number of ISFs identified over the last decade due to their apparent importance in the replication of medically important Flaviviruses (Blitvich and Firth, 2015) (Fig. 2).

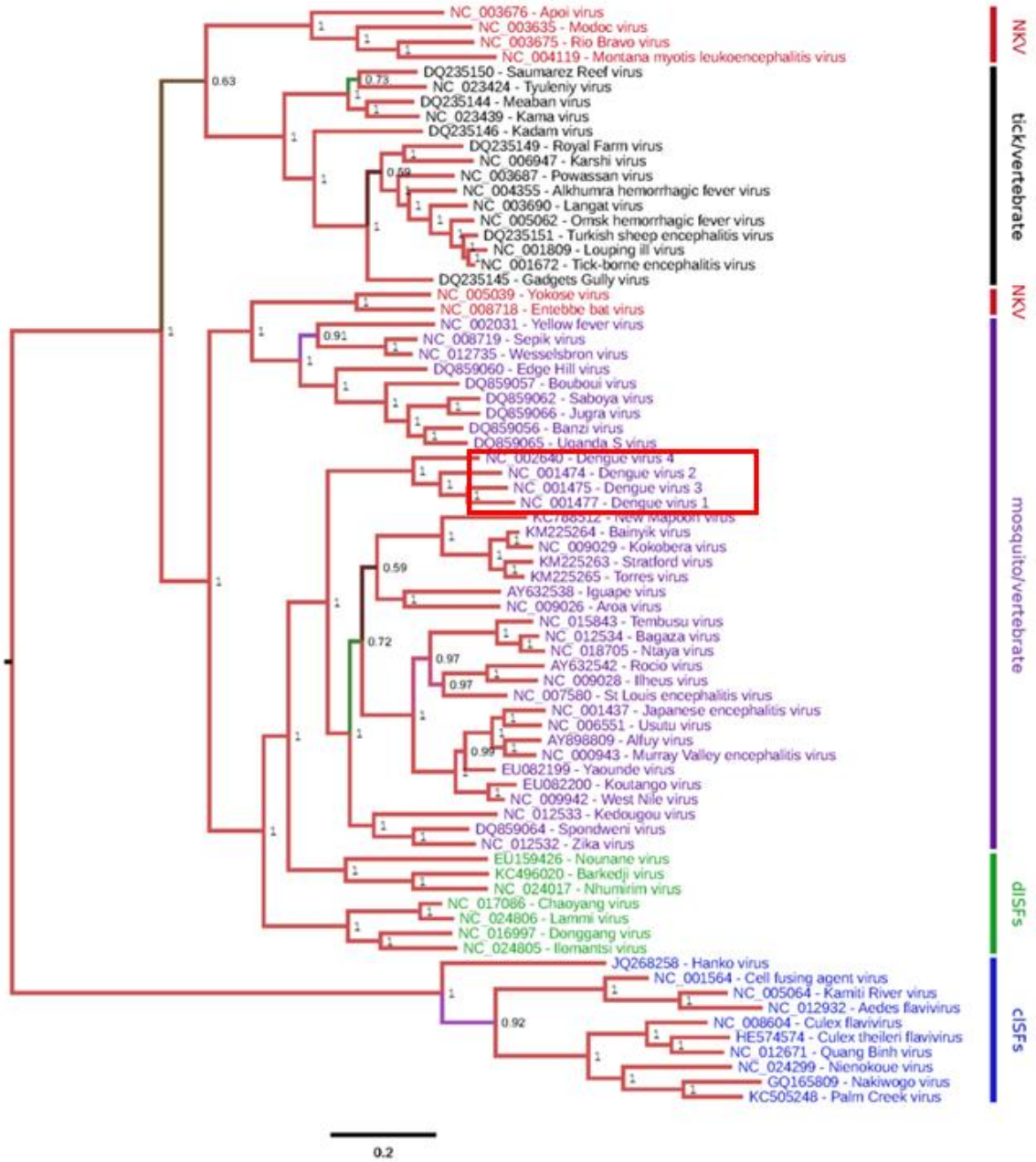



**Figure 2 Phylogenetic tree for the genus Flavivirus:**

*Complete polyprotein amino acid sequences were aligned using MUSCLE. Maximum likelihood phylogenetic tree. The tree is mid-point rooted, nodes are labelled with posterior probability values and branches are highlighted with alternative colours. DENV serotypes 1-4 are boxed in red. Species are colour coded: Classical Insect-specific Flaviviruses (cISFs) – blue; Dual-host affiliated Insect-specific Flaviviruses (dISFs) – green; No Known Vector Flaviviruses (NKV) – red; mosquito/vertebrate Flaviviruses – purple; tick/vertebrate Flaviviruses – black. Adapted from* (Blitvich and Firth, 2015)*.*

Globally, DENV is the most common arboviral disease, with an approximate 40% of the world's population at risk of viral infection (da Silva-Voorham *et al.*, 2009; CDC, 2018; WHO, 2018b). The origin of DENV in humans is unclear but it is proposed to have originated in Southeast Asian forests when humans first came into contact with the sylvatic-mosquito DENV transmission cycle (Gubler, 1998) (Figs. 3 and 4).

**Figure 3 Map of Southeast Asia:**

*Map of Southeast Asia. Humans are predicted to have come into contact with the sylvatic-mosquito transmission cycle of DENV in the jungles/forests of Southeast Asia annotated in green on the map. Taken from* (Greenpeace, 2018).

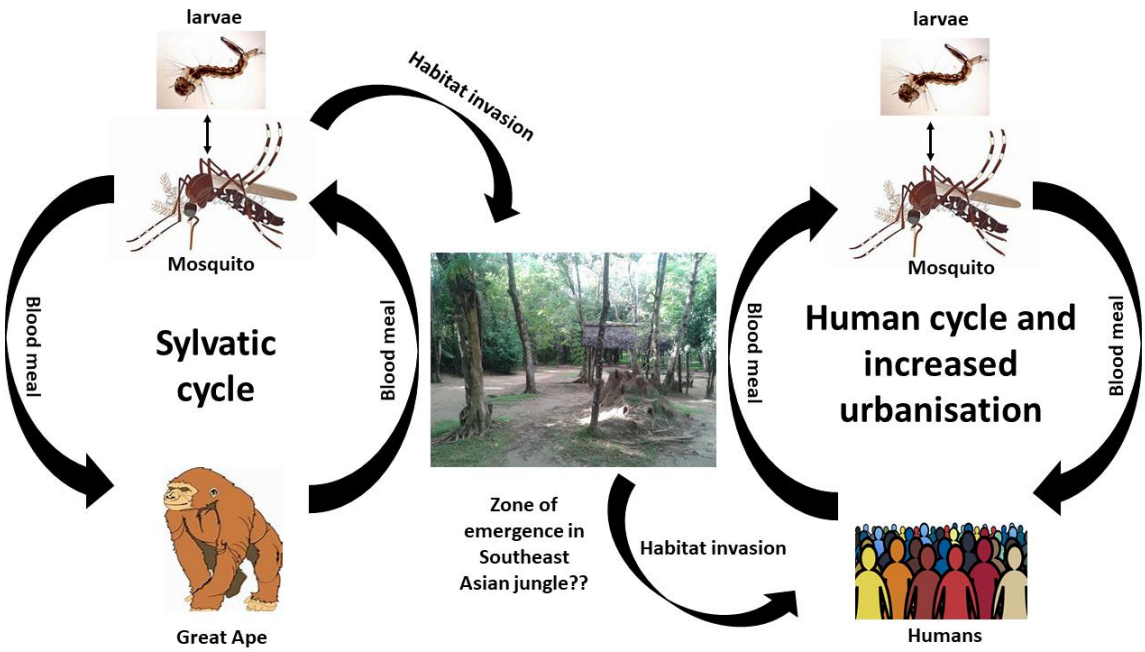



**Figure 4 The emergence and transmission cycle of DENV:**
*The emergence from sylvatic origin and transmission cycle of DENV.*

The first known records of dengue-like illness date back to 992 A.D, during the Northern Sung Dynasty of China, and was first described in an encyclopaedia of disease symptoms and remedies (Gubler and Clark, 1995; Gubler, 1998). The disease was termed “water poison” and was thought to be associated with flying insects residing in water habitats. Outbreaks of illness in the French West Indies in 1635 and in Panama in 1699 were also recorded to have presented dengue-like illness (Gilles, 1978; McSherry, 1982; Gubler, 1998; Brathwaite Dick *et al.*, 2012). It is predicted that DENV had a large geographic distribution prior to the 18$^{th}$ Century when the first truly known pandemic of dengue-like illness commenced.

The first epidemic of DENV occurred between 1827 and 1828 whereby the Virgin Islands, Jamaica, Cuba, USA, Venezuela and Mexico were affected, diagnosed on the basis of clinical presentation. During the latter half the 18$^{th}$ Century, DENV epidemics spread throughout Brazil, South America and the Caribbean and continued into the first half of the 19$^{th}$ Century (Gubler, 1998; Gubler & Clark, 1995).

During World War II (1939 – 45) and as a consequence of the movement of soldiers from England, the United States, Australia and Japan, DENV spread to

new areas, for example Australia and Hawaii, due to the soldiers' incursion into the sylvatic-mosquito cycle during the infiltration of Southeast Asian forests (Gibbons *et al.*, 2012). The spread of DENV during this time resulted in the co-circulation of multiple DENV serotypes within the population which had been previously separated and was associated with the clinical presentation of more serious DENV pathologies (further discussed in section 1.1.3) (Guzman and Alvarez, 2013). Consequently, the first epidemic of dengue shock syndrome/severe dengue was reported in the Philippines in 1953 which continued to spread throughout Southeast Asia over the next 20 years (Halstead, 2008). Subsequently, vector regulation initiatives, such as the use of mosquito bed nets and pesticides, controlled the transmission of DENV during the 1960-70s (Halstead, 2008). However, in recent years the number of DENV cases reported to the World Health Organisation (WHO) has risen dramatically and the transmission of DENV remains a significant threat to public health (Fig. 5).

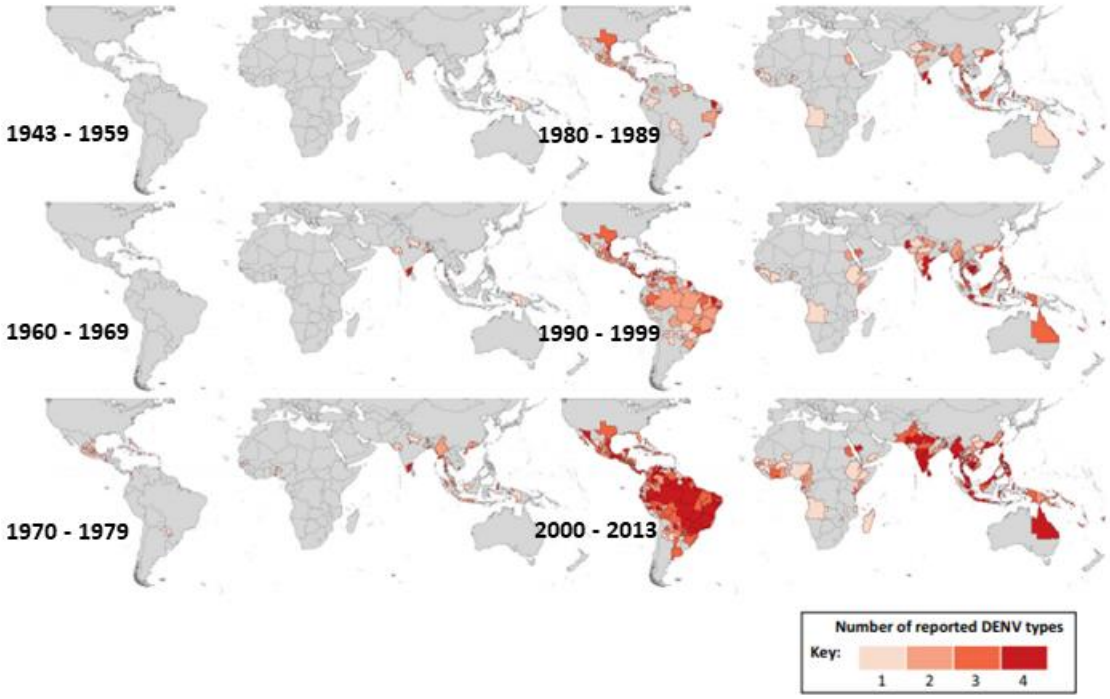



**Figure 5 The global emergence of DENV and co-circulation of serotypes from 1943 – 2013:**

*The cumulative number of DENV serotypes reported by decade since 1943. Taken from* (Messina *et al.*, 2014).

### 1.1.2 Epidemiology

There are four phylogenetically and antigenically distinct DENV serotypes numbered 1-4, with the possible identification of a fifth (Mustafa *et al.*, 2015), which co-circulate within the population as a multitude of DENV genotypes (Fig. 6). The ancestor of these viruses is thought to have evolved in an infectious cycle involving non-human primates and mosquitoes with subsequent transmission to humans having occurred independently for all separate virus types (discussed in section 1.1.1).

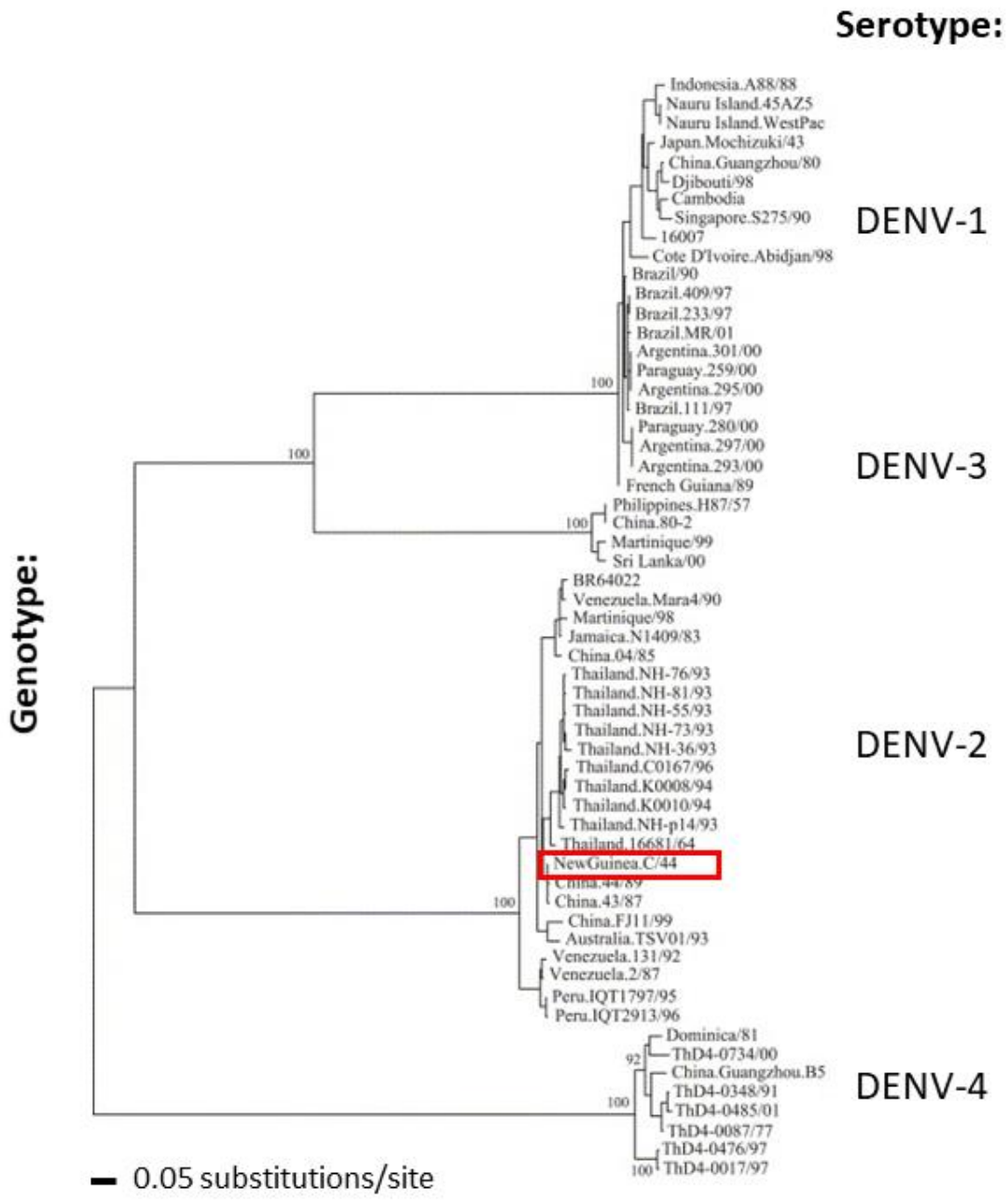



**Figure 6 Maximum likelihood phylogenetic tree of whole genome coding regions of 58 DENV strains representing all four DENV serotypes:**

*Horizontal branch lengths are proportional to the bar representing the number of nucleotide substitutions per site. The Southeast Asian DENV-2 New Guinea C genotype utilised during our experimentation is highlighted in red. Adapted from* (Klungthong *et al.*, 2004).

Before 1970, only 9 countries had experienced severe DENV epidemics. In contrast, by 2017 the disease was endemic in over 100 countries with Southeast Asia and Western Pacific regions baring the highest disease burden (Fig. 7) (Nealon *et al.*, 2016; WHO, 2018a). The annual number of dengue cases is predicted to be underreported and, in many cases, misclassified due to DENV infection commonly resembling symptoms of viruses such as Influenza virus or Chikungunya virus infections. However, the number of cases reported to the WHO has risen from 2.2 million in 2010 to 3.2 million in 2015 (WHO, 2017). It is estimated that 390 million DENV infections occur per annum of which only 96 million manifest clinically with any severity (Bhatt *et al.*, 2013; WHO, 2017) (pathogenesis further discussed in section 1.1.3).

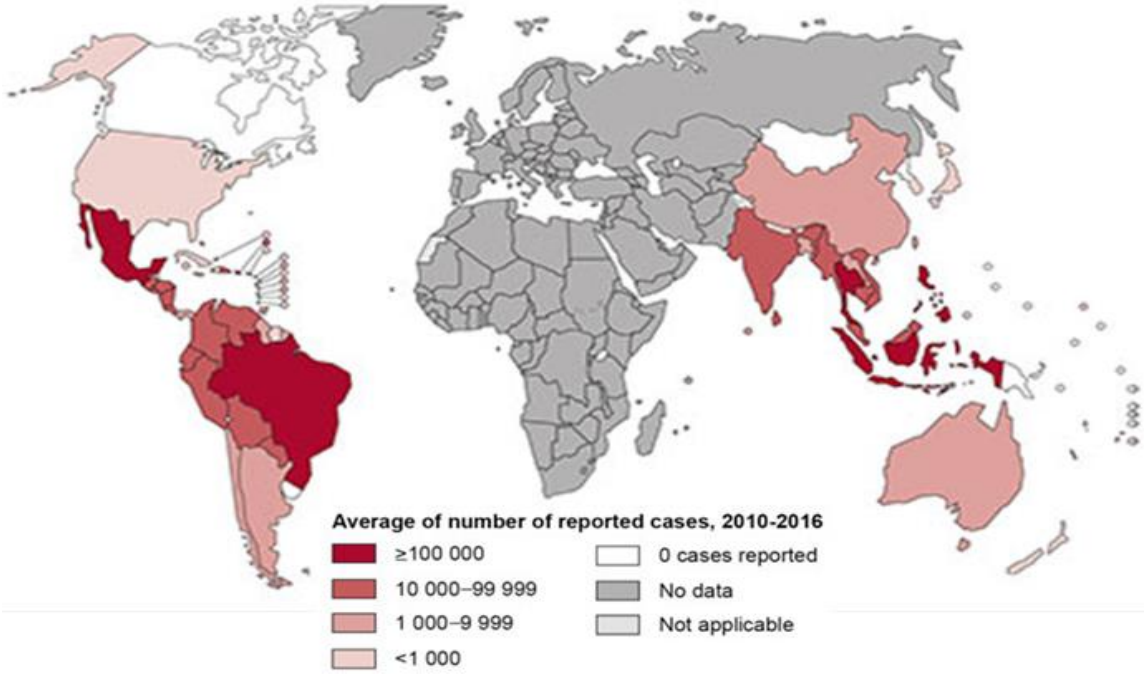



**Figure 7 The worldwide distribution of DENV in 2016:**

*The average number of suspected or confirmed DENV cases reported to WHO between 2010 and 2016. Adapted from* (WHO, 2018a)*.*

Importantly, the number of cases of DENV are increasing as the disease spreads geographically due to such factors as increased vector spread as a result of global warming and urbanisation. The threat of a possible outbreak of DENV now exists in Europe since local transmission was reported for the first time in 2010 in both France and Croatia (La Ruche *et al.*, 2010; Gjenero-Margan *et al.*, 2011; Semenza

*et al.*, 2014). In 2012 Portugal saw an outbreak of DENV on the Madeira Islands which culminated in over 2000 cases (Sousa *et al.*, 2012; Semenza *et al.*, 2014; WHO, 2017).

### 1.1.3 The pathogenesis of DENV infection

The true origin of the name dengue is unknown but is thought to be derived from the Swahili “ka-dinga pepo” which translates as “disease of the devil”. Furthermore, “dengue” is a Spanish word which describes the “fastidious and careful” gait exhibited by infected persons as they walked.

Infection with any of the DENV serotypes is, in the majority of cases, asymptomatic or conversely may result in a wide-spectrum of clinical symptoms (Harris *et al.*, 2000) (Fig. 8). The reported clinical symptoms usually occur after an approximate incubation period of 8-12 days and range from a mild-flu like illness and characteristic rash clinically diagnosed as dengue fever, to the most severe forms of the disease known respectively as dengue haemorrhagic fever (DHF) and dengue Shock Syndrome (DSS) (Chan and Johansson, 2012). DSS, the most serious pathogenesis of DENV, is most commonly associated with a second DENV infection of a different serotype predicted to be as a result of antibody-dependent enhancement (ADE) (discussed in more detail in section 1.1.3.1) (Martina, Koraka and Osterhaus, 2009).

DHF is characterised by coagulopathy, increased vascular fragility and permeability. DSS is characterised by hypovolemic shock as a result of sudden blood or fluid loss culminating in organ failure and death if uncontrolled (Martina, Koraka and Osterhaus, 2009). An estimated 500,000 people with severe dengue require hospitalisation each year associated with a 2.5% mortality rate (Zavala-Castro *et al.*, 2014; WHO, 2017).

Infection with different serotypes/genotypes, and the order of infection, with DENV are associated with different severities of disease. For example, a cross-sectional study on 485 confirmed dengue cases in Brazil found that cases of DENV-2 infection resulted in a higher proportion of severe dengue than among those of DENV-1 and 4 (Vicente *et al.*, 2016). Certain genotypes within different serotypes are associated with more severe disease outcomes. DENV-2 viruses of Asian

origin are capable of more robust replication *in vitro* relative to the less virulent American DENV-2 genotypes and consequently DENV-2 viruses of Asian origin are more frequently associated with severe dengue (OhAinle *et al.*, 2011). Additionally, the serotype of primary infection and the sequential serotype of a secondary infection may be associated with more severe disease outcomes. For example, a primary infection with DENV-1 and then subsequent second infection with either DENV-3 or DENV-4 was associated with DSS/DHF in Cuba (Alvarez *et al.*, 2006; Guzman and Alvarez, 2013). Early detection and diagnosis of DENV serotypes/genotypes circulating within the affected territory could be an important approach to prevent severe outcomes during dengue outbreaks.

**Figure 8 Bleeding manifestations upon the clinical presentation of DENV infection:** *Clinical presentations of DENV infection. DENV rash associated with dengue fever (left), Ecchymosis (bleeding under the skin) associated with DHF (top right) and subconjunctival haemorrhage associated with severe dengue (bottom right). Adapted from* (Rose *et al.*, 2017).

#### 1.1.3.1 Antibody-Dependent enhancement

The DENVs are a group of four closely related members of the Flavivirus genus. DENV serotypes 1-4 are antigenically distinct from each other, sharing approximately 60 – 70% nucleotide homology. Initial infections with any of the four serotypes raise a protective serotype specific immune response, characterised by the production of serotype specific antibodies. Despite conferring resistance to the serotype they are raised against, most of the antibodies produced are cross-

reactive and non-neutralising. Neutralising antibodies, however, are suggested to target a segment of the envelope protein located between residues 298 and 397 within domain III (Fig. 9) (Megret *et al.*, 1992). In 2- 4% of cases, upon a secondary DENV infection of a heterotypic serotype, the pre-existing antibodies act to enhance the infection. The DENV secondary infection is opsonised and targeted for uptake into Fc-receptor baring cells such as monocytes and macrophages, primary sites of DENV replication (Halstead, 2003, 2014) (Fig. 9).

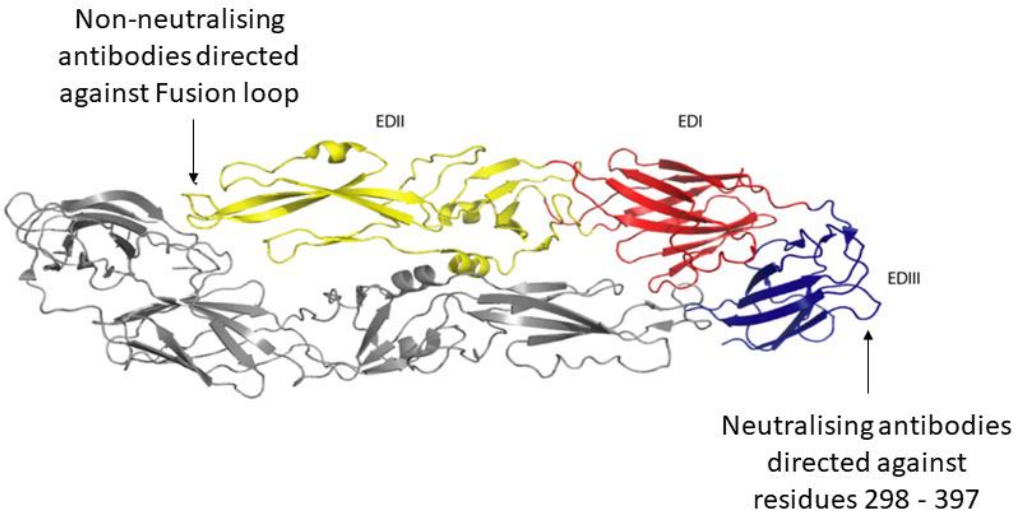



**Figure 9 Crystal structure of the dengue virus envelope protein:**

*The envelope protein on the mature virion is a homodimer (dimer not shown). Each subunit has three domains designated EDI (domain I) (red), EDII (domain II) (yellow) and EDIII (domain III) (blue). The position of the fusion loop is indicated, the immunodominant epitope of non-neutralising antibodies. Neutralising antibodies were found to be directed again residues 298 – 397 in Domain III as indicated. Adapted from* (Wahala and de Silva, 2011).

The immunodominant epitope of non-neutralising antibodies is predicted to be the Fusion Loop Epitope (FLE) of the DENV envelope protein, a linear epitope spanning the fusion loop, which is highly conserved across the DENV serotypes and across the Flavivirus genera (Fig. 9) (Oliphant *et al.*, 2007; Beltramello *et al.*, 2010; Dejnirattisai *et al.*, 2015; Rouvinski *et al.*, 2017).

FLE exposure takes place via the dynamic 'breathing' movement of envelope dimers at the surface of the virion (Rouvinski *et al.*, 2017). Since the FLE epitope is poorly presented by mature DENVs, anti-FLE antibodies often show a poor

neutralisation capability but are able to potently induce ADE (Oliphant *et al.*, 2007; Balsitis *et al.*, 2010; Beltramello *et al.*, 2010; Rodenhuis-Zybert *et al.*, 2011; Dejnirattisai *et al.*, 2015). ADE is therefore an important blockade to DENV vaccine development (discussed further in section 1.1.4.1). In this immunopathological phenomenon, the infection of permissive cells, such as macrophages, using infectious immune complexes, acts to suppress innate anti-viral systems and permits the intracellular growth of DENV. Therefore, due to the increase in viral load, ADE is commonly the cause of the most serious pathology of DENV, DSS (Fig. 10).

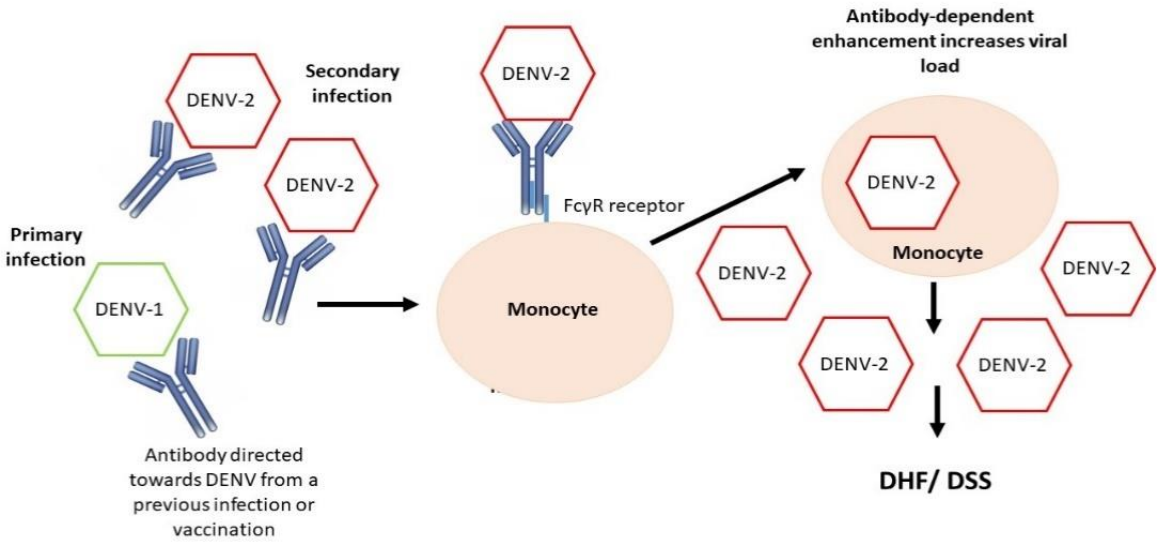



**Figure 10 DENV as a model for ADE:**

*DENV as a model for ADE. A primary DENV infection confers resistance to the same serotype. Following a second infection with a different DENV serotype the antibodies directed towards the DENV serotype of the primary infection are non-neutralising and enhance the infection of the second serotype. The virus is opsonised and targeted for uptake into Fc receptor baring cells. ADE is associated with a more severe clinical presentation of the disease due to increased viral load.*

Critically, it is now known that pre-existing antibodies directed against DENV enhance Zika virus infection *in vivo* and may increase disease severity (Paul *et al.*, 2016). Consequently, a clear understanding of the interplay between Flavivirus infections and the phenomenon of ADE will be critical in the orchestration of public health in regions where this family of viruses co-circulate.

### 1.1.4 Treatment

Direct acting anti-virals have not been licensed for the treatment of DENV infection. In most cases the infection is asymptomatic, however upon clinical presentation, treatment involves alleviating and controlling the symptoms. In non-severe cases painkillers, such as Paracetamol, are used to control the pain and relieve the fever. The intake of fluids is increased to prevent dehydration and bed rest is required until the infection clears. If the patient presents with critical DHF or DSS, the complications are fastidiously monitored.

Uncontrolled plasma leakage leads to shock. Studies reveal a reduction in plasma volume in more than 20% of severe cases (Soni *et al.*, 2001; WHO, 2018c), consequently early and effective replacement of plasma loss is required. Plasma expanders (a substance of high molecular weight which, on infusion, act to increase the volume of circulating fluid and increase osmotic pressure) may be used alongside increased fluid/electrolyte solutions. Thirst and dehydration can result from high fever and vomiting and therefore an increase in fluid and electrolyte intake may be required.

#### 1.1.4.1 Vaccination

Due to the lack of licensed direct acting anti-virals, a safe, effective and affordable vaccine directed against all four serotypes of DENV would represent a major advance in the control of the disease in endemic areas.

Dengvaxia® (CYD-TDV), developed by Sanofi Pasteur, was licensed for use in 2015 (Guy *et al.*, 2017). Approximately, five other DENV vaccine candidates are in clinical development. One such candidate developed by Takeda pharmaceutical company limited began phase III clinical trials in 2016 with TAK-003 in the United States and dengue endemic countries (Moi, Takasaki and Kurane, 2016). TAK-003 is a needle-free, tetravalent live-attenuated vaccine. The vaccine employs a DENV-DENV chimera (as opposed to the YFV-DENV chimera employed by Dengvaxia®) where the attenuated strain DENV-2 PDK-53 provides protection against DENV-2, alongside three chimeric viruses expressing the prM and E proteins of DENV-1,3 and 4. The inclusion of DENV-2 non-structural genes with TAK-003 is predicted to be a major differentiator between TAK-003 and

Dengvaxia® and may result in greater protection against DENV-2 strains, often associated with greater disease severity (Moi, Takasaki and Kurane, 2016).

Dengvaxia® is the first DENV vaccine to be licensed and was done so in Mexico in 2015 for use in persons living in endemic areas between 9-45 years of age. Dengvaxia® is a prophylactic, live, recombinant (YFV attenuated) tetravalent vaccine with a YFV backbone, administered in a 3-dose series on a 0/6/12-month schedule (Moi, Takasaki and Kurane, 2016). Critically, the vaccine is tetravalent and therefore aims to provide immunity towards all four DENV serotypes with the goal of preventing ADE and therefore protecting against the serious pathologies of DENV.

WHO recommends that countries should consider introduction of the vaccine only in geographical settings where epidemiological data indicate both a persistent and high intensity of DENV transmission. Additionally, the seroprevalence should be 70% or greater, in the age-group targeted for vaccination, maximising public health impact and cost-effectiveness. This recommendation was bought about due to the differential performance of Dengvaxia®, in terms of vaccine efficacy and possible safety concerns (WHO, 2016).

In December 2017, the Philippines halted its child vaccination campaign and recalled Dengvaxia® amid safety fears (Dyer, 2017). Vaccine efficacy for confirmed DENV cases was found to be lower in seronegative individuals than in individuals who were seropositive at baseline in the trials, therefore suggestive of a degree of ADE following immunisation (Aguiar, Stollenwerk and Halstead, 2018). The vaccine shows serotype specific efficacy with less protection against DENV serotype 2 (Sim and Hibberd, 2016; Capeding *et al.*, 2018; Sabchareon *et al.*, 2018). Consequently, prophylactic patient vaccination in geographically endemic areas is not without risk and DENV remains a significant public-health concern.

### 1.1.5 The mosquito vector and control strategies

Alongside vaccination and treatment strategies, preventing or reducing DENV transmission depends on controlling the mosquito vector population, *Aedes aegypti* and *albopictus* within which DENV replicates (Fig. 11), or interruption of human-vector contact. Aside from the use of mosquito repellents, insecticides and

bed nets, much effort has been undertaken in the genetic modification of the mosquito vector to control disease transmission.

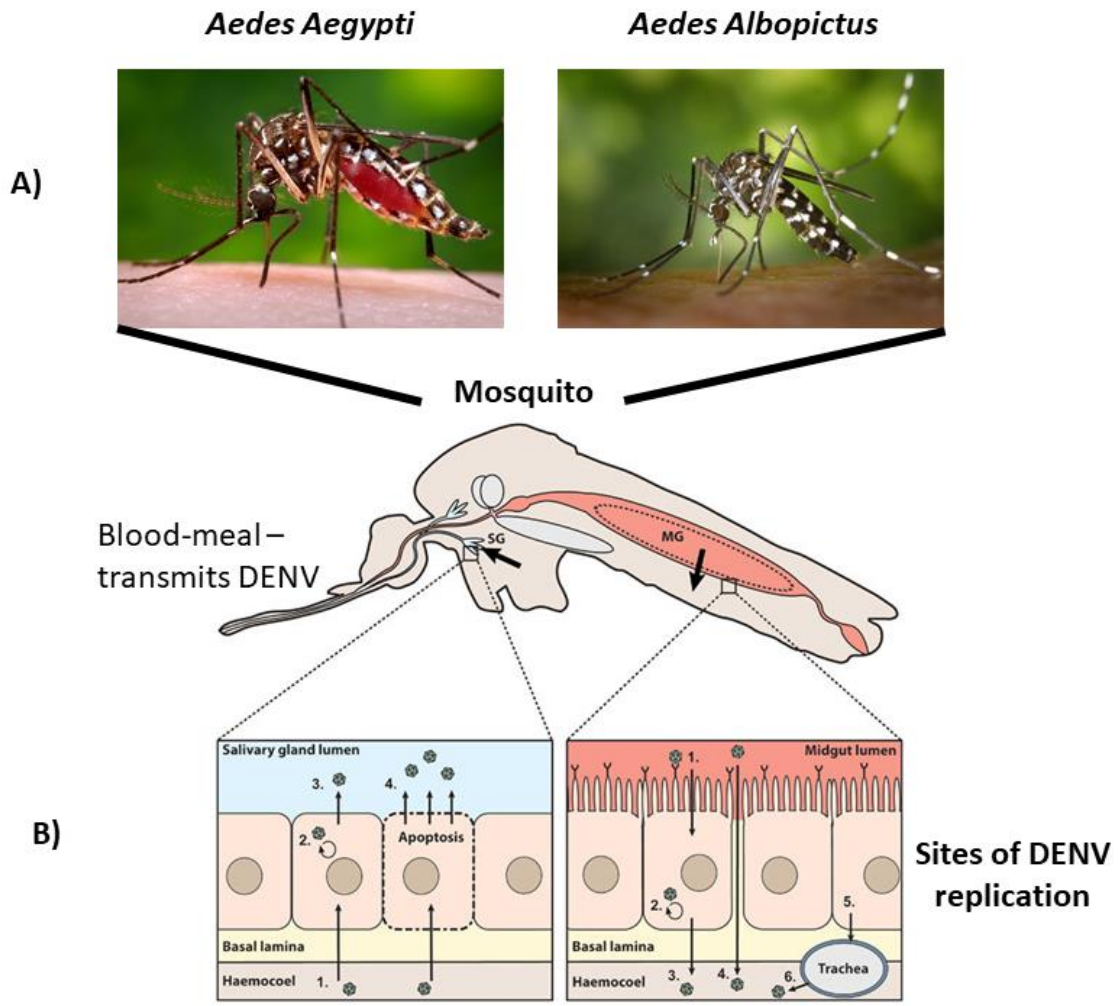



**Figure 11 The vectors of DENV infection and replication within mosquito host:**

***A)*** *Images courtesy of CDC. Image ID: 2165 and 9261.* ***B)*** *Schematic overview DENV replication sites within a mosquito vector. A longitudinal cross-section of a mosquito is displayed. The arrows indicate the passage of virions through both the midgut (MG) and the salivary gland. The circle within the MG represents the peritrophic membrane formed after taking a blood-meal. Midgut lumen –* ***1)*** *infection of midgut epithelial cells via envelope receptor binding.* ***2)*** *DENV replication within MG epithelial cells.* ***3)*** *Release of virus via budding from MG epithelial cells and passage through the basal lamina into the haemocoel (a haemolymph filled cavity of invertebrates).* ***4)*** *viral passage into the haemocoel via a 'leaky' MG.* ***5)*** *Viral infection of the trachea following budding from MG epithelial cells.* ***6)*** *Viral budding from the trachea into the haemocoel. Salivary gland lumen -* ***1)*** *infection of the SG epithelial cells following passage through the basal lamina.* ***2)*** *Viral replication in SG cells.* ***3)*** *Virus release following budding from SG cells into the SG lumen.* ***4)*** *Viral release from the SG cells into the SG lumen by apoptosis. Adapted from* (Vogels *et al.*, 2017).

Oxitec have produced a self-limiting strain of male *Ae.aegypti* mosquito termed OX13A (Oxitec, 2018a). Males which do not bite and transmit disease, carry the modification and are released to mate with females in endemic areas. The offspring die before reaching adulthood and cannot mature to transmit disease. The introduction of this technology was reported to reduce the wild mosquito population by 90-95% (Carvalho *et al.*, 2015; Oxitec, 2018a). OX513A has regulatory approvals for import and controlled testing in disease endemic areas such as Brazil, the Cayman Islands, USA and France (Oxitec, 2018a).

The OX13A vectors contain the insect specific, repressible, self-limiting gene tTAV (tetracycline trans-activator variant), expression of which, if unrepressed, is lethal through possible downstream interactions with transcription factors. Expression of tTAV is under the control of its own binding site, tetO. In the presence of tetracycline, tTAV binds tetracycline and cannot bind tetO. Consequently, the expression of tTAV remains at a basal non-lethal level. In the absence of tetracycline, tTAV binds tetO which drives the expression of tTAV in a positive feedback loop which is lethal. Male mosquitoes are reared in the presence of tetracycline prior to release. Following release, modified males mate with the wild population, where tetracycline is inaccessible. This action culminates in mosquito population reduction and a consequential decrease in disease transmission (Oxitec, 2018b).

The intracellular bacterium *Wolbachia pipientis* may be utilised to control DENV transmission (McGraw and O'Neill, 2013). *Wolbachia* is maternally transmitted and naturally found in approximately 40% of all insect species (Zug, Koehncke and Hammerstein, 2012). *Ae. agepti* do not carry *Wolbachia* naturally but have been shown to be stably *trans*-infected with the bacterium resulting in an endosymbiotic anti-viral effect (Xi, Khoo and Dobson, 2005; Walker *et al.*, 2011; Zug, Koehncke and Hammerstein, 2012). The mechanisms underlying Wolbachia's anti-viral properties remain un-clear. Wolbachia infection may prime the host innate immune system, manipulating the microRNA pathway and compete for host-resources critical to the establishment of a persistent DENV infection (Bian *et al.*, 2010; Caragata *et al.*, 2013; G. Zhang *et al.*, 2013; Ye *et al.*, 2015).

No single intervention, vaccination or vector control, will be sufficient to control the transmission of DENV (Achee *et al.*, 2015). Efforts must be effective and complementary whilst research is undertaken to treat the disease via the development of specific direct-acting anti-virals towards DENV infection.

### 1.1.6 Animal models of DENV infection

The development of animal models of DENV infection has not been without significant challenge and is critical to the efficacy testing of potential vaccine/ direct acting anti-viral candidates. Clinical isolates of DENV show low-level or lack of replication in wild-type mice and fail to present clinically in non-human primates.

Initial models in mice used intracranial injections of neurovirulent DENV into suckling and adult immunocompetent mice which culminated in neurological disease and paralysis, not representative of human infection (Raut *et al.*, 1996). Peripheral replication of DENV in mice was first reported by both intravenous and sub-cutaneous injection using a mouse model deficient for both IFN-α/β and γ receptors (J. Johnson and Roehrig, 1999). Intra-peritoneally administered mouse adapted DENV-2 was uniformly lethal in this model. However, immunised mice were protected from challenge with the same virus therefore demonstrating the promise of this model for use in both vaccine and anti-viral trials (J. Johnson and Roehrig, 1999). The development of mouse-adapted DENV strains has improved this model, allowing for the interrogation of DENV pathogenesis, the testing of therapeutic antibodies and anti-viral compounds (Shresta *et al.*, 2006; Yin *et al.*, 2009; Balsitis *et al.*, 2010; Zellweger, Prestwood and Shresta, 2010). The engraftment of human hematopoietic progenitors in immunodeficient mice allowed for the development of humanised mice. Humanised mice have been found to facilitate infection with clinical DENV isolates producing a dengue-like illness (Bente and Rico-Hesse, 2006; Kuruvilla *et al.*, 2007; Mota and Rico-Hesse, 2009; Zompi and Harris, 2012).

Non-human primate models do not present DENV disease clinically, but the same cells, for example monocytes and macrophages, are infected. Therefore, non-human primate models are useful for interrogating the immune response mounted

towards DENV infection and for testing vaccine efficacy in pre-clinical trials (Zompi and Harris, 2012).

## 1.2 DENV

### 1.2.1 Virion structure

Initial electron micrographs showed that DENV virions are characterised by a relatively smooth surface with a diameter of approximately 50 nm (Kuhn *et al.*, 2002). Furthering our knowledge, Cryo-electron microscopy (Cryo-EM) has been utilised to interrogate the virion structure and has revealed that in addition to the 10.7 Kb single-stranded positive-sense RNA genome, there are three structural proteins which occur in the particle and are arranged in icosahedral-like symmetry: capsid (C), membrane (M) and envelope (E). The lipid envelope contains the viral membrane proteins, M and E alongside lipids derived from host cell membranes (Kuhn *et al.*, 2002; X. Zhang *et al.*, 2013; Sztuba-Solinska and Le Grice, 2014; Reddy and Sansom, 2018) (Fig. 12).

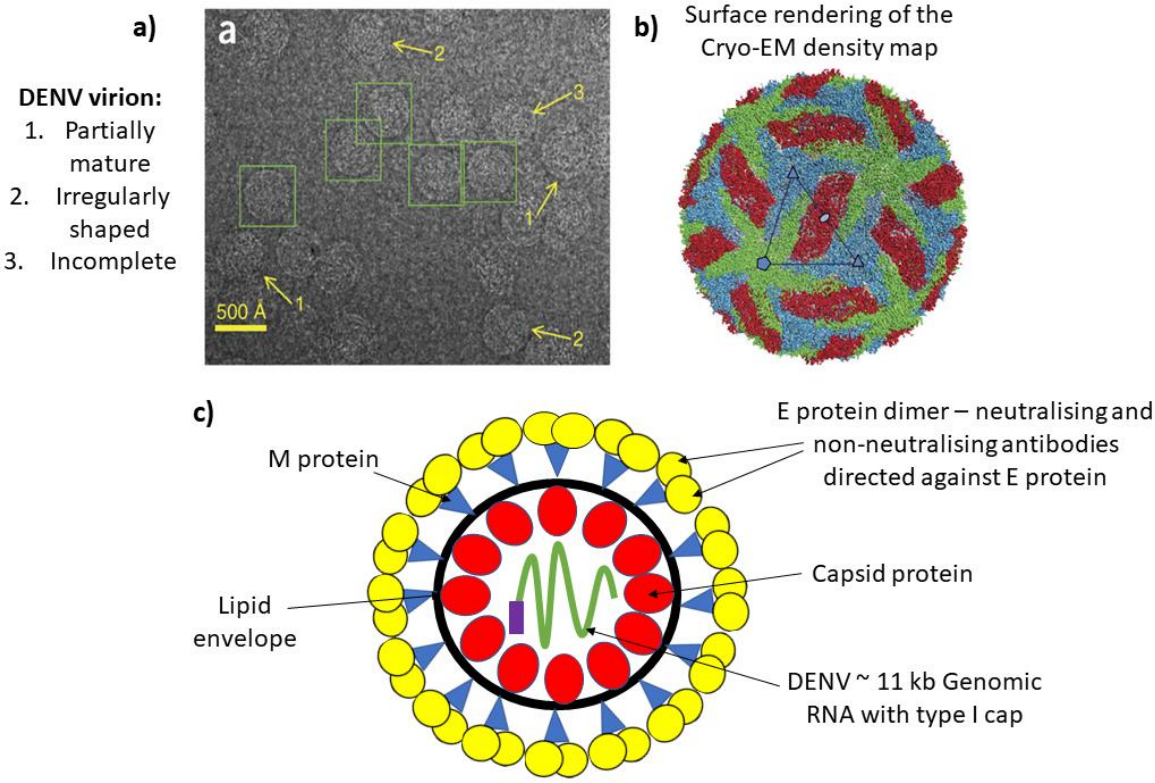



**Figure 12 Overview of the cryo-EM structure of the DENV virion and virion particle schematic:**

***A)*** *Cryo-EM image. Boxed particles were chosen for processing after excluding: 1) partially mature; 2) irregularly shaped or 3) incomplete viral particles. As indicated by yellow arrows.* ***B)*** *Surface rendering of the Cryo-EM density map. E:M heterodimers of the same colour are equivalent by icosahedral symmetry. Heterodimers of different colours are quasi-equivalent, with green E:M dimers falling on the icosahedral 5-fold axes, blue on the 3-fold, and red on the 2-fold. Adapted from* (X. Zhang *et al.*, 2013)*.* ***C)*** *A schematic of the DENV virion. E protein dimers indicated in yellow, M protein in blue, C protein in red, lipid envelope in black, DENV genomic RNA in green and type I cap in purple.*

### 1.2.2 The DENV life-cycle

Infection with DENV occurs when an infected female mosquito vector takes a blood meal and introduces the virus into the host's bloodstream. The delivery of DENV directly into the host's bloodstream is enhanced by the mosquito bite itself. For example, serine proteases within mosquito saliva act to proteolyse host extracellular matrix proteins, inducing cell migration towards the site of infection (Conway *et al.*, 2014; Mores, Christofferson and Davidson, 2014). DENV binds a permissive host cell, via the E protein, and is internalized by receptor-mediated endocytosis. Alternatively, DENV is opsonised for uptake via binding to a non-neutralising antibody during ADE (discussed in detail in section 1.1.3.1). A specific host cell receptor has not yet been identified. However, candidates include the mannose receptor of macrophages or the adhesion molecule of dendritic cells (DC-SIGN) (Miller *et al.*, 2008; Richter *et al.*, 2014). The fact that DENV can infect a wide variety of cell types suggests that DENV does not interact with a specific receptor but utilizes a wide variety of receptors, allowing the infection of a broad-range of cells of both mammalian and insect origin.

Following internalisation, acidification of the endosome triggers an irreversible trimerisation of the E protein which results in the fusion of the viral and host cell membrane (Allison *et al.*, 1995). Following fusion, the nucleocapsid is released into the host cell cytoplasm and DENV genomic RNA is exposed to the host cell environment.

The capped, positive-sense, single-stranded RNA genome is directly translated during early translation by host cell ribosomes resulting in the production of a polyprotein. The polyprotein is subsequently co and post-translationally cleaved by host cell and viral proteases to produce functional DENV structural and non-structural proteins. Following the initial translation of viral proteins, virus induced hypertrophy of the endoplasmic reticulum (ER) occurs allowing DENV to initiate negative-sense RNA synthesis in a concentrated, controlled environment termed the replication complex, facilitating viral replication (Welsch *et al.*, 2009; Gillespie *et al.*, 2010; Inoue and Tsai, 2013). The replication complex allows for the sequestration and concentration of catalysts, reactants and co-factors required for viral replication. Negative-sense RNA acts as a template for the production of

positive-sense RNA DENV genomes (vRNA). Successive rounds of translation from newly formed positive-sense RNA genomes results in the production of high levels of structural and non-structural viral proteins, critical for the life-cycle of the virus (discussed in detail in section 1.2.3).

The structural proteins, alongside the newly formed vRNA, are assembled into immature progeny virions which bud into the lumen of the ER and mature during secretion out of the host cell. During secretion, immature virions enter an acidic *trans*-Golgi compartment where furin cleavage occurs. Furin, a host proprotein convertase, cleaves prM to M generating E protein dimers that lie flat on the viral surface. However, the cleaved pr-protein remains associated with the virus particle blocking the fusion loop of the E protein. The function of pr-protein association with the virion prevents the fusion of the virion with intracellular membranes of the host cell. Upon release of the virion into the extracellular environment, the neutral pH allows for the dissociation of the pr-protein and the mature virus is then able to bind and fuse with a new host cell, continuing the DENV infection cycle (Yu *et al.*, 2008; Smith *et al.*, 2016).

However, furin cleavage, and therefore DENV maturation, is often inefficient as cells infected with DENV release a high proportion of immature prM- containing virions, suggested to be incapable of infection alone but are associated with increased disease pathology upon heterotypic infection (Rodenhuis-Zybert *et al.*, 2010). A schematic of the DENV life-cycle is depicted in figure 13.

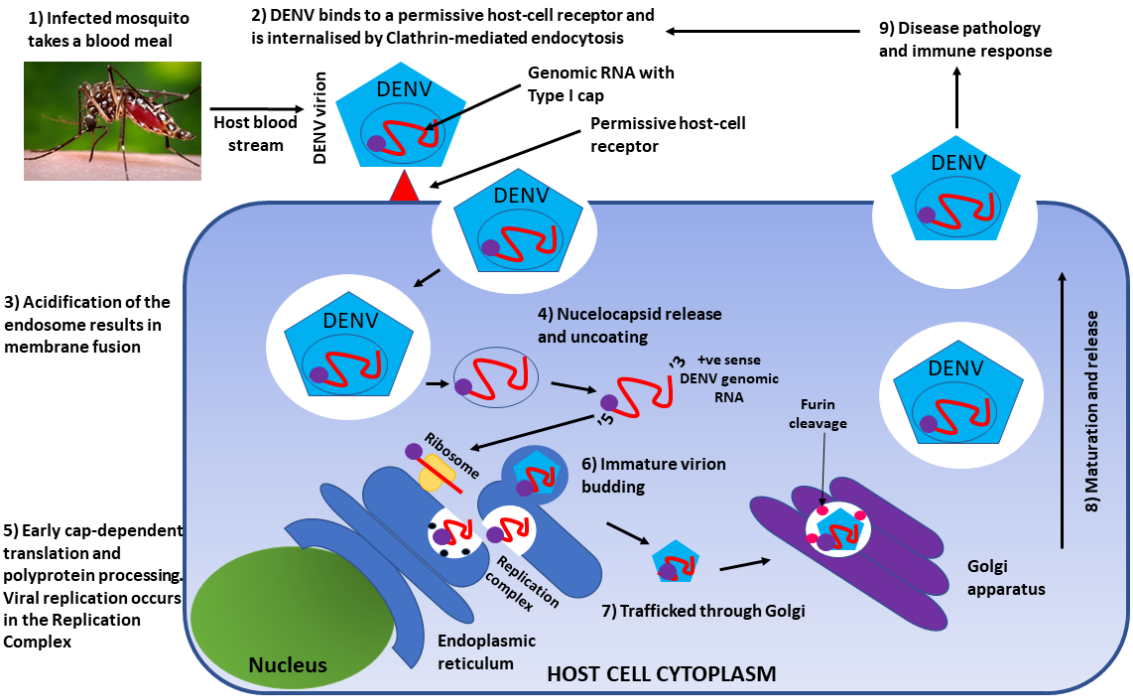



**Figure 13 The DENV life-cycle:**

*A schematic of the DENV life-cycle.* ***1)*** *A DENV infected mosquito takes a blood meal, introducing DENV into the host cell blood stream.* ***2)*** *DENV binds a permissive host cell receptor and is internalised by Clathrin-mediated endocytosis.* ***3)*** *Acidification of the endosome occurs which results in fusion of the viral and host cell membrane, releasing the nucleocapsid into the cytoplasm of the host cell.* ***4)*** *The nucleocapsid is released and DENV genomic RNA is uncoated, exposing DENV genomic RNA to the host cell environment.* ***5)*** *Early cap-dependent translation occurs in 5´-3´ orientation producing a polyprotein which is both co- and post-translationally cleaved by host and viral proteins into functional DENV structural and non-structural proteins. The formation of the replication complex is virally induced and occurs at the ER membrane, the site of DENV replication.* ***6)*** *The immature DENV virion is packaged and buds through the ER.* ***7)*** *The immature virion is trafficked through the Golgi apparatus where furin cleavage occurs cleaving prM to M however the pr-protein remains associated.* ***8)*** *DENV maturation and release into the extracellular environment where the pr-protein dissociates, and the mature virion can continue the infectious cycle. Mosquito image courtesy of the CDC image ID 9261.*

### 1.2.3 Genomic organisation

The capped positive sense, single-stranded RNA genome of DENV is approximately 10.7 Kb in length. The viral genome encodes a single, open reading frame (ORF), flanked by highly structured 5´ and 3´ un-translated regions (UTRs), vital to DENV replication (Fig. 14).

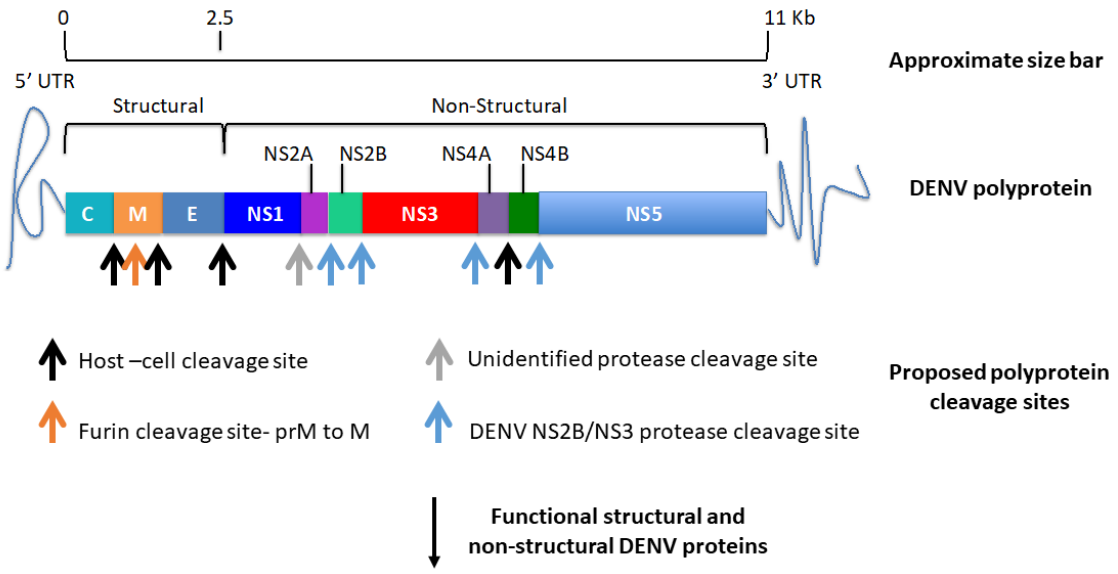



**Figure 14 DENV genomic organisation:**

*The organisation of the +ve sense, single-stranded RNA genome of DENV. Structural and non-structural proteins are indicated. 5´ and 3´ UTR regions are highly structured and fundamental to viral replication. Coloured arrows indicate proposed polyprotein cleavage sites of both host and viral origin. Cleavage of the DENV polyprotein results in functional structural and non-structural DENV proteins thereby allowing the continuation of the DENV life-cycle.*

Host cell early translation of the DENV ORF produces a single polyprotein which is both co and post-translationally cleaved by both viral and host cell proteases, for example host signal peptidases, furin protease and the viral NS2B/NS3 protease (Chambers *et al.*, 1990; Ryan, Monaghan and Flint, 1998) (Fig. 14). The structural proteins C, M and E are located towards the 5´ end of the genome whereas the non-structural (NS) proteins are located at the 3´ end and consist of NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5 (Fig. 14).

#### 1.2.3.1 The structural proteins of DENV

##### 1.2.3.1.1 Envelope (E)

E is the major envelope glycoprotein of DENV and alongside the M protein, forms the outer protein shell of the virion, as a dimer, anchored by an α-helical hairpin loop in the lipid membrane (Kuhn *et al.*, 2002). The E protein binds a permissive host cell receptor during DENV infection, directing the virion into the endocytic pathway. The E protein responds to the acidic environment of the endosome culminating in a large conformational rearrangement (Modis *et al.*, 2004). Firstly, the E protein inserts a hydrophobic anchor, the fusion loop, into the outer bilayer of the host cell membrane. Secondly, the E protein folds back on itself, directing its C terminal transmembrane anchor toward the fusion loop. This fold back forces the host cell membrane and the viral membrane against each other culminating in the fusion of the two membranes (Rey, 2003; Nayak *et al.*, 2009). This action results in the deliverance of viral genomic RNA into the host cell cytoplasm, allowing for the translation of viral proteins and the establishment of infection. Additionally, the E protein is the major epitope for both neutralising and non-neutralising antibodies, further discussed in section 1.1.3.1. Unusually the fusion of the E protein is regulated by the association with the viral protein prM.

##### 1.2.3.1.2 Pr-Membrane/Membrane (prM/M)

The DENV prM/M protein consists of seven antiparallel β-strands stabilised by three disulphide bonds (Perera and Kuhn, 2008). The outer glycoprotein shell of the mature DENV virion is known to consist of 180 copies of each the E protein and the M protein (Perera and Kuhn, 2008). During the life-cycle of DENV the initial virions produced are immature, with both E and prM proteins forming 90 heterodimers that give a spiky appearance to the shell of the virion. However, as the immature virion buds from the ER and travels through the *trans*-golgi network and as the virion exits the cell via the secretory pathway the virion matures, exposed to the acidic environment (Allison *et al.*, 1995; Rey, 2003; Modis *et al.*, 2004). It is believed that the acidic environment causes the conformational change of the E protein, dissociating it from the prM protein culminating in the formation of E homodimers. These homodimers lie flat on the surface of the virion giving the mature virus particle a smooth appearance (Allison *et al.*, 1995; Rey, 2003; Modis

*et al.*, 2004).. However, during the maturation process the pr-protein is cleaved from the M protein by the host cell protease furin. The M protein then acts as a transmembrane protein under the E protein shell of the newly formed mature virion. It is known that the pr peptide remains associated with the E protein until the viral particle is release into the extracellular environment (Stadler *et al.*, 1997; Yu *et al.*, 2008). This action prevents the fusogenic activity of the E protein to inappropriate host cell membranes as the virion is trafficked through the secretory pathway prior to the release of mature virions (Guirakhoo, Bolin and Roehrig, 1992; Heinz *et al.*, 1994; Wang, He and Anderson, 1999; Perera and Kuhn, 2008). However, the maturation of DENV virions is known to be inefficient and is further discussed in section 1.2.2.

##### 1.2.3.1.3 Capsid (C)

Interior to the outer bilayer of the DENV virion is the nucleocapsid core that comprises multiple copies of the capsid or core protein (C), and a single copy of the ~11 Kb DENV RNA genome. The C protein is essential to DENV assembly, ensuring the specific encapsidation of the DENV genome. However, the mechanism by which genomic encapsidation occurs is not understood but may require participation of the non-structural proteins as well as the C protein or a specific conformation of DENV genomic RNA may be necessary (Kuhn *et al.*, 2002; Ma *et al.*, 2004).

#### 1.2.3.2 The non-structural proteins of DENV

##### 1.2.3.2.1 Non-structural protein 1 (NS1)

NS1 is a ~ 49 kDa glycoprotein that, as a non-structural protein, is absent from the infectious viral particle (Avirutnan *et al.*, 2007). NS1 functions as a co-factor for viral RNA replication and co-localizes with the double-stranded RNA replicative form (Mackenzie, Jones and Young, 1996; Lindenbach and Rice, 1999; Avirutnan *et al.*, 2007). NS1 is synthesised in an infected cell as a soluble monomer which dimerises following post-translational modification in the lumen of the ER. Dimerised NS1 is subsequently transported to the cell surface and excreted into the extracellular milieu. The accumulation of soluble NS1 in serum is known to correlate with disease severity (Avirutnan *et al.*, 2007). However, the mechanisms

by which NS1 contributes to DENV pathogenesis remain unclear and may be associated with the orchestration of the immune response. For example, NS1 is suggested to attenuate the alternative pathway of complement activation following the binding of factor H (Chung *et al.*, 2006).

#### 1.2.3.2.2 Non-structural protein 2A (NS2A)

NS2A is a highly hydrophobic transmembrane protein that interacts with lipid bilayers and is known to co-localize with the double-stranded RNA replicative intermediate, with the 3´UTR of DENV genomic RNA and with other non-structural proteins within the replication complex (Xie *et al.*, 2013; Shrivastava *et al.*, 2017). NS2A is suggested to comprise a putative viroporin of DENV-2, allowing for the transport of nascent RNA to the site of viral assembly and membrane remodelling transporting RNA or partially assembled nucleocapsids to subcellular compartments during viral assembly and release (Leung *et al.*, 2008; Shrivastava *et al.*, 2017).

#### 1.2.3.2.3 Non-structural protein 2B (NS2B)

DENV NS2B is a transmembrane protein which localises on the cell membrane. NS2B mainly functions as a co-factor to form a stable NS3-NS2B serine-threonine viral protease which, for example, is mandatory for the cleavage of the DENV polyprotein into functional proteins and is therefore critical to the DENV life-cycle (Li *et al.*, 2015; Wu *et al.*, 2015).

#### 1.2.3.2.4 Non-structural protein 3 (NS3)

The NS3 protein functions to cleave the viral polyprotein in the cytoplasm. Cleavage of the polyprotein is mediated by the serine protease N-terminal domain of NS3, alongside a hydrophilic segment of 40 residues from NS2B which acts as a co-factor necessary for protease activity. Additionally, the domain required for ATPase/helicase and nucleoside 5’triphosphatase activity is located at the C terminus of NS3. Thus the NS3 protein is a critical determinant of the DENV life-cycle (Luo *et al.*, 2008; Perera and Kuhn, 2008).

#### 1.2.3.2.5 Non-structural protein 4 A/B (NS4A/B)

Membrane proteins NS4A/B are linked by a conserved 23 amino acid signal peptide (Zou *et al.*, 2015). The cleavage between NS4A and B is mediated by the NS2B/3 proteinase (Lin *et al.*, 1993). NS4A consists of 127 amino acids and contains two trans-membrane domains reported to form an amphipathic helix that mediates oligomerisation (Welsch *et al.*, 2009; Stern *et al.*, 2013). NS4B consists of 248 amino acids and contains three trans-membrane domains (Miller, Sparacio and Bartenschlager, 2006). It is known that both NS4A and B play a multitude of roles in viral replication and virus host interactions. For example, NS4A is known to induce ER membrane rearrangement and replication complex formation via interaction with cellular vimentin (Teo and Chu, 2014). NS4A induces autophagy acting to prevent cell death facilitating viral replication (McLean *et al.*, 2011). NS4B interacts with the helicase domain of NS3 dissociating it from single-stranded RNA (Zou *et al.*, 2015).

##### 1.2.3.2.6 Non-structural protein 5 (NS5)

NS5 encodes the methyltransferase and RdRp activity of DENV and is discussed in detail in section 1.3.1.

## 1.3 DENV replication

### 1.3.1 Non-structural protein 5 in detail

DENV non-structural protein 5 (NS5) is, at 100 kDa, the largest and most conserved (~ 70% amino-acid sequence identity) protein expressed during DENV infection, a significant anti-viral target. NS5 is comprised of two domains: an N-terminal methyltransferase domain (MTase) and a C'-terminal RdRp connected by a flexible linker region (Lim *et al.*, 2013; Zhao, Soh, Lim, *et al.*, 2015; El Sahili and Lescar, 2017) (Fig. 15).

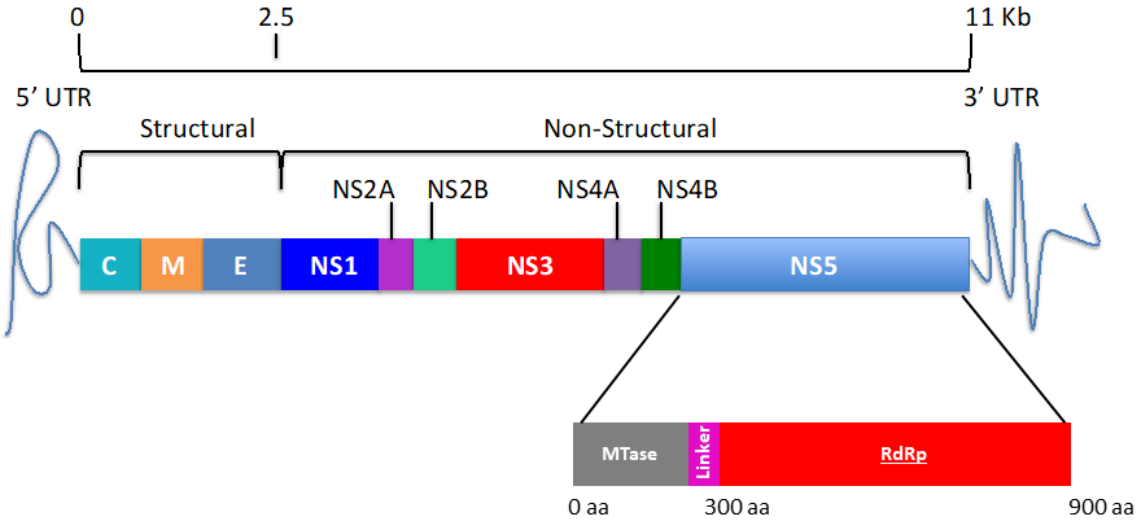



**Figure 15 DENV NS5 in the context of the DENV genome:**

*A schematic of DENV genomic RNA and NS5. MTase = Methyltransferase, RdRp = RNA-dependent RNA polymerase shown in grey and red respectively. The linker domain is highlighted in pink.*

#### 1.3.1.1 Methyltransferase domain

The DENV genome consists of a positive sense single stranded RNA genome with a type I cap structure, followed by the conserved nucleotide sequence 'AG': 5´$^{m7}$Gppp**A**$_{m2\text{-}O}$-**G**-3´ (Zhao, Soh, Lim, *et al.*, 2015). The MTase domain of NS5, spanning residues 1-262, is responsible for the addition of the cap moiety selectively on to the 5´ end of viral genomic RNA and is critical for viral replication via the direct early-translation of DENV genomic RNA. The addition of the cap moiety ensures host cell translation of viral genomic RNA and the production of

viral polyproteins, critical to the DENV life-cycle. The addition of the type I cap moiety protects the viral genomic RNA from 5´-3´ exoribonuclease degradation, concealing the viral phosphorylated genomic RNA from the host innate immune response (Pichlmair *et al.*, 2006; Goubau *et al.*, 2014; Zhao, Soh, Lim, *et al.*, 2015).

Following *de novo* RNA synthesis, catalysed by the RdRp, type I cap formation results from several sequential enzymatic reactions. Firstly, the RNA triphosphatase activity of the NS3 protease-helicase hydrolyses the γ-phosphate group of the viral 5´UTR, yielding a diphosphate nucleotide. Secondly, the guanyl-transferase activity of the MTase domain catalyses the transfer of a GMP molecule to the 5´diphosphate nucleotide. Finally, MTAse mediated sequential addition of N-7- and 2´-O- methylations occur (Zhao, Soh, Lim, *et al.*, 2015; El Sahili and Lescar, 2017) (Fig. 16).

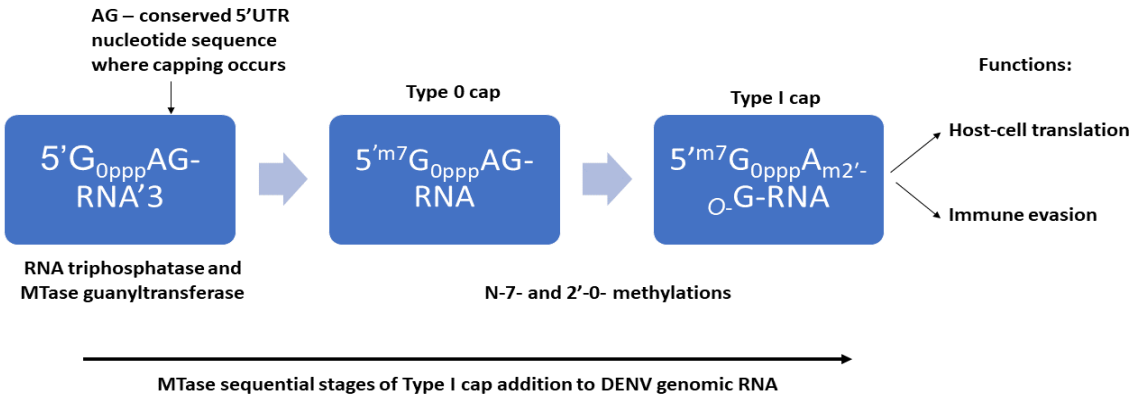



**Figure 16 Type I cap formation by the MTase domain:**

*A schematic of MTase domain Type I cap formation and its importance in the viral life-cycle. Firstly, the RNA triphosphatase activity of the NS3 protease-helicase hydrolyses the γ-phosphate group of the viral 5´UTR, yielding a diphosphate nucleotide. Secondly, the guanyl-transferase activity of the MTase domain catalyses the transfer of a GMP molecule to the 5´diphosphate nucleotide yielding a Type 0 cap. Finally, MTAse mediated sequential addition of N-7- and 2'-O- methylations occur yielding a Type I cap. The type I cap of DENV functions in early host cell ribosomal translation producing the DENV polyprotein. The Type I cap is also implicated in DENV immune evasion strategies by concealing the phosphorylated viral RNA from the host innate immune system.*

#### 1.3.1.2 NS5 linker domain

A putative linker domain (spanning residues 263-273) connects the MTase and RdRp domains of Flavivirus NS5 (Fig. 15). The sequence of this domain is poorly conserved although the length of the region is conserved across the Flavivirus genera (Lim *et al.*, 2013; Zhao, Soh, Zheng, *et al.*, 2015).

Crystal structures of full length JEV and DENV-3 NS5 have revealed differences in the conformation of the linker domain reflected by differences in the spatial arrangement of the two enzymatic domains, indicating conformational flexibility (Lim *et al.*, 2013; Lu and Gong, 2013; Zhao, Soh, Zheng, *et al.*, 2015). It is therefore proposed that the linker domain functions like a swivel, leading to the formation of different interfaces – enabling the MTase and RdRp to adopt differential orientations during the transition from RNA synthesis to RNA capping, binding to NS3 or to host cell factors during the viral life-cycle (Lim *et al.*, 2013; Zhao, Soh, Zheng, *et al.*, 2015) (Fig. 17).

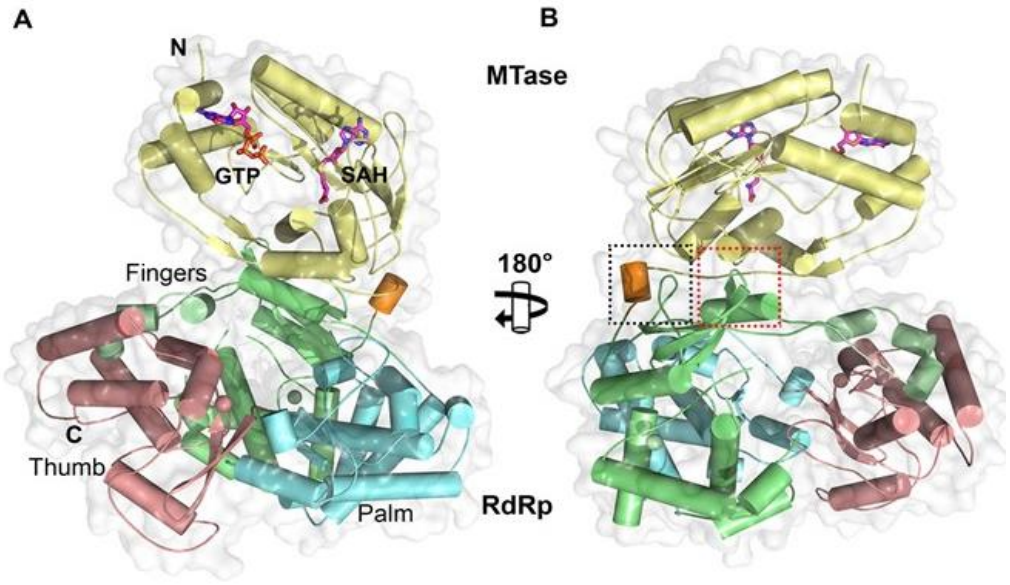



**Figure 17 Crystal structure of DENV-3 NS5:**

*A cartoon representation of the overall structure of the DENV3 NS5 protein.* ***A)*** *Viewing from the bottom of RdRp, MTase in yellow, RdRp fingers in green, palm in blue and thumb in pink. The linker domain is in orange. GTP and co-factor SAH are displayed as sticks and labelled respectively. Zinc ions are shown as spheres.* ***B)*** *View from the top of the NS5 molecule which is rotated by 180° around a vertical axis. The linker domain region is boxed. Adapted from* (Zhao, Soh, Zheng, *et al.*, 2015)

It is suggested that the enzymes, connected by the linker-domain region, promote the activity of one another. The RdRp domain was shown to enhance internal RNA methylation activity, in the context of full-length NS5 proteins from DENV and WNV (Dong *et al.*, 2012). However, results for the influence of the MTase domain on RdRp activity remain contradictory. On one hand, it is proposed that RdRp activity is higher in the context of full-length NS5, than in the context of the RdRp domain alone. This was confirmed using filter binding assays with poly(rC) and poly(rU) templates, detecting the incorporation of [$^3$H]-labelled nucleotides into the nascent RNA chain. The inclusion of the MTase domain was suggested to increase replication activity by 6-17 fold over that observed by the RdRp domain alone (Potisopon *et al.*, 2014). On the other hand, utilising a similar methodology measuring the incorporation of [$^3$H]-GMP on poly(rC)/ poly(rU) templates, it has been found that inclusion of the MTase domain does not have an effect on RdRp activity, displaying activity directly comparative to its full-length counterpart (Selisko *et al.*, 2006).

It was observed that the addition of linker residues to the N-terminus of the RdRp domain acted to stabilise the domain, enhancing the *de novo* RNA synthesis capability of the polymerase. A 2.6 Å resolution structure of an RdRp fragment of DENV-3 spanning NS5 residues 265-900 with enhanced *de novo* RNA synthesis activity was reported (Lim *et al.*, 2013). This observation is of significance, it is well documented that full-length purified DENV NS5 proteins often degrade substantially throughout the purification process, deeming functional analysis challenging. Therefore, the translation of this observation in the study of the RdRp domain in the context of other DENV serotypes could aid in determining the structural functionality of this critical viral replicative protein (Selisko *et al.*, 2006).

#### 1.3.1.3 Conformational dynamics of DENV RdRp during the initiation of DENV genomic replication

##### 1.3.1.3.1 Mechanistic process of DENV replication

DENV NS5 contains a C- terminal RdRp, responsible for viral replication and fundamental to the viral life-cycle. The RdRp, within the replication complex, first transcribes the viral genome, using the positive-sense RNA genome as a template (Fig. 18). The negative-sense RNA strand is then utilised as a template to

synthesise an excess of positive sense, single stranded genomic RNA, used for further translation and viral polyprotein production and for encapsidation into new progeny virions. The mechanistic process of DENV replication catalysed by DENV RdRp is described figure 18.

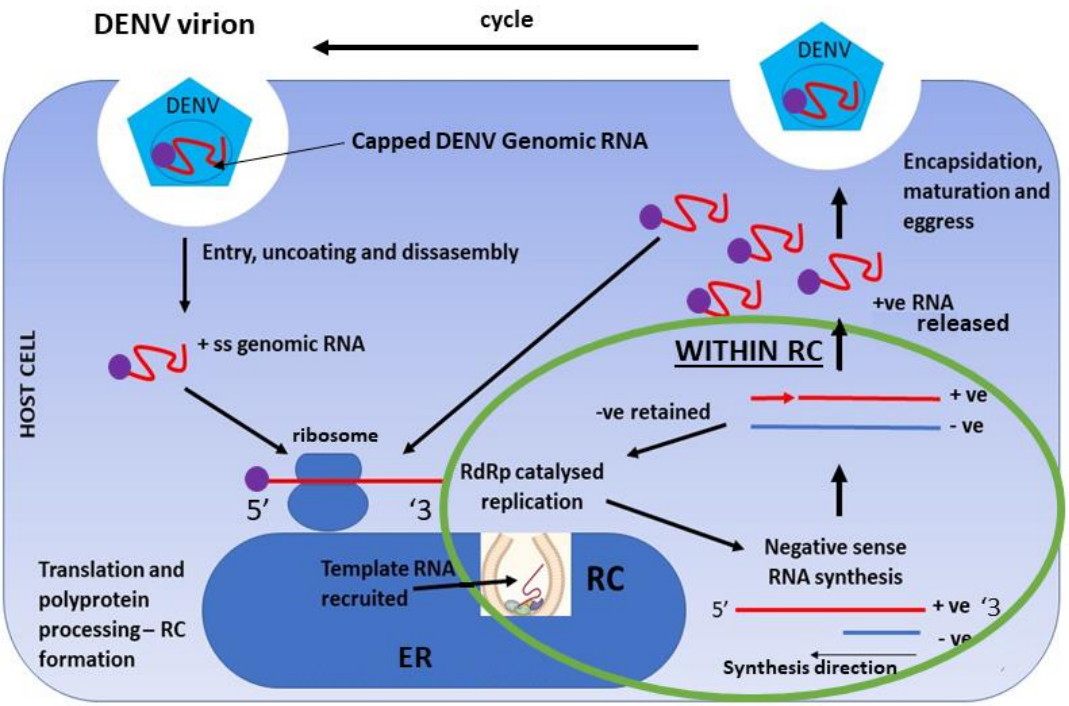



**Figure 18 DENV genome replication catalysed by the DENV RdRp:**

*A schematic representation of DENV replication in the context of the DENV life-cycle. Following infection of a permissive host cell and virion disassembly, capped DENV genomic RNA is released into the host cell cytoplasm. DENV genomic RNA is directly translated by host cell ribosomes resulting in the production of a polyprotein which is co- and post-translationally cleaved by host and viral proteases into functional viral proteins. The formation of the replication complex is virally induced, associated with the endoplasmic reticulum membrane (ER). Functional DENV non-structural proteins comprise the replication complex and DENV replication occurs solely within the RC. The mechanistic details of DENV replication within the replication complex are highlighted by the green circle. RC = Replication complex. +ve and -ve refer to positive and negative -sense genomic RNA templates respectively. Replication complex image adapted from* (Nagy and Pogany, 2011).

#### 1.3.1.3.2 The structure of the DENV RdRp

The first crystal structure of the DENV RdRp domain (residues 270 – 900) was published 11 years ago and revealed a right-hand enzyme architecture, structurally conserved amongst many DNA and RNA polymerases (Yap *et al.*, 2007). The DENV RdRp was found to consist of palm, thumb and finger sub-domains containing conserved motifs A-G with precise functions conserved throughout the *Flaviviridae* family (Fig. 19). Motif A contributes to the cation binding site (Yap *et al.*, 2007; Wu, Liu and Gong, 2015). Motif B aids in the process of template recruitment and positioning in the RdRp tunnel (Wu, Liu and Gong, 2015; Shu and Gong, 2017; Yang *et al.*, 2017). Motif C comprises the GDD catalytic residue – the active site (Yap *et al.*, 2007; Wu, Liu and Gong, 2015). The functionality of Motif D remains largely unknown but is proposed to aid in the release of the PPi by-product of replication (Castro *et al.*, 2009; Wu, Liu and Gong, 2015). Motif E comprises the structural zinc cation (Yap *et al.*, 2007; Wu, Liu and Gong, 2015). Motif F is involved in the stabilisation of the nascent RNA base pair (Iglesias, Filomatori and Gamarnik, 2011) and the functionality of Motif G is unknown but may be involved in the binding and translocation of the template RNA strand (Wu, Liu and Gong, 2015) (Fig. 19).

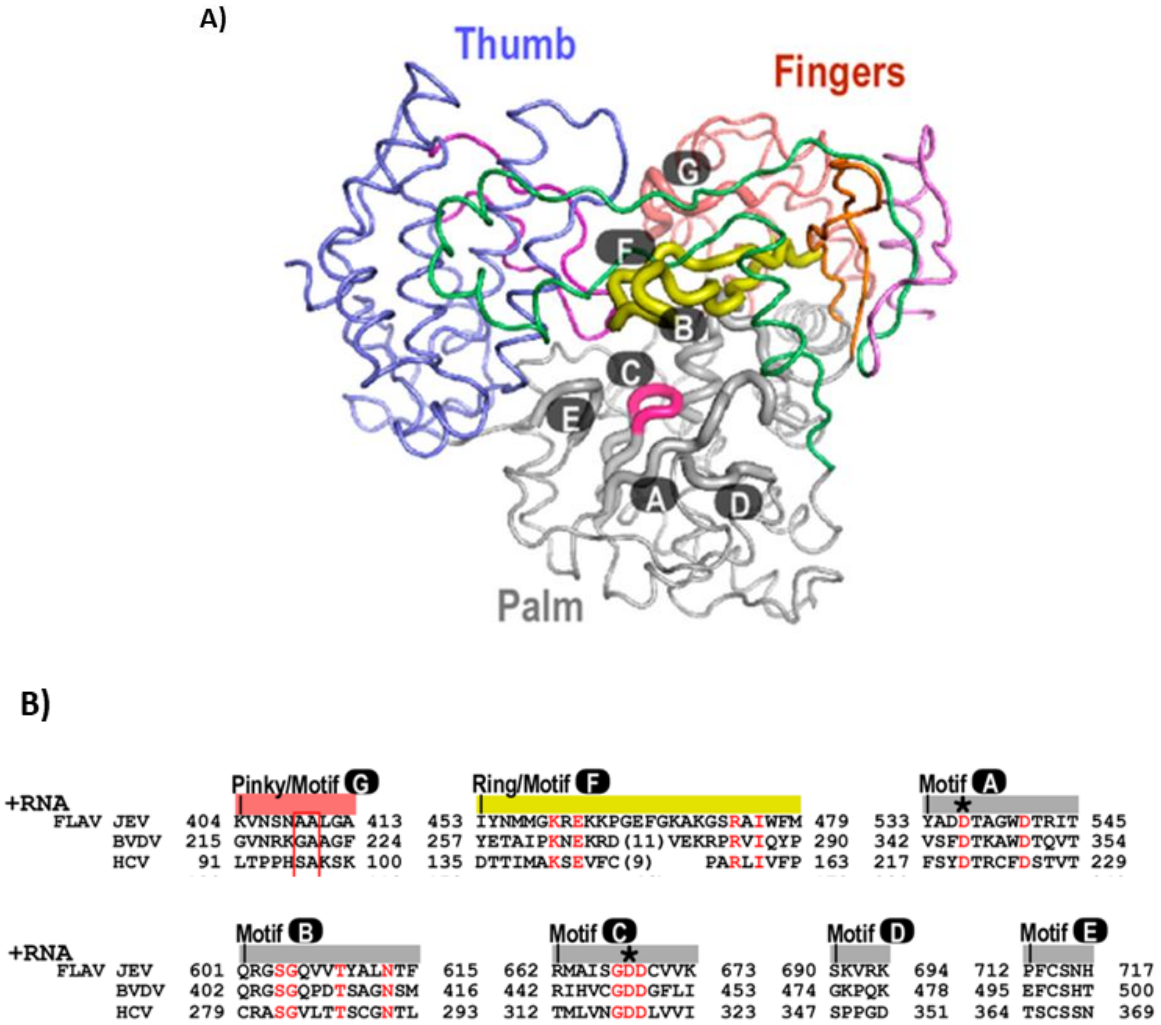



**Figure 19 Structural position/ conservation of *Flaviviridae* RdRp catalytic motifs:**

***A)*** *The spatial organisation of JEV RdRp catalytic motifs A-G (pdb accession number: 4K6M.* ***B)*** *Structure-based sequence alignment displaying the conservation of RdRp structural motifs. JEV (Japanese encephalitis virus) is a member of the Flavivirus genera. BVDV (Bovine viral diarrhoea virus) is a member of the Pestivirus genera. HCV (Hepatitis C virus) is a member of the Hepacivirus genera. Adapted from* (Wu, Liu and Gong, 2015)*.*

Together these regions and motifs shape a right-handed polymerase structure comprising three tunnels which allow access to the 3´ end of the template genomic RNA, incoming ribonucleoside tri-phosphates (rNTPs) and allow exit of the newly synthesised dsRNA product (Yap *et al.*, 2007). The palm domain contains the active site of the RdRp containing two aspartic residues which act to chelate two $Mg^{2+}$ metal ions allowing for the formation of the phosphodiester bond and therefore catalysing the formation of nascent RNA (Yap *et al.*, 2007) (Figs. 19 and 20).

Additionally, a “priming loop” extends from the thumb sub-domain, towards the active site. The presence of this loop is often considered the “hallmark” of RdRp domains capable of *de novo* RNA synthesis (Yap *et al.*, 2007; El Sahili and Lescar, 2017). The priming loop acts as a platform for the initiation of RdRp polymerisation activity and precise and relatively efficient initiation is achieved with the priming element, aiding in the stabilisation of the initiation complex, allowing for primer-independent viral replication (Yap *et al.*, 2007; Wu, Liu and Gong, 2015; El Sahili and Lescar, 2017) (Fig. 20).

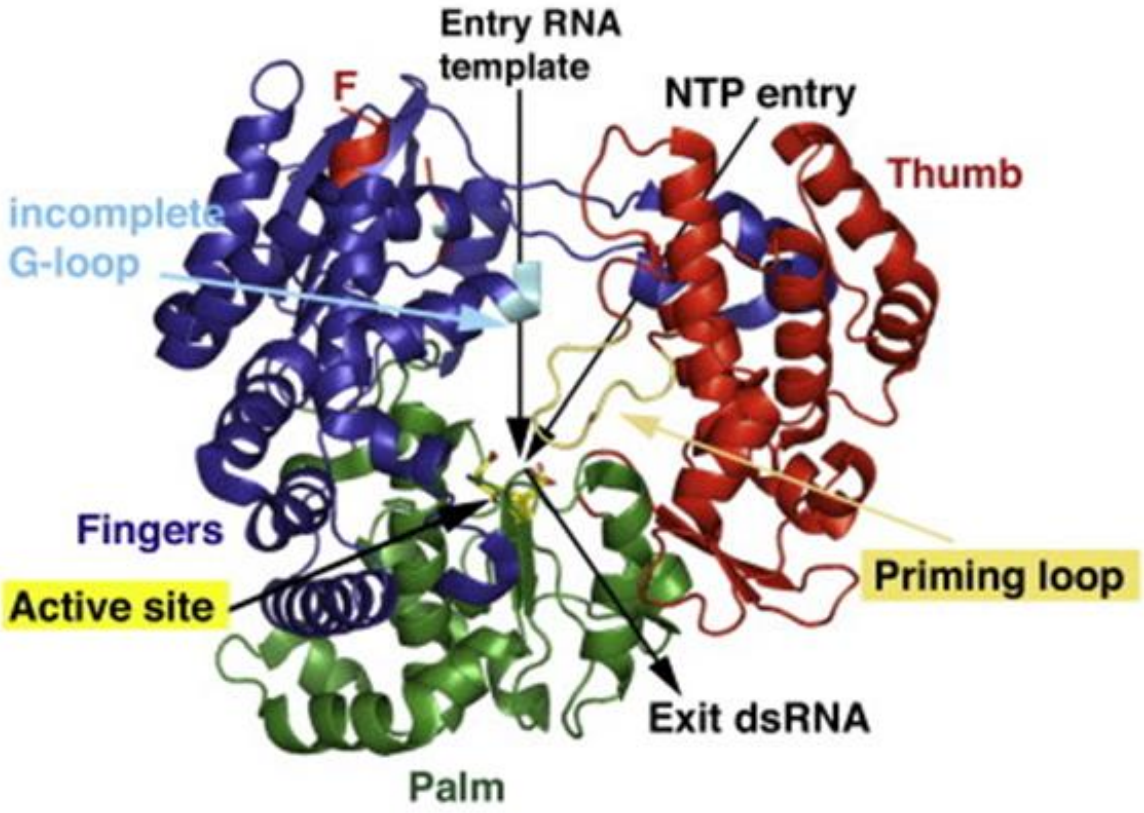



**Figure 20 Annotated crystal structure of the DENV-3 RdRp domain:**

*An annotated crystal structure of the DENV-3 RdRp domain. Fingers coloured in blue, palm in green and thumb in red. The priming loop extends towards the active site of the protein and is shown in orange. The position of the RNA template, NTP entry, and dsRNA exit tunnels are indicated by arrows. The active site aspartic acids of motifs A and C are shown as sticks and coloured according to atom type (oxygen- red, carbon – yellow). Motif F and the incomplete G loop are shown in red and cyan. Adapted from* (Selisko *et al.*, 2014).

#### 1.3.1.3.3 Primer-independent replication initiation and RdRp conformational change

The DENV RdRp binds selectively to stem-loop A (SLA), the promoter stem-loop in the 5´UTR of DENV genomic RNA, suggested to be via an arginine rich sequence in the thumb subdomain of DENV RdRp, as a result of the strong positive charge (Filomatori *et al.*, 2006; Lodeiro and Filomatori, 2009; Hodge *et al.*, 2016). DENV genomic cyclisation acts to transfer the RdRp to the 3´UTR of DENV genomic RNA, thereby allowing for the initiation of negative-sense RNA synthesis (Filomatori *et al.*, 2006; Villordo and Gamarnik, 2009; de Borba *et al.*, 2015) (discussed in detail in section 1.3.2.1). The RdRp synthesises new DENV RNA in the absence of a primer, termed primer-independent or *de novo* RNA synthesis (Kao, Singh and Ecker, 2001; Nomaguchi *et al.*, 2003; Selisko *et al.*, 2006).

The DENV RdRp synthesises RNA *de novo* in three main stages: *de novo* initiation (primer synthesis), transition and elongation (Jin *et al.*, 2011; El Sahili and Lescar, 2017). Firstly, the RdRp generates short primers over the 3´terminus of the viral RNA template ending in 5´-CU-`3 corresponding to the strictly conserved end of the Flavivirus genome (Selisko *et al.*, 2012). Then it is suggested that the RdRp undergoes structural/mechanistic transition from a “closed” *de novo* initiation complex to an “open” elongation conformation before continuing RNA synthesis in a more rapid and processive manner (Noble and Shi, 2012; Potisopon *et al.*, 2014).

The structural basis behind the DENV RdRp transition from a “closed” to “open” conformation remains unknown. The available crystal structures of Flavivirus full-length NS5 and the RdRp domain alone, display a pre-initiation, closed conformation, unbound to genomic RNA, with the priming loop positioned towards the active site. In this conformation the RNA binding tunnel of the DENV RdRp is predicted to be too narrow to accommodate the dsRNA replicative intermediate (Malet *et al.*, 2007, 2008; Yap *et al.*, 2007; El Sahili and Lescar, 2017).

It is unknown whether or not the transition from a closed conformation to an open conformation occurs following binding to SLA or directly after the synthesis of a priming dinucleotide/trinucleotide, or if a longer primer is required. It is expected that during the conformational change from a pre-initiation to a processive elongation state, an outward rotation of the fingers domain must occur, relative to

the palm subdomain, and a retraction of the priming loop is expected away, from the active site of the RdRp (Appleby *et al.*, 2015). The projected concerted conformational movements of the DENV RdRp would culminate in an increase in the volume of the dsRNA exit tunnel, enabling the translocation of dsRNA, the replicative intermediate, enabling continued viral replication (Appleby *et al.*, 2015; El Sahili and Lescar, 2017). The predicted DENV RdRp conformational changes are depicted in figure 21.

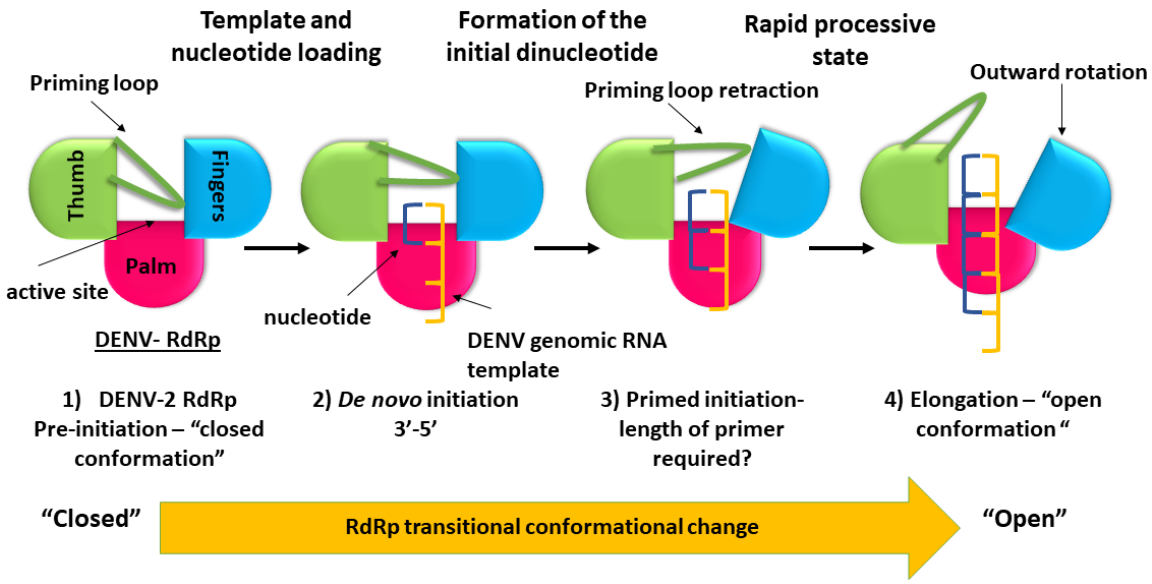



**Figure 21 Predicted DENV-RdRp conformational transition during replication**

*Schematic of representative steps during RNA synthesis indicating the proposed RdRP conformational changes.* ***1)*** *The DENV RdRp in a pre-initiation 'closed' conformation state. The priming loop extends towards the active site of the RdRp.* ***2)*** *DENV RdRp de novo initiation in the 3`- 5´ direction.* ***3)*** *DENV RdRp in the primed initiation state. The length of the required primer is unknown.* ***4)*** *DENV RdRp in the elongation 'open' conformation. The priming loop retracts and an outward rotation of the fingers is expected to accommodate the double-stranded replicative intermediate.*

Following many years of research, the only member of the *Flaviridae* family for which the RdRp has been observed structurally at different replicative initiation and elongation stages, is the Hepacivirus HCV NS5B RdRp protein (Appleby *et al.*, 2015). Subsequently, the obtained structural and mechanistic data have been exploited in the development of HCV nucleotide therapies with a higher barrier to the selection of drug resistance (Appleby *et al.*, 2015). Therefore, our understanding of DENV replication demands comparative interrogation, following the binding of DENV RdRp to SLA, and is consequently a focus of this project.

### 1.3.2 RNA structure and DENV genomic replication

Besides encoding for viral proteins, positive-sense single stranded RNA genomes contain a wide variety of *cis*-acting RNA elements, for example stem-loops, and pseudoknots, which are known to regulate fundamental processes of the viral life-cycle and are vital to the overall secondary and therefore tertiary structure of viral genomes (Villordo, Alvarez and Gamarnik, 2010). These elements are known to have a multitude of functions such as promoters, enhancers and as repressors of translation, transcription, genome replication and encapsidation (Andino, Rieckhof and Baltimore, 1990; Song and Simon, 1995; McKnight and Lemon, 1998; Dreher, 1999; Alvarez, Laura De Lella Ezcurra, *et al.*, 2005; Villordo, Alvarez and Gamarnik, 2010). RNA secondary structure is an extremely versatile viral tool, facilitating the viral life-cycle without sacrificing the limited coding capacity of the viral genome.

Important structural elements of viral RNA are RNA stem-loops and pseudoknots (Fig. 22). RNA stem-loops are formed by intramolecular base-pairing (both Watson-Crick and non-Watson-Crick wobble base pairs) and can also be referred to as a hairpin or hairpin loops (Leontis, Stombaugh and Westhof, 2002; Manas *et al.*, 2015). An RNA stem-loop occurs when two regions, usually complementary, of the same nucleotide strand base pair to form a double-stranded helix that comprises a single-stranded nucleotide terminal loop. The resulting stem-loop structure is often vital in terms of both nucleotide sequence and structure to the viral life-cycle (Clyde and Harris, 2006; Filomatori *et al.*, 2006; Wang *et al.*, 2017). An RNA pseudoknot is defined as an RNA structure formed by the nucleotide base-pairing of a stem-loop terminal loop to a region located upstream or downstream

of the stem flanking the loop (Staple and Butcher, 2005). Pseudoknots have also been implicated in a multitude of biological functions for example gene expression and viral replication (Staple and Butcher, 2005; Alla and Alexander, 2014) (Fig. 22).

Additionally, long-range nucleotide base pairing interactions allow for the extraordinary flexibility displayed by viral genomic RNA. For example, *cis*-acting RNA elements, vital to negative-sense RNA synthesis, can be found at the 5´ genomic RNA terminus despite the fact that negative sense RNA synthesis is initiated from the 3´ terminus of genomic RNA (Alvarez, Lodeiro, *et al.*, 2005) (Fig. 22).

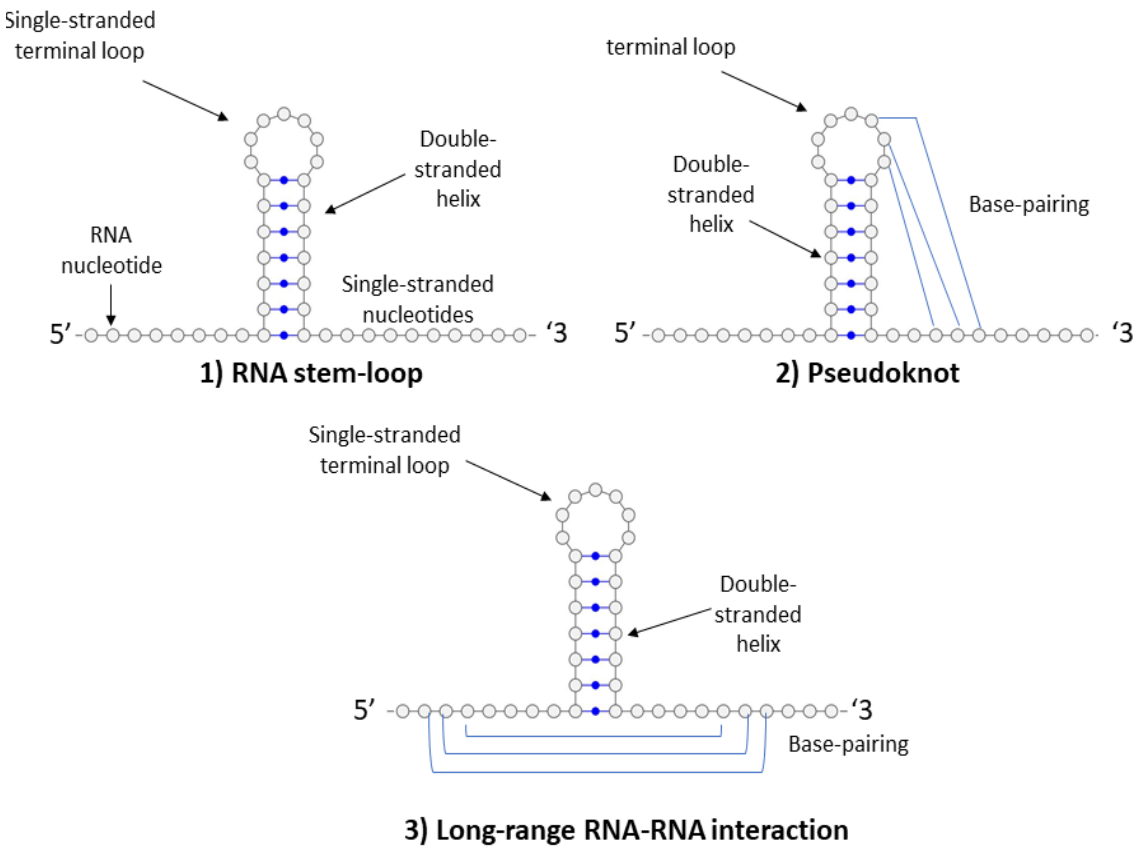



**Figure 22 A schematic of RNA secondary structural elements:**

*A schematic of RNA structural elements.* ***1)*** *An RNA stem-loop is depicted showing the base-paring of RNA nucleotides forming a double-stranded helix also comprising a single-stranded nucleotide terminus.* ***2)*** *An RNA pseudoknot is depicted showing the base-pairing of an RNA stem-loop to a nucleotide region located downstream of the stem flanking the terminal loop.* ***3)*** *An example long-range RNA-RNA interaction is shown. RNA nucleotide base-pairing indicated by blue lines. Figure created using VARNA.*

#### 1.3.2.1 Genomic cyclisation and DENV replication

The linear to circular genomic conformational switch is yet another function of Flavivirus secondary structure, facilitating the viral life-cycle, without sacrificing the limited coding capacity of the viral genome. Following extensive mutational analysis, it is known that both the linear and circular genomic conformations are a fundamental requirement of the DENV life-cycle and therefore the ability to form both genomic conformations exerts a significant selective pressure (Alvarez, Lodeiro, *et al.*, 2005; Villordo, Alvarez and Gamarnik, 2010). Critically, what regulates the switch between the two genomic conformations is unknown and is consequently a subject of our experimental analysis utilising *in vitro/* intracellular **S**elective 2' **H**ydroxyl **A**cylation Analysed by **P**rimer **e**xtension (SHAPE) mapping (methodology discussed in detail in Chapter 4.1) to interrogate the requirements of the conformational switch in the context of DENV replication.

Viral genomic cyclisation for life-cycle facilitation is not a novel concept. For example, an important translational feature of HCV is likely conserved from the model of classical messenger RNA (mRNA) circularisation involving the highly structured HCV 3´ UTR and interactions with translation initiation factors such as polypyrimidine tract binding protein, to facilitate cap-independent translation (Ito, Tahara and Lai, 1998; Wells *et al.*, 1998). The interaction of the HCV 3´ UTR with translational initiation factors functions to circularise the HCV genome during translation, therefore increasing translational efficiency of viral proteins, allowing for ribosome recycling during successive rounds of translation (Ito, Tahara and Lai, 1998; Bradrick, Walters and Gromeier, 2006). In contrast, Flavivirus genomes are able to cyclise by long-range direct RNA-RNA interactions (Villordo and Gamarnik, 2009; de Borba *et al.*, 2015; Liu *et al.*, 2016).

A conserved feature of Flavivirus genomes, is the presence of inverted complementary sequences, termed cyclisation sequences at both termini of the genomic RNA, which mediate long-range RNA-RNA interactions functioning to stabilise genome cyclisation by direct base-pairing (Alvarez, Lodeiro, *et al.*, 2005). The cyclisation sequences are termed: 5´ upstream of AUG region (UAR) which binds to the 3´UAR, 5´ downstream of AUG region (DAR) which binds to the 3´DAR and the 5´cyclisation sequence (CS) which binds to the 3´CS with the combined

action of genomic cyclisation (Alvarez, Lodeiro, *et al.*, 2005). The structural changes associated with genomic cyclisation in the context of full-length DENV genomic RNA are depicted in figure 23.

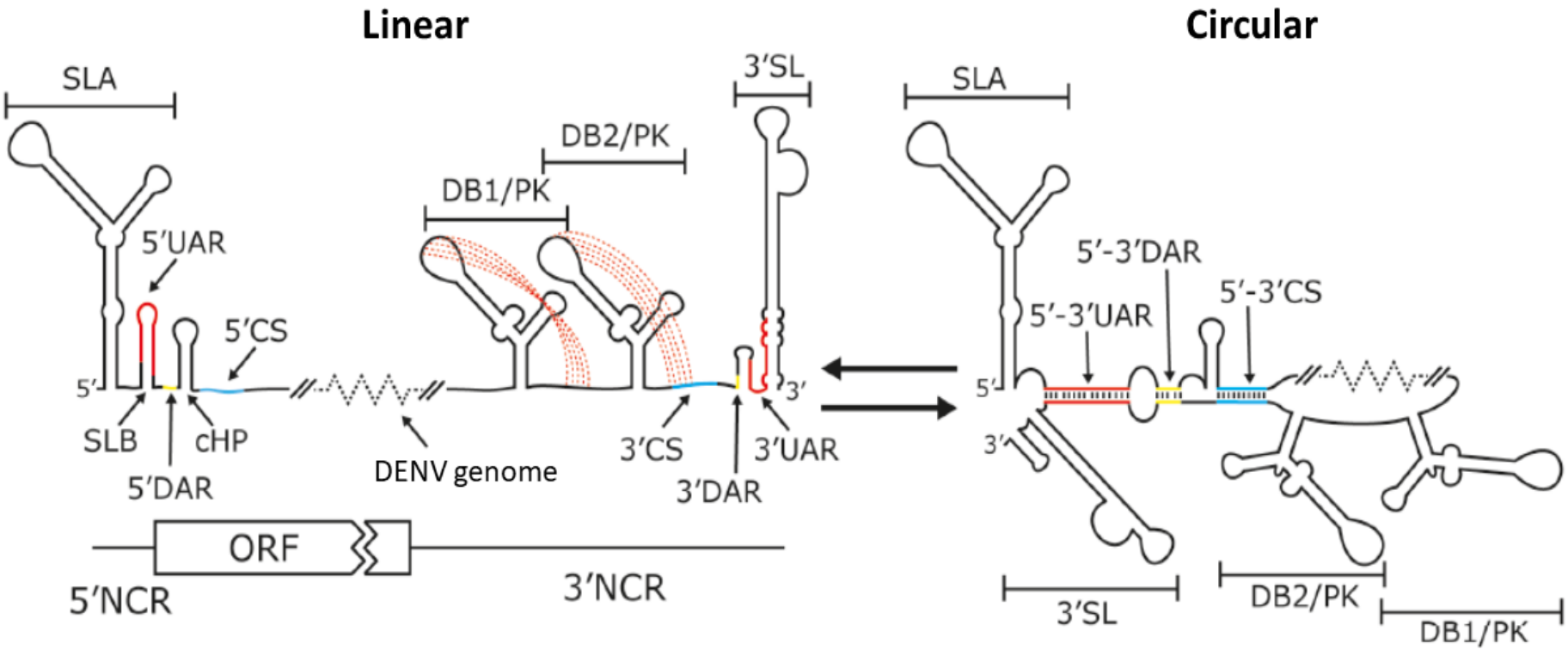
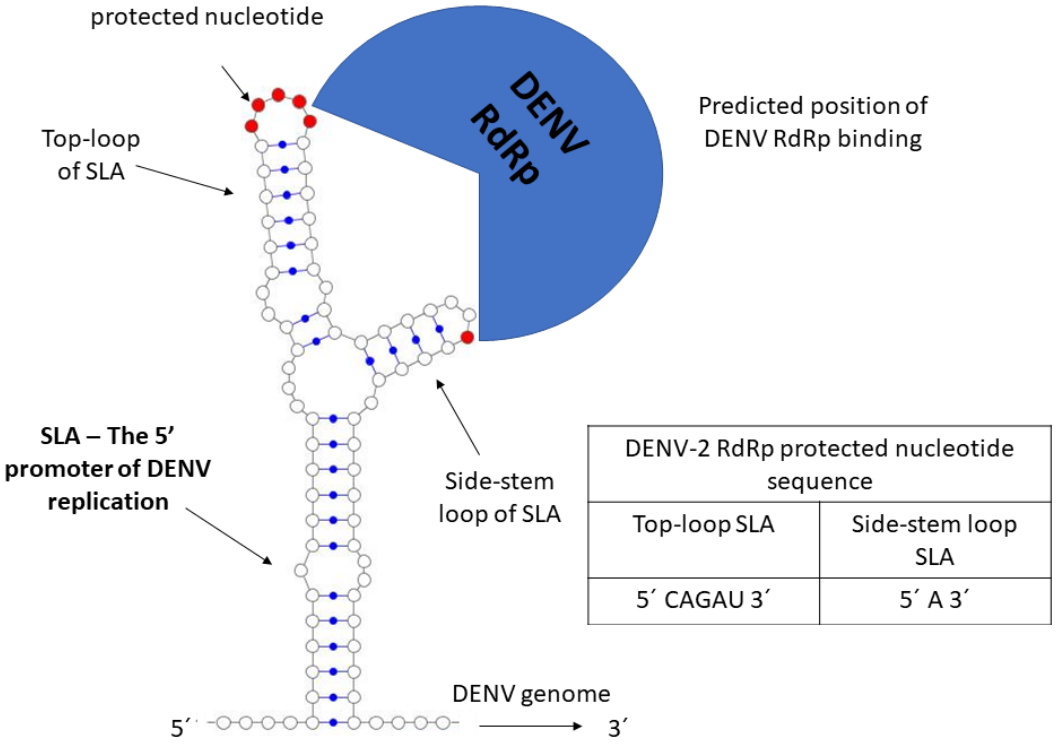

Linear
Circular
SLA
3′SL
DB2/PK
DB1/PK
5′UAR
5′CS
5′
SLB
cHP
5′DAR
DENV genome
3′CS
3′DAR
3′UAR
3′
ORF
5′NCR
3′NCR
SLA
5′-3′DAR
5′-3′UAR
5′-3′CS
5′
3′
3′SL
DB2/PK
DB1/PK

**Figure 23 A schematic representation of DENV genome cyclisation:**

*The predicted linear and circular genomic conformations of DENV genomic RNA, including the predicted dumb-bell (DB) and pseudoknot (PK) interactions of the 3´UTR. Red, yellow and blue lines represent complementary sequence positions and interactions acting to cyclise genomic RNA. Red is representative of the UAR interaction, yellow is representative of the DAR interaction and blue is representative of the CS. Stem-loops (SL) are indicated. NCR = non-coding region. Adapted from* (Tuplin, 2015)*.*

The circular genomic form is the most thermodynamically favoured conformation. Visualisation of individual full-length DENV RNA genomes using Atomic Force Microscopy (AFM) revealed that it was the physical interaction between complementary cyclisation sequences that cyclised DENV genomic RNA and that the genome was able to cyclise spontaneously in the absence of any other factor (Alvarez, Lodeiro, *et al.*, 2005; Villordo and Gamarnik, 2009; Nicholson and White, 2014).

Following extensive mutational analysis and structural stabilisation techniques utilising both mini-genome and replicon transcripts and during infection with intact cyclisation sequences, it is predicted that both linear and circular genomic conformations are required throughout DENV replication. (Alvarez, Lodeiro, *et al.*, 2005; Villordo, Alvarez and Gamarnik, 2010). Moreover, the circular genomic form is critical for negative-sense RNA synthesis. Genomic cyclisation, via complementary cyclisation sequence binding, acts to bring the selectively RdRp bound SLA into close proximity of the 3´stem-loop (3´SL) and is suggested to induce the opening of the bottom half of the 3´SL structure (selective RdRp binding to SLA is depicted in Fig. 24) (Filomatori *et al.*, 2006, 2011). This action transfers the RdRp to the conformationally prepared 3´SL, allowing for negative-sense RNA synthesis and therefore viral replication (Filomatori *et al.*, 2006; Lodeiro and Filomatori, 2009) (Fig. 25). Additionally, both SLA and the cyclisation sequences are highly conserved throughout the Flavivirus genus suggesting that genomic cyclisation, and the NS5-RNA interaction to be a vital criterion in general within the Flavivirus family (Hodge *et al.*, 2016; Villordo *et al.*, 2016).




| DENV-2 RdRp protected nucleotide sequence | |
|---|---|
| Top-loop SLA | Side-stem loop SLA |
| 5´ CAGAU 3´ | 5´ A 3´ |

**Figure 24 The binding site of DENV RdRp on SLA RNA:**

*A schematic of the position of DENV RdRp binding to the DENV 5´ promoter element SLA. The DENV RdRp is known to bind selectively and with high affinity to the top-loop and side-stem loop of the 5´ promoter SLA, as determined by footprinting and Electromobility shift assay (EMSA). EMSA protected nucleotides are depicted in red according to the work of (Filomatori et al., 2011). Figure created using the RNA structure visualisation applet VARNA.*

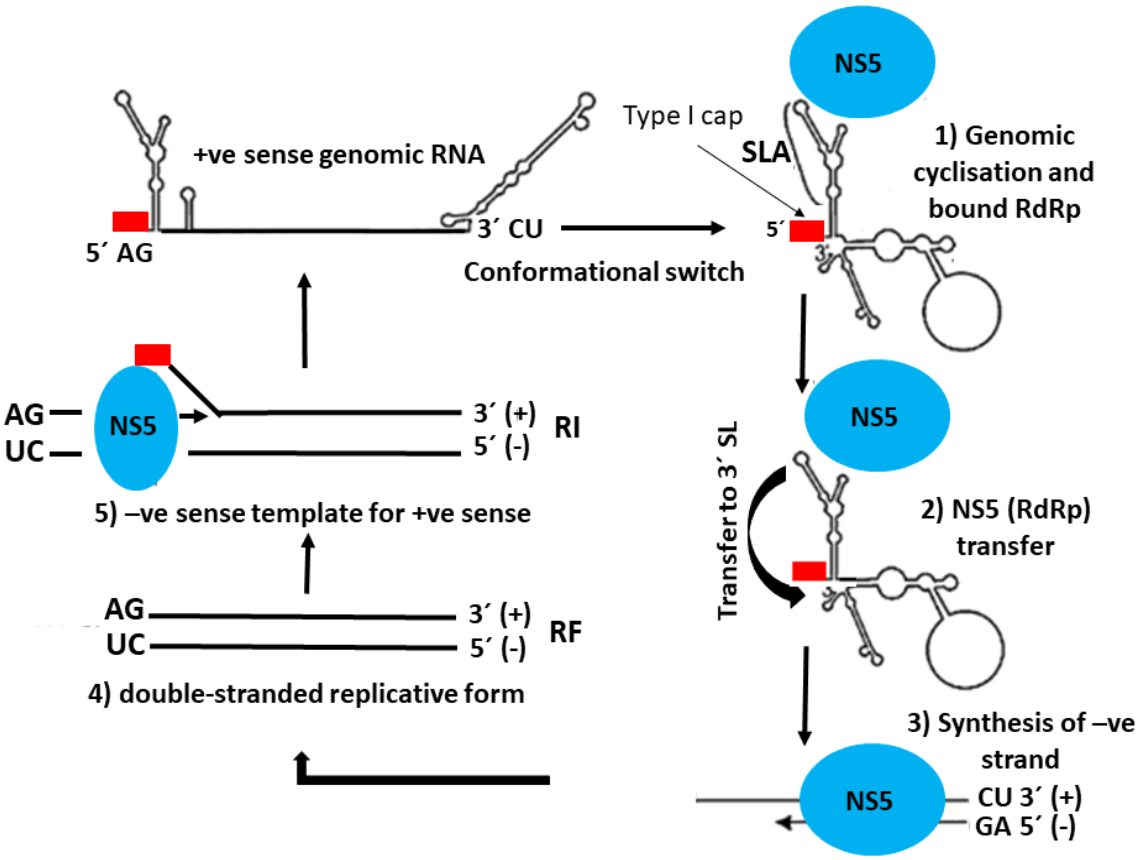



**Figure 25 A schematic of DENV replication:**

***1)*** *Capped DENV genomic RNA is cyclised by complementary RNA-RNA interactions.* ***2)*** *NS5 binds selectively via the RdRp domain to SLA, the 5`promoter of DENV replication and is transferred to the 3´SL to initiate negative-sense RNA synthesis. De novo RNA synthesis is initiated using the conserved 3´genomic end as a template.* ***3)*** *The RdRp synthesises the complete negative-sense strand.* ***4)*** *The double-stranded replicative form is synthesised, consisting of the positive-sense template genomic strand annealed to the de novo synthesised negative-sense strand (RF).* ***5)*** *The negative-sense RNA form serves as a template for the synthesis of positive-sense genomic RNA (RI). Synthesised positive-sense genomic RNA can be encapsidated or used for further rounds of host cell translation. Adapted from* (Malet *et al.*, 2008).

The linear genomic form is required for efficient replication, yet its apparent function remains unknown (Villordo, Alvarez and Gamarnik, 2010). The linear conformation may act to inhibit negative-sense RNA synthesis, providing a mechanism to control the 100:1 ratio of positive-sense versus negative-sense RNA in infected cells. Additionally, the same genomic RNA is utilised as a template for early host cell translation. It may be that the linear conformation is required during host cell translation and is stabilised by the interaction of the DENV 5´ UTR with translation initiation factors (Villordo, Alvarez and Gamarnik, 2010).

Moreover, it is unclear how the DENV capsid protein recruits genomic RNA for subsequent encapsidation and virion release. It is plausible that a certain genomic conformation is required for encapsidation and capsid recognition. This RNA conformational selectivity for encapsidation has been demonstrated in the life-cycle of retroviruses (Huthoff and Berkhout, 2001; D'Souza and Summers, 2004).

Alternatively, although evidence strongly suggests that base-pairing between the complementary 5` and 3` cyclisation sequences results in genomic cyclisation, it is proposed that such base-pairing does not necessarily require genomic cyclisation. It is suggested that the functional requirement for base-pairing between 5` and 3` cyclisation sequences could be equally fulfilled by the formation of multimeric structures in *trans.* The cyclisation sequences could be used to assemble an anti-parallel homoduplex or to assemble head-to-tail multimeric concatemers from multiple genome copies (Lott and Doran, 2013) (Fig. 26). Both the circular conformation and the formation of a genomic homoduplex would unfold SLB and sHP stem-loop structures. In contrast, both the linear genomic conformation and the formation of genomic concatamers would allow these structures to be maintained. These proposed biologically active genomic assembles, although not yet demonstrated to exist *in vitro,* would maintain replicative ability due to the proximity of 5 ´ SLA to the 3´ SL for RdRp transfer and subsequent negative-sense RNA synthesis and are depicted in figure 26.

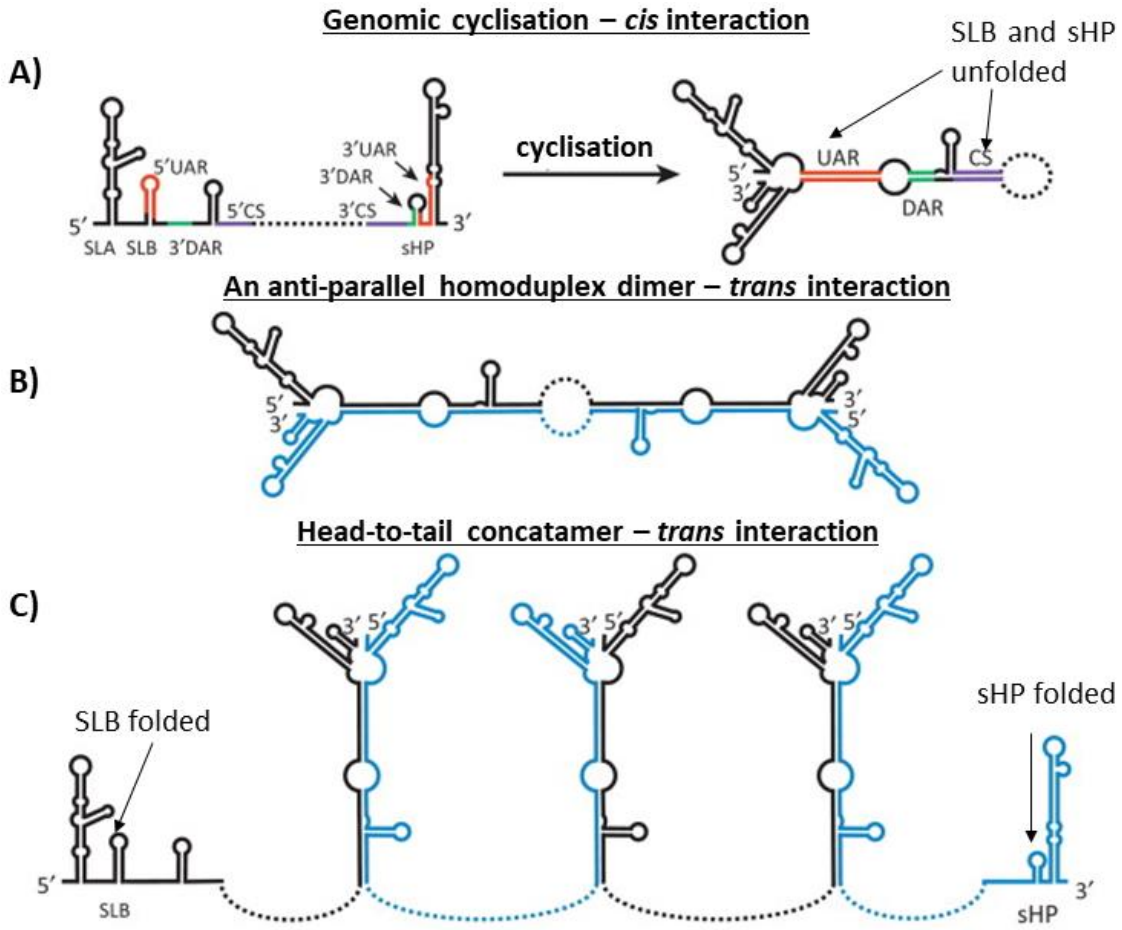



**Figure 26 The four potential biologically active forms of DENV genomic RNA:**

***A)*** *Complementary cyclisation sequences at the termini of DENV genomic RNA intramolecularly base pair forming the circular genomic form from the linear genomic form. The locations and complementarity of the cyclisation sequences are depicted. The UAR is shown in red, the DAR is shown in green and the CS in purple. The dashed line represents the intervening 10 kb of the viral genome.* ***B)*** *An anti-parallel homoduplex formed by the intermolecular base-pairing of two DENV genomic RNAs using the same complementary sequences. The top and bottom strands are identical but are shown in black and blue respectively.* ***C)*** *A head-to-tail concatemer is also possible through intermolecular base-pairing. Genomes are identical and depicted in black and blue respectively. SLB and sHP which are unfolded in either cyclised or dimeric form are able to coexist in concatemeric formation. Adapted from* (Lott and Doran, 2013).

Although single molecule cyclisation was directly visualised by AFM, the *in vitro* experimental conditions utilised were poorly representative of the physiological environment (Alvarez, Lodeiro, *et al.*, 2005; Lott and Doran, 2013). This action was suggested to strongly bias the result towards the desired observation of genomic cyclisation, for example dilution of the genome ensured only a *cis* interaction could be observed (Alvarez, Lodeiro, *et al.*, 2005). Additionally, it was shown using an *in vitro* replication assay that efficient replication initiation was observed when the DENV 5´ terminal region was added in either *cis* or *trans* (You and Padmanabhan, 1999). Furthermore, as discussed previously, deletion and compensatory mutation experiments using Flavivirus replicon systems or mini-genomes containing intact cyclisation sequences further demonstrated this base-pairing requirement for viral replication (Khromykh *et al.*, 2001; Lo *et al.*, 2003; Lott and Doran, 2013; de Borba *et al.*, 2015; Liu *et al.*, 2016). However, none of the assays excluded the possibility that the complementary cyclisation sequences incorporated in *cis* were acting in *trans* (Lott and Doran, 2013). Despite this suggestion of *trans* molecular interactions, the linear to circular transition of the DENV genome is the most experimentally evidenced and supported theory.

As evidenced, the DENV genome acquires and regulates different genomic conformations during replication as a fundamental requirement of the viral life-cycle and these genomic conformational changes have not been interrogated in the context of full-length DENV RNA, the most physiologically relevant form. The molecular constraints of why, how and what regulates the linear to circular genomic switch have yet to be examined and is therefore a focus of this project with the aim of increasing our knowledge of DENV, and therefore Flavivirus, RNA structure at the molecular level, further elucidating a significant anti-viral target.

#### 1.3.2.2 SLB and the UFS switch model

Previous *in vitro* footprinting studies carried out on DENV-2 genomic RNA suggested the importance of a poly-uracil (poly-U) rich nucleotide region between the SLA and SLB elements of the 5´terminus of DENV genomic RNA. The poly-U nucleotide tract is suggested to function as a spacer/hinge, providing flexibility for the SLA promoter element to transfer the bound RdRp to the 3´SL (Filomatori *et al.*, 2011) and is consequently the reason why most DENV linear structural

presentations utilise this structural model in publications. However, opinion remains divided regarding the function of this sequence and the stem-loop structures downstream of SLA, such as SLB and cHP, during DENV replication (Fig. 27).

A novel RdRp-SLA binding requirement model was presented by Liu and colleagues in 2016. This model proposes the strict structural requirement of the base of SLB, including the poly-U tract, which is hybridised in this format, in DENV replication. This region has been named the 5´UAR flanking stem (UFS) because of its location which constrains the 5´UAR and SLB between its proposed two strands (Fig. 27). It is suggested that the UFS regulates the binding of DENV RdRp to SLA by dynamically regulating DENV genomic conformation during viral replication (Liu *et al.*, 2016).

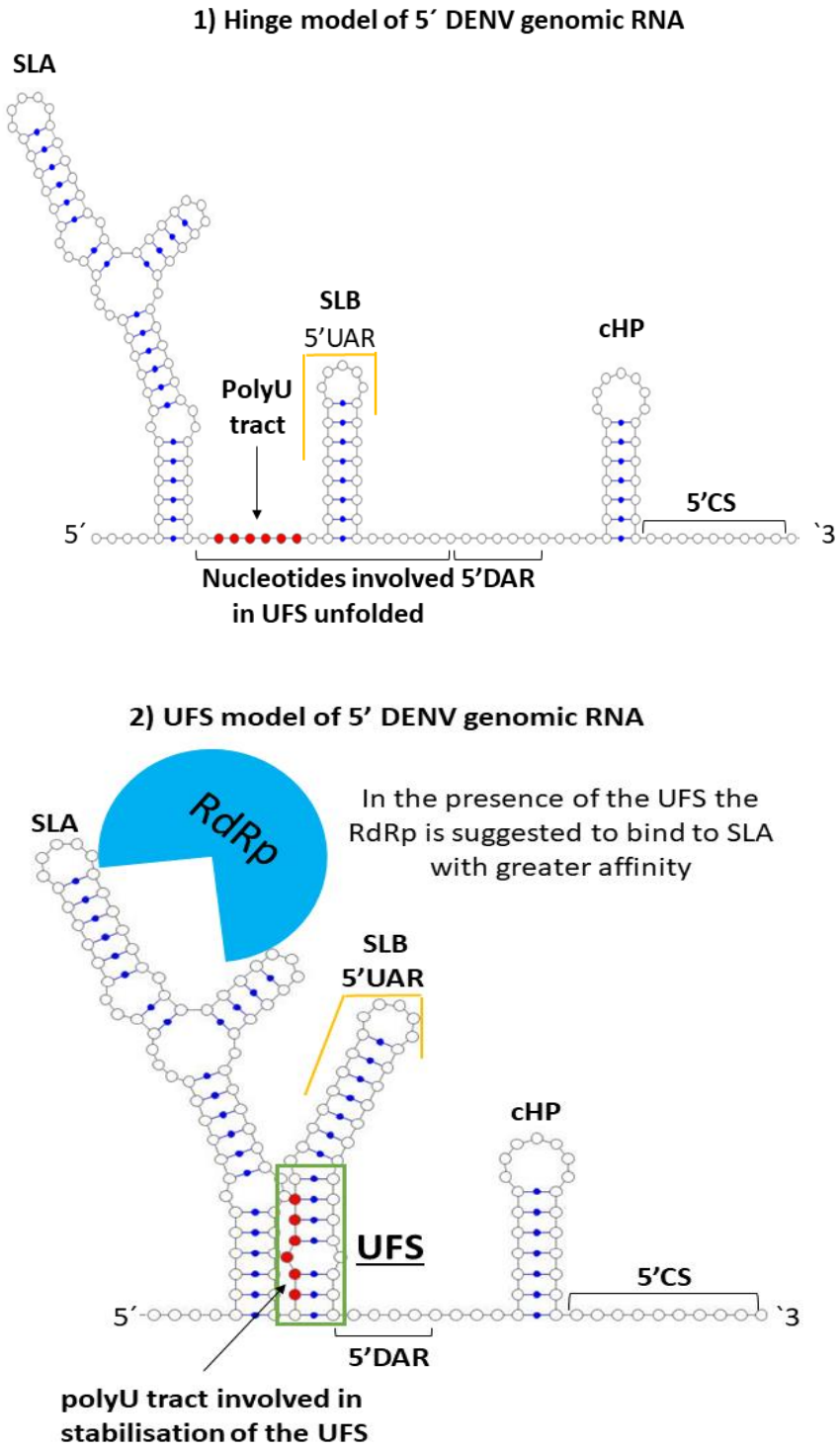



**Figure 27 Differing conformations of the 5´ terminus of DENV genomic RNA:**

***1)*** *The hinge model – the polyU tract is available as a spacer/hinge during DENV replication.* ***2)*** *The UFS model – the polyU tract is hybridised due to the proposed formation of the UFS. The polyU tract is indicated in red. The position of cyclisation sequences are highlighted. The presence of the UFS is suggested to increase the affinity of the RdRp to SLA. The UFS is highlighted in the green box. Figure was created using VARNA.*

*In vitro* SHAPE mapping (methodology further discussed in Chapter 4 section 4.1) approaches were utilised, alongside a bioinformatic approach, to infer the presence of the DENV UFS structurally, in the linear genomic conformation, conserved throughout the four serotypes *in vitro* utilising truncated genomic transcripts. However, it is of note that the DENV-2 UFS was inherently unstable in this context (Liu *et al.*, 2016).

The function of the UFS in DENV replication was interrogated using mutational analysis of the UFS during full viral replication whereby it was found that over-stabilisation of the UFS hindered genomic cyclisation and therefore viral replication. Furthermore, the influence of the UFS on RdRp affinity for SLA was tested using EMSA, utilising mini-genome templates, able to cyclise, or truncated transcripts. It was found that the presence of the UFS increased the affinity of the RdRp-SLA interaction. A decrease in this apparent affinity was observed in the absence of the UFS (Liu *et al.*, 2016). These findings subsequently led to the proposal of the UFS switch model for DENV replication via an ‘on’ ‘off’ mechanism.

The UFS switch model proposes that when viral RNA is uncoated and released into the host cell cytoplasm, the genome is translated resulting in the production of a sufficient level of viral proteins and the formation of the replication complex. This process leads to the initial binding of RdRp to SLA and therefore the initiation of negative-sense RNA synthesis. Completion of negative strand synthesis generates the double-stranded replicative intermediate which serves as a template for the synthesis of positive sense genomic RNA. During the elongation of the nascent positive strand, the 5´ terminus of the parental positive strand is displaced first, and it is suggested that the UFS duplex folds at that point and is in the ‘on’ state. This ‘on’ state of the UFS subsequently facilitates the recruitment of NS5 via the RdRp domain to the available 5´ genomic terminus. After completion of nascent chain synthesis, the NS5 bound genomic RNA is suggested to undergo genomic cyclisation, setting the element to the ‘off’ state. Hybridisation of complementary cyclisation sequences unwinds the UFS, and therefore SLB, decreasing the affinity of the RdRp for SLA and promoting its translocation to the 3´SL, initiating the next round of negative sense-RNA synthesis (Liu *et al.*, 2016). A schematic depiction of this model is detailed in figure 28.

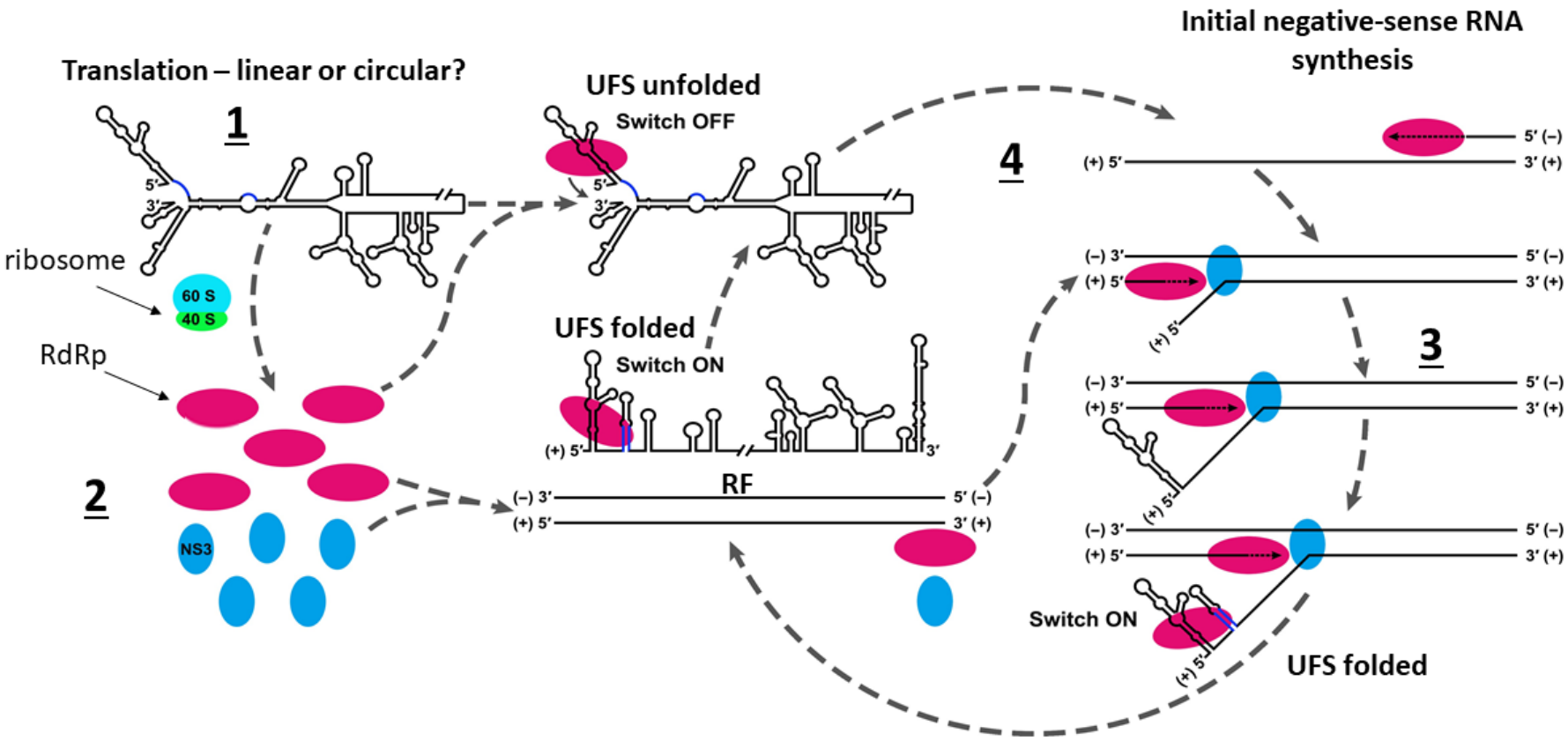

Translation – linear or circular?
1
ribosome
60 S
40 S
RdRp
2
NS3
UFS unfolded
Switch OFF
UFS folded
Switch ON
RF
(–) 3′
(+) 5′
5′ (–)
3′ (+)
4
Initial negative-sense RNA synthesis
3
Switch ON
UFS folded

**Figure 28 A schematic describing the function of the UFS switch model – a proposed mechanistic model of Flavivirus replication:**

***1)*** *Following entry, the viral genome is released into the cytoplasm followed by host cell translation producing a sufficient level of viral proteins for viral replication within the replication complex.* ***2)*** *Accumulation of viral replication proteins (NS3 and NS5 shown) culminates in the formation of an RNA-viral replicase complex initiating negative-sense RNA synthesis generating the double-stranded replication form (RF).* ***3)*** *Positive sense-RNA synthesis is initiated by the replicase complex. During this process the 5´ termini of positive genomic RNA is released inducing the formation of the UFS and a free NS5/replicase complex is able to bind the 5´ displaced end of the positive-sense genomic RNA.* ***4)*** *Following the completion of positive strand RNA synthesis, the released positive sense-single stranded RNA genome is suggested to cyclise and the next round of negative-sense synthesis is able to commence. Adapted from* (Liu *et al.*, 2016)

Consequently, this proposed structurally dynamic model of DENV replication may be of great significance throughout the viral life-cycle. The UFS switch model may allow for the maintenance of the appropriate level of negative-sense genomic RNA, held within the replication complex and ensure an overall dynamic balance of DENV replication.

The DENV-2 UFS was noted to be structurally unstable and may present a novel DENV replicative model with altered replicative kinetics and therefore the interrogation of this serotype, at the RNA structural level, during viral replication *in vitro* and during virus infection is an exciting avenue explored in this project.

#### 1.3.2.3 cHP

It has been suggested that the cHP stem-loop element, a 5´stem-loop within the capsid-coding region adjacent to SLB, is a critical determinant of both DENV and WNV RNA synthesis. It was shown, utilising a combination of reporter replicons and infectious clones that an intact cHP structure is required for efficient RNA synthesis, acting independently of its sequence (Clyde, Barrera and Harris, 2008). The cHP may function structurally within the viral life-cycle as part of the overall topology of the 5´ terminus and may even act to stabilise the circular conformation of the viral genome (Clyde, Barrera and Harris, 2008).

#### 1.3.2.4 RNA structure and the host environment during DENV replication

Members of the Flavivirus genus cycle between both mammalian and insect hosts. Establishing an efficient infection in such different physiological environments, with different selection pressures, demands these viruses to have evolved strategies for rapid adaptation. However, the mechanistic basis behind this clear adaptation process remains to be truly determined at the genomic RNA structural level.

The mosquito-human adaptation process has been studied previously utilising deep sequencing and RNA structure analysis (conducted using SHAPE mapping) together with fitness evaluation, acting to define the process of host specialisation defined by RNA elements within the 3´UTR. In one such study, it was found that single nucleotide substitutions within the 3´sHP structure, which maintained the structure of the stem-loop, did not affect replication in mammalian cells but abolished replication in mosquito cells indicating a sequence dependence, and

possible interaction, in mosquito cells which is not required during mammalian DENV infection (Villordo and Gamarnik, 2013). DENV adaptation to mosquito or mammalian cells and subsequent deep-sequencing showed a strong selection of specific mutations in the viral 3´UTR in each host, while the *cis*-acting structural elements present at the 5´ terminus of genomic RNA (SLA,SLB,cHP) remained unchanged. Additionally, a feature of the 3´UTR of all Flaviviruses is the presence of short direct repeats and long RNA element duplications. It is suggested that these stem-loop duplications incur a fitness advantage in insect cells whereas they are redundant in mammalian replication (Villordo *et al.*, 2015).

Therefore, the host environment and its clear structural implications demand to be elucidated during the structural interrogation of DENV replication and was subsequently considered during this presented study of DENV replication.

## 1.4 Project aim and objectives

As evidenced, the RdRp-SLA interaction is a fundamental requirement of DENV replication and subsequently presents an understudied, valid anti-viral target, desperately lacking in the treatment of DENV infection.

This study aims to structurally interrogate this essential protein-RNA interaction using DENV-2 as a model serotype.

Objectives:

**1)** To determine the structure of DENV-2 RdRp bound and unbound to SLA promoter RNA by X-ray crystallography.

**2)** To identify the conformational requirements of the full-length DENV-2 genome in the presence of *trans*-activating factors, during active virus replication and in the replication complex, considering mammalian and insect host environments using both *in vitro* and intracellular SHAPE.

# Chapter 2 Materials and methods

## 2.1 Materials

### 2.1.1 Plasmids and vectors

The DENV-2 New Guinea C (NGC) (accession number AF038403) replicon GFP expressing and full-length infectious plasmids were kindly provided by Dr. Andrew Davidson (University of Bristol). The NS5 and RdRP ORFs were sub-cloned into the pET-28a-his-SUMO expression vector. The identity of all sequences was confirmed prior to experimentation via DNA Sanger sequencing.

### 2.1.2 Bacterial strains

Plasmid DNA constructs were amplified, and recombinant DENV-2 NS5 or RdRp, was expressed following transformation of competent *Escherichia coli (E.coli)* cells. High efficiency DH5α *E.coli* (NEB) were used for the initial transformation of ligation reactions. For routine re-transformation procedures sub-cloning efficiency DH5α *E.coli* (NEB) were utilized. For protein expression, the BL21-Gold derivative, BL21-CodonPlus (DE3)-RIPL expression strain for optimised isopropyl β-D-1-thiogalactopyranoside (IPTG) inducible protein expression of DENV-2 RdRp was used.

### 2.1.3 Continuous mammalian cell lines

HEK293T (Human Embryonic Kidney cell origin, expressing a mutant SV40 large T antigen (DuBridge *et al.*, 1987)) cells were used for the generation of a DENV-2 replicon expressing stable cell line and for host cell S100 cytoplasmic enrichment. BHK-21 (Baby Hamster Kidney fibroblast cell origin (Macpherson and Stoker, 1962)) cells were used for DENV infection and titre determination via plaque assay. Cells were routinely cultured, harvested by trypsinisation and split when 80% confluency was reached, in complete media consisting of Dulbecco's Modified Eagle Medium (DMEM) high glucose (Sigma Aldrich), 10% foetal bovine serum (FBS) and 1% Penicillin Streptomycin (P/S) at 37 °C and 5% $CO_2$.

### 2.1.4 Continuous mosquito cell lines

C6/36 (Aedes albopictus origin (Igarashi, 1978)) cells were used for DENV propagation. Cells were routinely cultured, harvested by scraping and split at 80% confluency, in Lebovitz's L-15 medium containing 10% FBS, 10% Tryptose Phosphate Broth and 1% P/S at 28 °C.

### 2.1.5 Primers

Oligonucleotide primers were synthesised for described experimental procedures (Sigma Aldrich). The oligonucleotide sequences are provided as follows in 5´-3´ orientation:

**Table 1 Oligonucleotides**

| | **Forward** | **Reverse** |
|---|---|---|
| **DENV-2 NS5** | GGAACTGGCAACATAGGAGA | CTACCACAGGACTCCTGCCT |
| **DENV-2 RdRp** | ATCGGAATTGAAAGTGAGAT | CTACCACAGGACTCCTGCCT |
| **5´ 160 nt** | GCTAATACGACTCACTAT | GCATATTGAAAGGCGTATTTC |
| **5´ 500 nt** | GCTAATACGACTCACTAT | |
| **DENV-2 NS3** | GCTGGAGTATTGTGGGATGTCC | CTTTCTTCCAGCTGCAAACTCC |
| **ΔSLA** | TAATACGACTCACTATAGCAGATCTCT | TTAATGGTCCTCGTCCCTGCAGC |
| **3´UTR** | TAATACGACTCACTATAGCACTGGAAG | GCCGCTCTAGAACCTGTTG |
| **SHAPE** | | HEX/FAM TGTACAGTCGACACGCGGTTTC |

### 2.1.6 Antibodies

Antibodies were used for protein identification via western blot. Primary and secondary antibodies utilized, as per manufacturer's instructions, are listed below:

**Table 2 Primary and secondary antibodies for western blot**

| Target | Primary antibody | Secondary antibody |
|---|---|---|
| **DENV-2 NS5** | Rabbit PA5-32200 (Thermo Scientific) | Donkey anti-Rabbit 800 nm (Li Cor) |
| **β-Actin** | Mouse A1978 (Sigma Aldrich) | Donkey anti-Mouse 680 nm (Li Cor) |
| **NUP98** | Rabbit C39A3 (Cell Signalling Tech.) | Donkey anti-Rabbit 800 nm (Li Cor) |

### 2.1.7 Buffer, media and stain preparation

Listed as appropriate in Appendix I.

## 2.2 Molecular biological methods

### 2.2.1 Agarose gel electrophoresis

The integrity, purity and size of plasmid DNA, PCR products, restriction digests and ligation reactions were determined by native agarose gel electrophoresis using 1% agarose gels (0.5 g analytical grade agarose (Sigma Aldrich)) in 50 ml 1 X TAE with SYBR Safe DNA stain (Life Technologies at 1: 10,000 dilution). Samples were combined with a final concentration of 1 X DNA Gel loading dye (NEB). Samples were loaded alongside the ladder of choice for analytical size comparison (either 1 Kb DNA ladder (NEB) or 1 Kb exACTGene ruler (Fisher)) and run at 90 V for 45-60 min in 1 X TAE buffer. UV transillumination was used for DNA visualisation.

### 2.2.2 RNA quality analysis using denaturing formaldehyde MOPS agarose gel electrophoresis

Products of *in vitro* RNA transcription reactions and Trizol (Thermo Scientific) extracted RNA samples were purity and quality analysed using denaturing formaldehyde agarose gel electrophoresis in 3-(N-morpholino)propanesulfonic acid (MOPS) buffer. A 1% agarose gel was prepared using analytical grade agarose (Sigma Aldrich) and 1 X MOPS in nuclease free water. 1.4 ml 37% Formaldehyde was added following the melting of the agarose. RNA samples for analysis were combined to a concentration of 1 X denaturing RNA loading dye

(NEB) and denatured at 65°C for 10 min. Samples were loaded alongside either a Millennium or Century RNA ladder (Ambion) as appropriate for relative size comparison and electrophoresis performed at 90 V for 1 hr in 1 X MOPS buffer. Blue light transillumination was used for RNA visualisation.

### 2.2.3 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

5 ml SDS-PAGE gels were made with a 7%, 10% or 12% resolving gel (10 ml) where appropriate (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). Prior to analysis, protein samples were mixed 1:1 with 2 X denaturing sample buffer. Protein samples were denatured by heating to 95 $^oC$ for 5 min prior to loading on to gels. Samples were loaded alongside 5 µl colour pre-stained protein standard, broad range marker (NEB) for size analysis. Electrophoresis was performed in SDS running buffer for 50-60 min at 200 V.

### 2.2.4 Protein visualisation by Coomassie stain

Proteins previously resolved by SDS-PAGE electrophoresis were visualized directly following incubation in Coomassie stain at room temperature for 15 min followed by de-staining for 16 hrs at room-temperature in de-stain solution. Gels were rehydrated in $ddH_20$ prior to imaging.

### 2.2.5 Western blot analysis

Proteins for identification and analysis were transferred from SDS-PAGE gels to fluorescence compatible polyvinylidene fluoride membranes (Immobilon-FL Transfer membranes; Millipore) using a Trans-blot semi-dry cell (Bio-Rad) in Towbin buffer for 1 hr at 15 V. The membrane was blocked for 1 hr at room temperature in 50% (v/v) Odyssey blocking buffer- Phosphate buffered Saline (PBS) Licor in PBS. Blocking buffer was then exchanged with appropriately diluted antibody in 75% (v/v) PBS, 25% (v/v) Odyssey blocking buffer and incubated for 1 hr at room temperature or overnight at 4 $^oC$. Membranes were then washed with PBS before incubation with the appropriate fluorescently tagged secondary

antibody for 1 hr at room temperature in the dark. Secondary antibodies were directly visualised using a LiCor Odyssey Sa Infrared imaging system (LiCor).

### 2.2.6 Restriction endonuclease digestion

Plasmid DNA and Polymerase Chain Reaction (PCR) products were subjected to restriction enzyme digestion for cloning, plasmid linearization in preparation for *in vitro* transcription and for diagnostic purposes. For cloning, the pET-28a-SUMO expression vector was digested with *Bam* HI and *Xho* I (New England Biolabs (NEB), concentration as per manufacturer's instructions), in a 200 µl total volume containing 20 µg of DNA, 1 X CutSmart buffer (NEB) and nuclease free water (Ambion) and incubated overnight at 37 °C. The digested backbone was then dephosphorylated using 3 µl Calf Intestinal Alkaline phosphatase (10,000 u/ ml, NEB) and incubated at 37°C for 1 hr. NS5 and RdRp ORF PCR products for sub-cloning were digested in a total volume of 100 µl containing 50 µl PCR reaction, 1 X CutSmart buffer, *Bam* HI, *Xho* I (NEB, concentration as per manufacturer's instructions) and nuclease free water and incubated for 3 hrs at 37 °C. For plasmid linearization in preparation for *in vitro* transcription, 20 µg DNA was linearized in a 200 µl total volume containing 1 X CutSmart buffer (NEB), *Xba* I (NEB) and nuclease free water. All digestion reactions were purified, following agarose gel electrophoresis and gel extraction if required, utilising a Wizard® SV Gel and PCR Clean-Up System (Promega) as per manufacturer's instructions.

### 2.2.7 The *in vitro* transcription of capped and uncapped RNA transcripts

Capped full-length infectious DENV-2 RNA was synthesised in preparation for transfection to produce infectious DENV-2 virus. Following linearization of the DENV-2 plasmid containing the full-length infectious clone, capped RNA was synthesized utilizing a mMessage mMachine T7 transcription kit (Thermo Scientific) following manufacturer's protocol.

Uncapped RNA was synthesized for use in SHAPE mapping, polymerase activity assays and EMSA competition analysis. Uncapped RNA was produced, following plasmid linearization or PCR amplification and purification, utilizing a T7 RiboMax™ Express Large Scale RNA Production System (Promega).

All transcripts were purified using Lithium Chloride (Thermo Scientific) precipitation as per manufacturer's instructions. Following purification, a Nanodrop 1000 (Thermo Scientific) was used to determine the concentration and denaturing formaldehyde agarose gel electrophoresis was used to determine the purity and quality of the sample.

### 2.2.8 Polymerase Chain Reaction

PCR was utilized to amplify both the DENV NS5 and RdRp ORFs from the DENV replicon plasmid to clone into the pET-28a-SUMO expression vector for protein purification. Additionally, PCR was used to generate 5´160 nt, 5´500 nt and electromobility shift assay (EMSA) template transcripts for use as templates for *in vitro* transcription for polymerase activity determination, *in vitro* SHAPE experimentation and EMSA competition study respectively.

For protein expression ORF PCR amplification, primers were designed to incorporate terminal restriction endonuclease sites to facilitate sub-cloning into the pET-28a-SUMO expression plasmid. PCR reactions were performed in a 50 µl total volume containing 200 ng template DNA, 2.5 µl forward primer, 2. 5 µl reverse primer (both 10 µM), 1 µl 10 mM dNTPs, 10 µl 5 X GC buffer (NEB) and 1 µl Phusion DNA polymerase (NEB). Reaction cycles were performed in a thermocycler (Eppendorf); 30 seconds 98°C denaturation; 30 cycles of 98°C, 10 seconds - 55°C, 30 seconds and a 1.5 minute 72°C elongation. A final 10 min extension at 72°C was performed before being cooled to 4 °C.

For amplification of the 5´160 nt, 5´500 nt and EMSA transcript templates forward primers were designed to incorporate a T7 promoter sequence for *in vitro* transcription. PCR reactions were performed as described above in a total reaction volume of 50 µl. Reaction cycles were performed in a thermocycler (Eppendorf); 30 seconds 98°C denaturation; 25 cycles of 98 °C, 10 seconds, 55°C 30 seconds and a 2.5 minute 72°C elongation. A final 10 min extension at 72°C was performed before being cooled to 4 °C. All PCR products were purified utilising Wizard® SV Gel and PCR Clean-Up System (Promega) as per manufacturer's instructions and analysed by agarose gel electrophoresis and eluted in 20 µl nuclease free water.

### 2.2.9 Bacterial transformation

For routine plasmid amplification low efficiency *E.coli* DH5α cells were transformed according to manufacturer's protocol. Briefly, approximately 50 ng of plasmid DNA or 1 µl of ligation reaction product was mixed with the DH5α strain of choice and incubated for 30 min on ice; followed by a 30 second 42°C heat shock and then recovered on ice for 2 min. 500 µl of Super Optimal broth with Catabolite repression (SOC) was added to the transformation mix and incubated in a shaking incubator at 37°C for 1 hr. 200 µl of the transformation mixture was then spread onto Luria-Bertaini (LB) and 1.5% bacteriological agar plates containing the appropriate antibiotic and incubated for 16 hrs at 37 °C.

### 2.2.10 Ligations

Following restriction endonuclease digestion, PCR amplified and restriction endonuclease digested NS5 and RdRp ORF products were ligated into the prepared pET-28a-SUMO backbone for protein purification. Ligation reactions were performed at a 5:1 insert:vector molar ratio. The ligation reactions contained 100 ng backbone DNA, the appropriate amount of insert, 2 µl T4 ligase (NEB), 1 X T4 ligase buffer (NEB) to a total volume of 20 µl. Ligation reactions were incubated at room-temperature for 3 hrs and transformed into high-efficiency chemically competent *E.coli* following standard protocol.

### 2.2.11 Plasmid DNA amplification

Overnight cultures, grown at 37°C from either transformed colonies, glycerol stocks or a starter culture for large scale protein expression, were pelleted by centrifugation at 4000 *xg* at 4°C for 20 min. Plasmid DNA was isolated and purified using either a Maxi or Midi preparation kit (Thermo Scientific) following manufacturer's instructions. Following purification, a Nanodrop 1000 (Thermo Scientific) was used to quantify the purity and concentration of the isolated DNA by spectrophotometry.

### 2.2.12 Reverse transcription PCR (RT-PCR)

RT-PCR was utilized to confirm the successful establishment of a DENV-2 stable cell line and for confirmation of DENV infection. First strand cDNA synthesis was performed using Random hexamer primers (NEB) as per manufacturer's

instructions (SuperScript III First-strand synthesis system for RT-PCR). PCR reactions were performed using 10 µl GC buffer, 1 µl dNTP mix, 1 µl NS3 forward primer, 1 µl NS3 reverse primer (both 10 mM), 1 µl (2 U) Phusion polymerase, 2 µl cDNA from the first-strand synthesis reaction and nuclease free water up to a total volume of 50 µl. Reaction cycles were performed in a thermocycler (Eppendorf); 30 seconds 95°C denaturation; 30 cycles of 95°C 30 seconds, 55°C 30 seconds and a 2 minute 72°C extension. A final 10 min extension at 72°C was performed before being cooled to 4 °C. All PCR reactions were purified utilising a Wizard® SV Gel and PCR Clean-Up System (Promega) as per manufacturer's instructions and analysed by agarose gel electrophoresis and eluted in 20 µl nuclease free water.

### 2.2.13 Establishment of DENV-2 stable cell line

A DENV-2 stably expressing replicon cell line was established for S100 protein extraction and for use in intracellular SHAPE experimentation. 8 µg of capped *in vitro* transcribed DENV replicon RNA was diluted in 200 µl Opti-MEM (Gibco). 12 µl of Lipofectamine® 2000 reagent was diluted in 200 µl Opti-MEM. Both solutions were incubated at room temperature for 5 min. The solutions were then combined and incubated for a further 20 min at room-temperature. The final solution was then divided in half and added to 2 mls of HEK293T cells (400,000 cells/ml) in DMEM and 10% FBS. The suspension was further incubated for 10 min at room-temperature prior to plating in a 12 well format. The plate was incubated for 6 hrs at 37°C before the media was replaced with complete media. The transfected cells were incubated for 72 hrs at 37°C prior to GFP fluorescence visualisation to confirm successful transfection. Following GFP visualisation, the transfected HEK293T cells were expanded into a 10 $cm^3$ petri dish and maintained under Puromycin selection (3.5 µg/ml) in conditioned growth media. DENV-2 non-structural protein expression was further confirmed using western blot and RT-PCR.

### 2.2.14 Isolation of HEK293T DENV-2 Replicon Cell Cytosol (S100) protein extract

Cytosolic cell extract was harvested from cultured HEK293T DENV-2 Replicon cells for use in SHAPE mapping following the method of Zhou *et al* (Zhou *et al.*,

2000). In brief, 6 confluent T175 flasks were washed with PBS, trypsinized and pelleted by centrifugation at 1000 *xg*. The pelleted cells were washed twice with ice cold PBS and resuspended in 5 volumes of Buffer D and incubated on ice for 20 min. The cells were homogenised using approximately 6 strokes of a Dounce homogenizer and pelleted by centrifugation at 1000 *xg* for 10 min at 4 °C. The supernatant was removed and pelleted again at 100,000 *xg* for 1 hr at 4 °C. The supernatant was aliquoted and stored at – 80°C prior to use.

### 2.2.15 RNA transfection and generation of infectious DENV-2

Infectious DENV-2 virus was generated for use in intracellular SHAPE. Capped full-length infectious DENV-2 RNA was generated as previously described. BHK-21 cells were seeded in a 6 well plate prior to transfection at a density of 0.15 x $10^6$ cells/ ml and incubated until 60- 80% confluency was reached at the time of transfection. Duplicate wells of cells were transfected with 3.5 µg DENV RNA using Lipofectamine® 2000 (Thermo Scientific) as per manufacturer's instructions. The remaining cells were used as a Lipofectamine 2000® only negative control and incubated for 3 hrs at 37 °C. The transfection mix was replaced with complete growth medium and incubation at 37°C continued for 72 hrs. The supernatant was harvested, and cell samples lysed in 1 X Passive lysis buffer (Promega (PLB)) for western blot analysis, and used to infect a T175 flask of C6/36 cells and incubated for a week at 28 °C. The supernatant was harvested, aliquoted and frozen at – 80 °C. The virus stock was titred using a plaque assay as described in section 2.2.16.

### 2.2.16 Titration of infectious DENV-2 and multiplicity of infection determination

Generated infectious DENV-2 was titred to determine the plaque forming units/ ml (PFU/ml) of the virus stock for use in intracellular SHAPE. 24 hrs prior to infection, BHK-21 cells were plated at a density of 200,000 cells/ ml in a 6 well plate and incubated until 80% confluency was reached. The virus stock was thawed at room temperature and 10-fold dilutions were prepared in complete media. Cell monolayers were washed with PBS and 200 µl of virus dilution was added to each well in order. The plates were then rocked to ensure the cell monolayer was covered in the diluted virus and incubated at 37°C for 1 hr to allow virus entry. Following incubation, the diluted virus was removed from the cell monolayer and

replaced with 50: 50 1.6% Methyl cellulose (MC) to complete growth medium and incubated for 6 days at 37°C until plaque formation. The MC/media solution was removed and fixed in 10% Formaldehyde in water. The formaldehyde solution was removed, the cell monolayer stained with 0.25% Crystal violet and washed gently with $ddH_20$ until plaques became visible. The plaques were then counted, using an appropriate dilution, and titre determined using the below equation:

$$\text{PFU/ ml} = \frac{Number\ of\ plaques\ observed}{dilution\ factor\ (10^{-x}) \times volume\ of\ virus\ added\ (ml)}$$

Multiplicity of infection (MOI) is calculated to determine the number of virions that are added per cell during an infection, a means of standardizing experimentation. MOI is determined using the equation below:

$$\text{MOI} = \frac{pfu\ per\ ml\ used\ for\ infection}{number\ of\ cells}$$

## 2.3 Protein expression and purification in *E.coli*

### 2.3.1 Bacterial expression strain selection

Both DENV-2 NS5 and RdRp were tested for soluble protein expression using a range of *E.coli* expression strains including: BL21- Gold, BL21- Star, BL21-Rosetta 2 and BL21-CodonPlus (DE3)-RIPL (Agilent) and the most successful used for routine protein production. Briefly, 2 colonies per construct were picked and used to inoculate 10 ml LB starter cultures, containing the appropriate antibiotic and incubated for 16 hrs at 37°C whilst shaking. The bacterial culture was pelleted by centrifugation at 4000 *xg* and resuspended in 10 ml of fresh LB containing the appropriate antibiotic. Cultures were further incubated whilst shaking at 37°C until an optical density (OD 600nm) of 0.6- 0.8 was reached, analysed using a spectrophotometer (Bio-rad) and divided in half. Half was induced with 1 M IPTG and the other utilised as an uninduced control. The cultures were then further incubated at 37°C whilst shaking for 16 hrs. The cultures were then pelleted, resuspended in 1 X denaturing loading buffer and stored at – 20°C prior to analysis via SDS-PAGE gel electrophoresis and Coomassie stain.

### 2.3.2 Optimisation of IPTG concentration and temperature for induction

Optimisation of IPTG concentration and temperature of induction was carried out to increase protein yield for crystallography. Optimisation was carried out using the DENV-2 NS5 expression construct and the optimised protocol followed for the expression of RdRp. Following construct transformation into BL-21 CodonPlus (DE3)-RIPL (Agilent) *E.coli*, a 20 ml starter culture was inoculated and grown whilst shaking at 37°C for 16 hrs. The culture was pelleted by centrifugation at 4000 *xg* and resuspended in fresh LB containing Kanamycin (50 µg / ml) and grown at 37°C until an OD of 0.6 – 0.8 was reached. 1 ml aliquots were taken and induced at varying IPTG concentrations (0.2 – 1 mM) and temperatures (15 – 25 °C) and grown whilst shaking for 16 hrs. Following incubation, the cultures were pelleted by centrifugation at 4000 *xg* and resuspended in 1 x denaturing loading buffer and stored at – 20°C prior to analysis by SDS-PAGE gel electrophoresis and western blot and the best conditions selected for protein purification procedures.

### 2.3.3 Bacterial induction and lysis for protein purification

Both DENV-2 NS5 and RdRp proteins were optimised for expression and purified in the same manner with the aim of producing a high concentration of pure protein for use in crystallographic and SHAPE procedures. Due to the fundamental instability of DENV-2 NS5, the purification of RdRp is described throughout.

Freshly transformed BL21-CodonPlus (DE3)-RIPL *E.coli* were used to inoculate a 10 ml LB starter culture containing Kanamycin (50 µg / ml) and grown for 16 hrs whilst shaking at 37 °C. The starter culture was used to inoculate 2 L of LB broth containing the appropriate antibiotic and incubated at 37°C until an OD of 0.6-0.8 was reached. The cultures were induced for protein expression using 0.2 mM IPTG, allowed to cool to room temperature, and incubated at 15°C for 16 hrs whilst shaking. The culture was pelleted by centrifugation at 4000 *xg* for 20 min and stored at – 20°C until lysed for protein purification.

The pelleted cells were thawed and resuspended in lysis buffer (10 ml / L of culture). Cells were incubated, whilst agitated, in lysis buffer at 4°C for 1 hr followed by a sonication procedure, on ice, of 10 cycles of 10 seconds 10 µm amplitude

bursts, and 10 seconds off. Lysates were then clarified at 25, 000 *xg* at 4°C for 45 min. The clarified supernatant containing the soluble expressed protein was retained for RdRp purification via Nickel-affinity chromatography and size-exclusion chromatography.

### 2.3.4 Nickel affinity chromatography

Nickel (Ni) affinity chromatography was used as an initial stage to isolate and purify the expressed, his-tagged RdRp protein from contaminants. 1 prepacked 1 ml High Performance HisTrap™ Ni column (GE Healthcare) attached to a peristaltic pump (Bio-rad), at 4 °C, was used per 2 L of initial culture, and washed with 5 column volumes (CV) $ddH_20$ and equilibrated with 5 CV of binding buffer. The clarified supernatant was allowed to flow- through at a rate of 1 ml/ minute. The column was then washed with 10 CV binding buffer. 10 ml of wash buffer was prepared of identical composition to the binding buffer but containing 50 mM Imidazole and allowed to flow over the column. In order to remove any RNA bound to the target a 2 M NaCl, 20 mM Tris-HCl wash was included. The column was then back equilibrated with another 10 CV 50 mM imidazole wash. Elution of bound protein was performed with 2 CV elution buffer containing 20 mM Tris-HCl, 300 mM NaCl and 120, 300 and 500 mM Imidazole respectively. The eluted fractions were pooled prior to removal of the his-SUMO tag using SUMO protease and dialysed against 5 L gel filtration buffer for 16 hrs at 4 °C. The pooled fractions were then further purified by size exclusion chromatography. All buffers were prepared at 4 °C. If required, samples of all fractions were taken and analysed by SDS-PAGE electrophoresis and visualised by Coomassie stain.

### 2.3.5 Size exclusion chromatography

Size exclusion chromatography was used to further purify the DENV RdRp from both the his-SUMO tag and the SUMO protease. A S200 Superdex 26/60 Column (GE Healthcare) was used with an Akta Prime pump and collection system at 4 °C, with a 280nm absorbance sensor to determine protein concentration in eluted samples and for trace analysis. Prepared gel filtration buffer was de-gassed and filtered using a 0.2 µm sterile filter and the S200 column washed with de-gassed $ddH_20$ and equilibrated with de-gassed gel filtration buffer. Sample protein was sterile filtered using a 0.2 µm filter and loaded into an equilibrated 5 ml injection

loop. The pump and collection system was programmed to collect 3 ml fractions with a flow rate of 0.3 ml/min. The S200 column was routinely calibrated using a gel filtration buffer standard kit (Bio-rad). Fractions of interest were analysed by SDS-PAGE gel electrophoresis and pure protein concentrated, following the addition of 1 mM Tris (2-carboxyethyl)phosphine (TCEP), to the required concentration using a 50 kDa cut off Vivaspin concentration column (Sartorious Stedium) prior to experimentation.

### 2.3.6 Storage of purified DENV RdRp

Protein purified for X-ray crystallography studies was used for experimentation immediately. Preparations purified for activity studies and SHAPE mapping experimentation were stored at the desired concentration, following snap-freezing on dry ice, at – 80 °C.

## 2.4 Determination of DENV-2 RdRp activity

### 2.4.1 Circular Dichroism

Circular Dichroism was carried out to investigate the secondary structure of purified DENV RdRp. Purified DENV RdRp was diluted to a final concentration of 0.2 mg/ml in 300 mM NaCl, 20 mM Tris pH 7.5 and 1 mM TCEP in 200 µl. Circular dichroism was measured between 250 and 180 nm in a 1 mm path length cuvette using a Chirascan CD spectrometer (Applied Photophysics).

### 2.4.2 Polymerase activity assay

*De novo* RNA synthesis capability was determined in a total volume of 25 µl containing 50 mM HEPES (pH 8.0), 10 mM KCl, 5 mM $MgCl_2$, 10 mM Dithiothreitol (DTT), 40 U RNase inhibitor, 500 µM (each) GTP, CTP, UTP, 10 µM [α-$^{32}$P] ATP, 0.5 µg un-capped 5´160 nt DENV template RNA and 1 µg of purified DENV-RdRp. The reactions were incubated at 30 ºC for 30 min and RNA purified using an RNAeasy mini kit (Qiagen) following manufacturer's protocol. The RNA was eluted in RNase free water and α-$^{32}$P incorporation was measured by scintillation counting.

Positive and negative controls of purified Hepatitis C Virus (HCV) NS5B and HCV NS5B GND and corresponding RNA template were kindly provided by Dr. C. Bartlett (University of Leeds). HCV NS5B GND refers to a GDD to GND mutation in the active site of HCV NS5B rendering the polymerase unable to replicate.

### 2.4.3 Fluorescence Anisotropy

The RNA binding capability of DENV-2 RdRp was tested using Fluorescence Anisotropy (FA). Binding of RdRp to short transcripts (5-20 nucleotides) of 3´-fluorescein labelled RNAs representing the known RdRp binding site on SLA were synthesised (Dharmacon). Reactions were carried out in triplicate in 384- well format, in 20 mM Tris-HCl pH 7.5, 150 mM NaCl and 0.01% (v/v) Triton X-100, with 20 nM 3´-Fl-labelled RNA and increasing concentrations of DENV-2 RdRp (0.08 nM to 9 µM). Plates were incubated at room temperature for 30 min prior to data collection on a Spark 10M Multimode Microplate Reader (Tecan) with a 480nm excitation filter and S/P- channel emission filters at 530 nm. All graphs were plotted using OriginPro 9.1 (Origin Lab) according to:

$$y = \frac{A1-A2}{1+\left(\frac{x}{x^0}\right)p} + A2;$$

y is the fraction of RNA bound, $x$ is the protein concentration, *A*1 is the initial value of y, *A*2 is the final value of y, $x^0$0 is dissociation constant ($K_d$) and *p* is the Hill coefficient.

### 2.4.4 Electromobility Shift assay (EMSA)

The RNA binding and template specificity, at physiological temperatures, of purified recombinant DENV-2 RdRp was further confirmed using protein/RNA EMSAs.

#### 2.4.4.1 *In vitro* transcription of unlabelled RNA templates for competition EMSAs

All templates were approximately 160 nts in length. Un-labelled transcripts without SLA (Δ SLA) and the 3´ UTR of DENV genomic RNA were transcribed for analysis by EMSA competition.

#### 2.4.4.2 *In vitro* transcription synthesis of radioactive + SLA transcript

A radioactive template of approximately 160 nts in length containing the SLA promoter (+ SLA) was synthesised for shift analysis by EMSA. A DNA template was synthesised using PCR, utilising primers incorporating the T7 promoter sequence. The radioactively labelled +SLA RNA was *in vitro* transcribed in a 40 µl total volume containing 1 µg template, 1 X T7 polymerase buffer, 2 µl T7 polymerase, 1 µl ATP, 2 µl CTP, GTP and UTP,10 µCi [α-$^{32}$P]-ATP [Perkin Elmer], 1M $MgCl_2$, 2 µl Yeast inorganic pyrophosphatase (NEB), 1 µl (40 U) RNasin RNAse inhibitor (Promega) and nuclease free $H_20$ up to volume. The reaction was incubated for 3 hrs at 37°C before purification using RNAeasy mini kit (Qiagen) following manufacturer's protocol and eluted in 20 µl nuclease free water. Concentration was determined using a Nanodrop 1000 (Thermo Scientific) spectrophotometer. Purified transcripts were analysed for quality using denaturing MOPs agarose gel electrophoresis following fixing (46% (v/v) methanol, 46% (v/v) water and 8% (v/v) galacial acetic acid) and drying before exposing the gel to film and placed in an imaging cassette. RNA was visualised on a Fujifilm FLA-5000 phosphorimager using a Fujifilm image reader under default settings and a white laser set to 635 nm.

#### 2.4.4.3 Protein/RNA binding conditions

EMSA shift assays were carried out in a total volume of 30 µl containing 100 ng radioactive template RNA, 1 X SHAPE folding buffer, 0.5 X TE buffer, 1 µl (40 U) RNasin RNase inhibitor (Promega) and increasing concentrations of DENV-2 RdRp from 0 to 20-fold molar ratio protein to radiolabelled +SLA template RNA. Competition assays were carried out at 20-fold molar ratio protein to hot +SLA RNA template and the hot transcript competed out with either cold Δ SLA or 3´ UTR at 20, 40 or 80-fold hot + SLA concentration. All reactions were combined prior to the addition of protein and folded at either 37°C or 28°C for 30 min. The protein was added, and the reactions incubated for a further 60 min at the required physiological temperature (28°C or 37 °C). The reactions were combined to a concentration of 1 X RNA loading dye (non-denaturing, NEB) prior to analysis by native gel electrophoresis.

#### 2.4.4.4 Preparation of native gel

A 4.5% native acrylamide gel was prepared in a 40 ml total volume as follows; 6 ml Acrylamide mix (30%; 29:1 acrylamide: bisacrylamide), 4 ml 5 X Tris-borate /EDTA (TBE) buffer, 2 ml Glycerol (20% v/v), 28 ml nuclease free $H_20$, 300 µl APS and 30 µl TEMED.

#### 2.4.4.5 Electrophoresis parameters and visualisation

The native gel was pre-run in 1 X TBE buffer at 95 V for 30 min. The samples were loaded and the gel was run at 95 V for a further 3.5 hrs using a Biorad protean® II xi cell rig. Throughout the electrophoretic process the tank was packed around the outside with ice to prevent the gel from over-heating. EMSA assays were analysed following fixing and drying before exposing the gel to film in an imaging cassette and imaged as previously described (refer to section 2.4.4.2)

## 2.5 DENV-2 RdRp crystallographic techniques

Protein X-ray crystallography is a technique used to obtain the three-dimensional structure, and therefore interrogate the functionality, of a protein of interest via the X-ray diffraction of its crystallized form. We therefore sought to harness this technique to determine the structure of DENV-2 RdRp unbound and bound to DENV SLA RNA.

### 2.5.1 Sparse matrix screening

Sparse matrix screening of DENV -2 RdRp was carried out using the JCSG core suite screens I-IV (Qiagen). Sitting drops were set up in 96-well MRC-3 drop plates (Molecular Dimensions) using a Mosquito crystal Nanolitre protein crystallisation robot (TTP Labtech) with 30µl reservoir solution. Unbound protein was concentrated to 9 mg/ml and dispensed into 0.2-3 µl droplets containing protein: reservoir at ratios of 1:1, 2:1 and 1:2. Additionally, equimolar (200 µM) protein and SLA RNA was combined and dispensed in the same manner. Plates were sealed with Viewseal pressure adhesive clear seals (Grenier Bio-One) and incubated at 20 °C.

### 2.5.2 Crystal optimisation

JCSG core suite screen sitting drops were monitored using a light microscope for crystal growth 24 hours after they were set up and then periodically for 1 month following using a Rockimager 1000. The most successful condition from the screen was further optimised. Sitting drops were set up in 96-well MRC-3 drop plates using a Formulatrix NT8 crystallisation robot. Unbound protein was concentrated to both 9 mg/ml and 4.5 mg/ml and dispensed in 1 µl drops containing protein: reservoir at a ratio of 1:1. Equimolar protein and RNA was combined and dispensed in 0.75-1 µl droplets containing protein: reservoir ratios of 1:1 and 2:1. Plates were sealed with Viewseal pressure adhesive clear seals (as before) and incubated at 20 °C. Crystal growth was monitored periodically over a 4-week period.

### 2.5.3 Crystal streak seeding

To obtain large single diffracting quality crystals, streak seeding was employed. Streak seeding is a method to introduce pre-formed crystal nuclei into a hanging-drop to control nucleation in the aim of improving crystal growth. Poor quality crystals were aspirated from sitting-drop vapour diffusion plates and crushed into micro-seeds using either a cat whisker or Microseed beads (Molecular dimensions) to produce a concentrated seed stock. The seed sock was subsequently diluted 1:5 and 1:10 in Gel filtration buffer and streaked using a cat hair into sitting-drops comprised of 0.5 – 1 M tri-sodium citrate (Na3 Cit) and N-Cyclohexyl-2-aminoethanesulfonic acid (CHES) pH 9-10. Hanging drops were routinely monitored for crystal growth using a light microscope over a one-month period.

### 2.5.4 Crystal harvesting and data collection

Macromolecular crystals undergoing X-ray diffraction rapidly suffer radiation damage when collected at room temperature. This often results in the need for multiple crystals to collect a sufficient quantity of diffraction data for a complete dataset for the generation of the experimental electron density maps. The use of data collection at cryo-genic temperature has now been the principle method of data collection since the early 1990s. To facilitate the use of the cryo-genic temperature, the crystals must first be transferred to a condition typically containing the mother liquor with a suitable cryo-protectant prior to flash-freezing in liquid

nitrogen. The selection of cryo-protectants for any given protein crystal is often accomplished by trial and error. Cryo-protectant screening was performed using DENV-2 RdRp crystals and transferred into mother liquor containing different cryo-protectants consisting of 25% (v/v) glycerol, 25% (v/v) ethylene glycol, 25% (v/v) PEG 400 and 25% (v/v) methyl-2 4-pentanediol. This initial cryo-protectant screening resulted in disordered diffraction patterns of low resolution at approximately 5.0 Å. Further screening of suitable cryo-protectant using Paratone-N oil prior to cryo-cooling in liquid nitrogen yielded higher resolution and an ordered diffraction pattern. Paratone-N was subsequently used for cryo-cooling of all DENV-2 RdRp unbound and RNA bound crystals.

All data were collected at 100K°C at the Diamond Light Source on macromolecular crystallography beamline station I03 using a wavelength of 0.9795 Å. For unbound DENV-2 RdRp, a complete dataset was recorded to 2.2 Å resolution from a single rod- shaped crystal.

### 2.5.5 Data processing and structure solution

The diffraction images for DENV-2 RdRp were integrated, scaled and reduced using three different sets of programs implemented as part of the CCP4 program suite XDS, AIMLESS, and CTRUNCATE or Xia2 (Potterton *et al.*, 2003; Winn *et al.*, 2011). Five percent of the reflections were selected at random and excluded from the refinement using the program FREERFLAG and constituted the $R_{free}$ set (Potterton *et al.*, 2003; Winn *et al.*, 2011). The crystal structure of DENV-2 RdRp was determined by molecular replacement (MR) using the program *PHASER* with the DENV-2 RdRp domain bound to compound 27 (Lim *et al.*, 2016) (Protein Data Bank (PDB) accession number 5K5M) as the search model (McCoy *et al.*, 2007). One single MR solution was obtained from *PHASER* and after initial rounds of rigid body and restrained refinements using REFMAC5, iterative cycles of manual model building and refinement were carried out using COOT and REFMAC5 respectively (Emsley and Cowtan, 2004; Murshudov *et al.*, 2011). Figures were generated using PyMOL.

## 2.6 SHAPE mapping

*In vitro* and intracellular SHAPE mapping was utilised to interrogate the structure of DENV RNA, acting to further decipher the molecular constraints behind the linear to circular genomic switch.

### 2.6.1 *In vitro* SHAPE mapping

#### 2.6.1.1 Preparation of NMIA

N-Methylisatoic anhydride (NMIA) (Sigma-Aldrich) was resuspended in anhydrous DMSO to a concentration of 1 M and stored at – 80°C prior to experimentation.

#### 2.6.1.2 NMIA modification

Full-length or 5´500 nt (~ 60 µg) DENV genomic RNA in nuclease free $H_20$ RNA was denatured at 95°C for 2 min and snap-cooled on ice for 2 min. Following this 103 µl $ddH_20$, 45 µl 3.3 X folding buffer (333 mM HEPES, 30 mM $MgCl_2$, 330 mM NaCl), and 2 µl (80 U) RNase inhibitor (either RNase Out (Invitrogen) or RNasin RNase inhibitor (Promega) was added and incubated at either 37°C or 28°C for 30 min. If DENV-2 RdRp or *trans*-activating factors were to be added 10 µl of 2 mg/ml DENV-2 RdRp or 10 µl cellular extract was added and incubated for a further 30 min at the required temperature. Following incubation, the reactions were equally divided into positive and negative reactions, to which 8 µl 50 mM NMIA (positive) or DMSO (negative) was added respectively. The reactions were incubated for 50 min at the required temperature (28°C or 37 °C) and modified RNA purified using phenol-chloroform (Life Technologies) following manufacturer's protocol. RNA pellets were resuspended in 10 µl 0.5 X TE buffer.

#### 2.6.1.3 Reverse transcription and preparation of sequencing ladder

For both the positive and negative reactions 5 µl of modified RNA was incubated with 1 µl 10 µM 5´ FAM labelled fluorescent primer, and 6 µl $ddH_20$ at 85°C for 1 minute, 60°C for 10 min and 30°C for 10 min. A master mix of 4 µl Superscript III RT buffer, 1 µl 100 mM DTT, 0.5 µl 100 mM dNTPs, 1 µl (40 U) RNasin RNase inhibitor (Promega), 1 µl $ddH_20$ and 1 µl (200 U) Superscript III reverse transcriptase (Thermo Scientific) was prepared and added to each reaction. The reactions were incubated for 30 min at 55 °C.

For preparation of the sequencing ladder, ~15 µg of unfolded *in vitro* transcribed template RNA was combined with 1 µl 10 mM 5´HEX labelled fluorescent primer and 2 µl $ddH_20$. The reaction was incubated at 85°C for 1 minute, 60°C for 10 min and 30°C for 10 min. A master mix of 4 µl Superscript III RT buffer, 1 µl 100 mM DTT, 0.5 µl 100 mM dNTPs, 2 µl ddCTP, 1 µl (40 U) RNasin RNase inhibitor (Promega), 1 µl $ddH_20$ and 1 µl Superscript III reverse transcriptase (Thermo Scientific) was prepared and added before incubation at 55°C for 30 min.

1 µl 4 M NaOH was added to all reverse transcription reactions and heated at 95°C for 3 min. The reactions were then cooled to 4°C and pH adjusted with 2 µl 2 M HCl. cDNA was precipitated in 4 µl 3 M NaAc, 4 µl 100 mM EDTA, 1 µl 20 mg/ml Glycogen (Thermo Scientific) and 60 µl 100% ethanol for 16 hrs at – 80 °C. cDNA was pelleted by centrifugation and supernatant removed. The resulting cDNA pellet was resuspended in 40 µl deionised formamide (Hi-Di formamide (Thermo Scientific). Positive and negative samples were combined with 10 µl of corresponding sequencing ladder and stored at -80 °C.

### 2.6.2 Intracellular SHAPE

#### 2.6.2.1 NAI synthesis

2-methylnicotinic acid imidazolide (NAI) was produced as a 1:1 mixture with imidazole in DMSO stock solution as described in (Spitale *et al.*, 2013). Briefly, 137 mg (1 mmol) of 2-methylnicotinic acid was dissolved in 0.5 ml anhydrous DMSO. A solution of 162 mg (1 mmol) 1,1'-carbonyldiimidazole in 0.5 ml anhydrous DMSO was added dropwise over 5 min. The resulting 1 M solution was stirred at room temperature for 1 hr to allow for complete de-gassing and then further stirred at room temperature for a further 1 hr. The 1 M stock solution was aliquoted and stored at - 80°C

#### 2.6.2.2 NAI treatment and RNA extraction of stably expressing HEK293T DENV-2 replicon cells

6 X T175 flasks of stably expressing DENV-2 HEK293T replicon cells were maintained and allowed to reach 80% confluency prior to NAI treatment. For RNA structure probing, cell monolayers were washed twice with ice cold PBS and scraped into PBS. The cells were pelleted by centrifugation at 1000 *x g* and half of

the pellets resuspended in 2 ml 200 mM NAI (in 10% DMSO made up in PBS) (positive) or 10% DMSO in PBS (negative) and allowed to incubate at 37°C for 15 min. Following incubation, the cells were pelleted by centrifugation at 1000 *x g* and supernatant discarded, and the cell pellets resuspended in Trizol (Thermo Scientific). The reactions were then freeze-thawed at – 20°C prior to RNA extraction, as per manufacturer's instructions. Extracted RNA was resuspended in 20 µl nuclease free $H_20$. Quality of extracted RNA was analysed by denaturing MOPs agarose gel electrophoresis prior to primer extension. Stable expression of DENV-2 replication proteins was further confirmed by RT-PCR and via cell lysis with PLB and western blot analysis.

#### 2.6.2.3 NAI treatment and RNA extraction of DENV-2 infected mammalian/ insect cells

BHK-21 or C6/36 cells were routinely maintained as previously described and allowed to reach approximately 70% confluency prior to DENV-2 infection. 6 X T175 flasks of BHK-21 or C6/36 cells were infected at an MOI of 0.0002 and incubated at either 37°C or 28°C as appropriate for 72 hrs prior to NAI treatment and RNA extraction as described in section 2.6.2.2. DENV-2 infection was confirmed following RNA extraction by RT-PCR and infective culture supernatant analysed via lysis with PLB and western blot analysis

#### 2.6.2.4 Enrichment of DENV-2 replication complexes from stably expressing DENV-2 replicon cells

6 X T175 flasks were maintained of stably expressing DENV-2 replicon cells and grown to 80% confluency prior to NAI treatment. Prior to 16 K fractionation, cell monolayers were washed twice with ice cold PBS and scraped into PBS. The cells were divided in half and pelleted by centrifugation at 1000 *x g* and half resuspended in 6 ml 200 mM NAI (in 10% DMSO made up in PBS) (positive) or 10% DMSO (negative) and allowed to incubate at 37°C for 15 min. Following incubation, the cells were pelleted by centrifugation, supernatant discarded.

Active DENV-2 replication complexes were enriched based on the protocol of (Uchil, Kumar and Satchidanandam, 2006). Each cell pellet was resuspended in 5 ml Tris-HCl/NaCl/Mg (TNMg) buffer and allowed to swell on ice for 20 min. The cells were disrupted by passing through 19- and 25-gauge needles 20 times each

and 10 times through a 27-gauge needle. The cell lysate was centrifuged at 800 x *g* for 12 min at 4 °C, generating supernatant and nuclear pellet (Nuc) fractions. The supernatant fraction was transferred to a new tube and further centrifuged at 16,000 x *g* for 15 min, resulting in a cytoplasmic fraction (Cyt) and a heavy membrane pellet fraction containing the viral replication complex (16K). The Nuc and 16K pellets were resuspended in 125 µl of TNMg buffer. The two positive and negative 16 K fractions were divided into 3 prior to the addition of 200 µl Trizol (Thermo Scientific). Total RNA was extracted following freeze-thaw at – 20°C as per manufacturer's instructions and resuspended in 20 µl nuclease free $H_20$. RNA quality was analysed by denaturing MOPS agarose gel electrophoresis.

All extracts were aliquoted and stored at -80 °C. Fractions were further analysed for enrichment by SDS-PAGE gel electrophoresis and western blot.

#### 2.6.2.5 Reverse transcription and preparation of sequencing ladder

20 µl NAI/DMSO treated, Trizol extracted RNA was combined with 4µl 10 µM FAM labelled fluorescent primer and incubated at 85°C for 1 minute, 60°C for 10 min and 30°C for 10 min. A master mix of 8 µl (200 U/ µl) Superscript III RT buffer, 2 µl 100 mM DTT, 2 µl 10 mM dNTPs, 1 µl (40 U) RNasin RNase inhibitor (Promega), 1 µl $ddH_20$ and 2 µl Superscript III reverse transcriptase (Thermo Scientific) was prepared and added to each reaction. The reactions were incubated for 30 min at 55 °C.

Sequencing ladders were prepared, using the appropriate *in vitro transcribed* RNA template, and combined with probed samples as described in 2.6.1.3.

### 2.6.3 SHAPE data analysis

*In vitro* and intracellular SHAPE fragment analysis was conducted by capillary electrophoresis by DNA sequencing and services, Dundee University. SHAPE data was analysed in the program QuSHAPE (Karabiber *et al.*, 2013) and RNA structure prediction was carried out using RNAstructure software (Reuter and Mathews, 2010) using the generated SHAPE reactivity profile as a pseudo free energy constraint. Alternatively, reactivities were overlaid onto the thermodynamically predicted structure (Mfold, (Zuker, 2003)) of the first 145 nts of

DENV genomic RNA using the RNA structure applet VARNA. All figures were created using the RNA structure applet VARNA.

# Chapter 3 Expression, purification, functional and structural analysis of DENV-2 RdRp

## 3.1 Introduction

DENV is the most prominent arboviral disease in the world (European Centre for Disease Prevention and Control, 2018). To date, neither direct acting anti-virals or a universally accessible vaccine exists in the treatment of DENV. DENV replication provides one such significant anti-viral target and is therefore the focus of our studies.

A fundamental requirement of DENV replication is the binding of the DENV-2 RdRp domain to 5´SLA RNA. Genomic cyclisation acts to transfer the RdRp to the 3´SL to initiate negative-sense RNA synthesis and therefore replicate the viral genome. Consequently, this protein-RNA interaction provides a significantly understudied structure based direct acting anti-viral target.

This chapter describes the expression, purification, activity analysis and structural determination using X-ray crystallography, of the DENV-2 RdRp domain and our RdRp and SLA co-crystallisation efforts. Additionally, purified, active DENV-2 RdRp was used during *in vitro* SHAPE mapping procedures (Chapter 4).

### 3.1.1 Introduction to the bacterial expression of DENV-2 RdRp

Structural studies such as X-ray crystallography require the production of milligram quantities of homogeneous pure protein. The most common method of protein expression is the over-expression of recombinant proteins utilising prokaryotic *E. coli* based expression systems and subsequent purification.

The plasmid used here is a pET-28a based vector whereby the expression of the DENV-2 RdRp target protein was under the control of a lac operator. Utilising the pET system, expression of the protein of interest is driven by bacteriophage T7 RNA polymerase which binds to the specific promoter region upstream of the target DENV-2 RdRp ORF for expression. pET systems can therefore be introduced into various strains of *E. coli* which express T7 RNA polymerase, such as the strain utilised here BL21-CodonPlus (DE3)-RIPL. Expression of DENV-2 RdRp was induced via the introduction of IPTG which binds to the lac repressor instigating an allosteric change in the repressor, catalysing its removal from the lac operator. This allows for the production of T7 RNA polymerase, the synthesis of mRNA and

subsequent protein expression from lac operator linked genes. A schematic of the IPTG induced expression of the DENV-2 RdRp construct is depicted in figure 29.

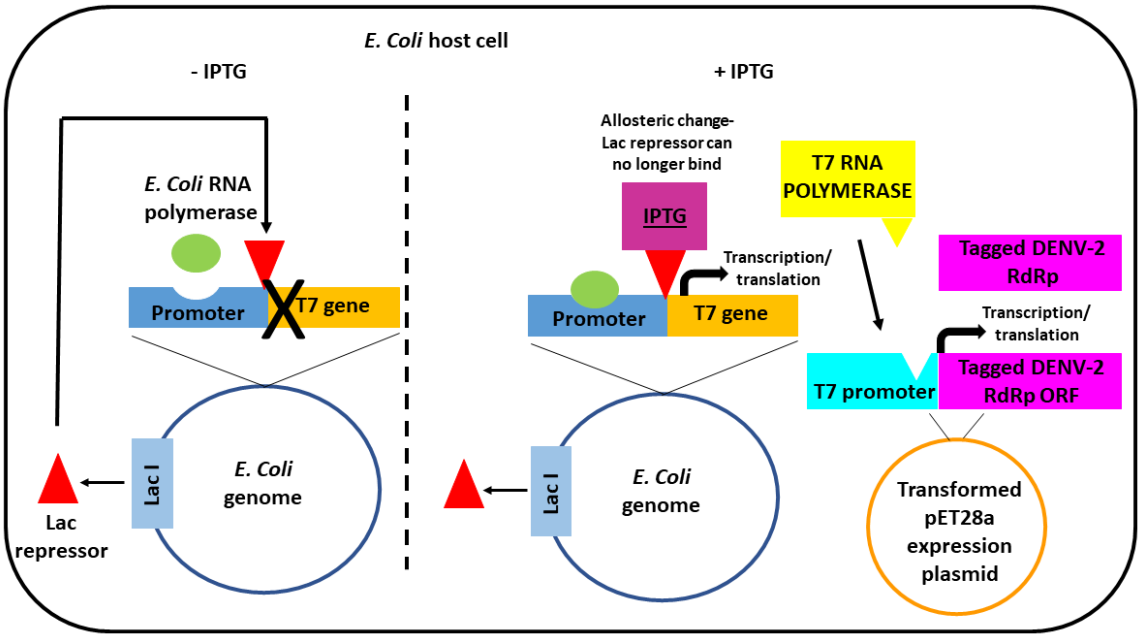



**Figure 29 Expression of DENV-2 RdRp utilising the pET expression system:**

*The expression of his-SUMO-DENV-2 RdRp construct following IPTG induction in BL21-Codon Plus (DE3)-RIPL E.coli prokaryotic cells. In the absence of IPTG the lac repressor prevents the transcription of T7 RNA polymerase (left). Following IPTG induction, IPTG binds to the lac repressor catalysing an allosteric change rendering the lac repressor unable to bind, allowing for the production of T7 RNA polymerase. The T7 RNA polymerase then binds to the T7 promoter allowing for the expression of the target protein (right).*

### 3.1.2 Introduction to the purification of DENV-2 RdRp

The establishment of a successful purification protocol demanded stringent testing of multiple growth conditions, considering both temperature, concentration of IPTG induction and chromatography techniques, to yield pure, functional DENV-2 RdRp of a sufficient yield for both X-ray crystallography and *in vitro* SHAPE studies (discussed in Chapter 4).

Purification optimisation was carried out using full-length DENV-NS5 however the full-length protein was inherently prone to degradation. When the stability of this protein could not be improved, DENV-2 RdRp was shown to purify suitably utilising the same protocols without further optimisation (Fig. 30). This is therefore the

reason why the optimisation of IPTG temperature and concentration is described using the full-length DENV-NS5 protein.

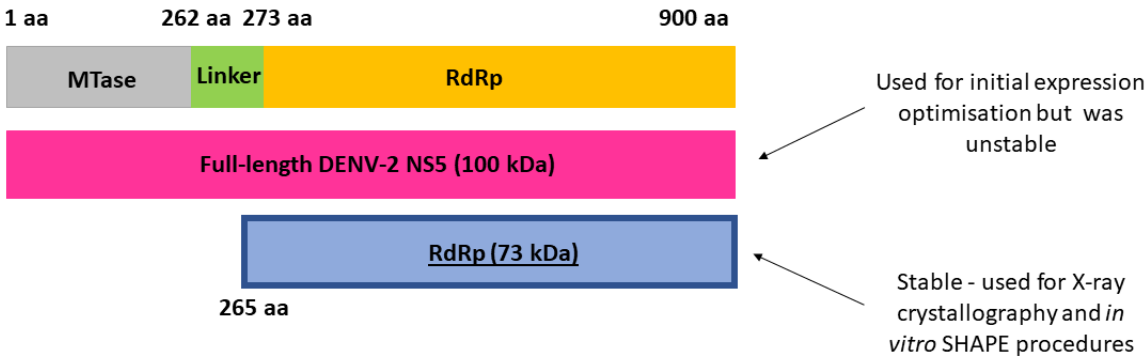



**Figure 30 Schematic of DENV-2 full-length NS5 and DENV-2 RdRp constructs:**

*Full-length DENV-2 NS5 was used for the optimisation of protein expression and purification. Following purification, the full-length protein was unstable. The DENV-2 RdRp construct remained stable following the same purification procedure and was therefore used for both X-ray crystallography and in vitro SHAPE mapping experimentation.*

The final optimised DENV-2 RdRp purification procedure is based on two separate purification techniques – nickel affinity chromatography followed by size exclusion chromatography (SEC):

Nickel affinity chromatography exploits the binding of poly-histidine tagged proteins to resin embedded with immobilised $Ni^{2+}$. Histidine, through its side chain, binds to these immobilised ions because of electron donor groups within the imidazole ring. The use of poly-histidine tagged DENV-2 RdRp allows for the capture of the target protein on immobilised nickel resin and therefore isolation from contaminating proteins of *E. coli* origin. Following the binding of his-SUMO-DENV-2 RdRp to nickel resin, sequential washing with buffers containing an increasing concentration of imidazole competes with the bound target protein. Imidazole competes with the histidine-$Ni^{2+}$ binding site gradually removing weaker binding partners, facilitating the isolation of the target protein (Fig. 31).

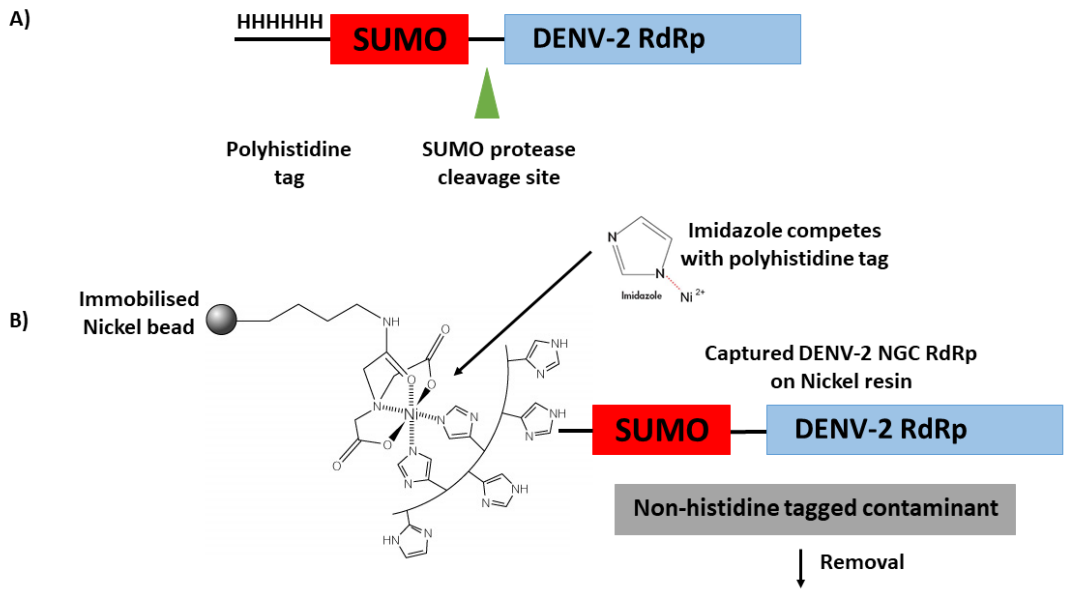



**Figure 31 The *E. coli* expressed DENV-2 RdRp construct and nickel affinity chromatography principle:**

***A)*** *The E. coli expressed poly-histidine and SUMO tagged DENV-2 RdRp.* ***B)*** *Nickel affinity purification of his-SUMO tagged DENV-2 RdRp on immobilised nickel resin. Nickel affinity diagram adapted from* (Promega, 2018)*, Imidazole ring adapted from* (IGEM, 2018)*.*

Following nickel affinity chromatography, the poly-histidine tag was cleaved using SUMO protease during dialysis which also removed the imidazole from the protein preparation. The cleavage of the SUMO tag, comprising the 6xhis-SUMO portion of the fusion protein from the DENV-2 RdRp target protein generated two separate proteins – one ~73 kDa protein comprising the DENV-2 RdRp and the his-SUMO tag of ~15 kDa, in addition to the SUMO protease of ~25 kDa. The SUMO tag is cleaved through a tertiary structure driven interaction, allowing for the increased specificity of the protease (Panavas, Sanders and Butt, 2009). Size exclusion chromatography (SEC) was then utilised to separate the three distinct proteins away from each other on account of their relative size.

SEC utilises a bead matrix with varying microscopic pores present within the beads. The physical structure of the matrix allows for proteins of different size to flow through the resin at different rates (Wang *et al.*, 2010). Depending on the chosen column type, multimeric and large proteins are eluted first as they flow

through the column without restriction. In contrast, smaller proteins flow through the pores in the matrix, reducing their flow rate and therefore their elution time. Protein elution is monitored by UV absorption at 280 nm and size confirmed in comparison to a standard protein calibration curve (Fig. 32).

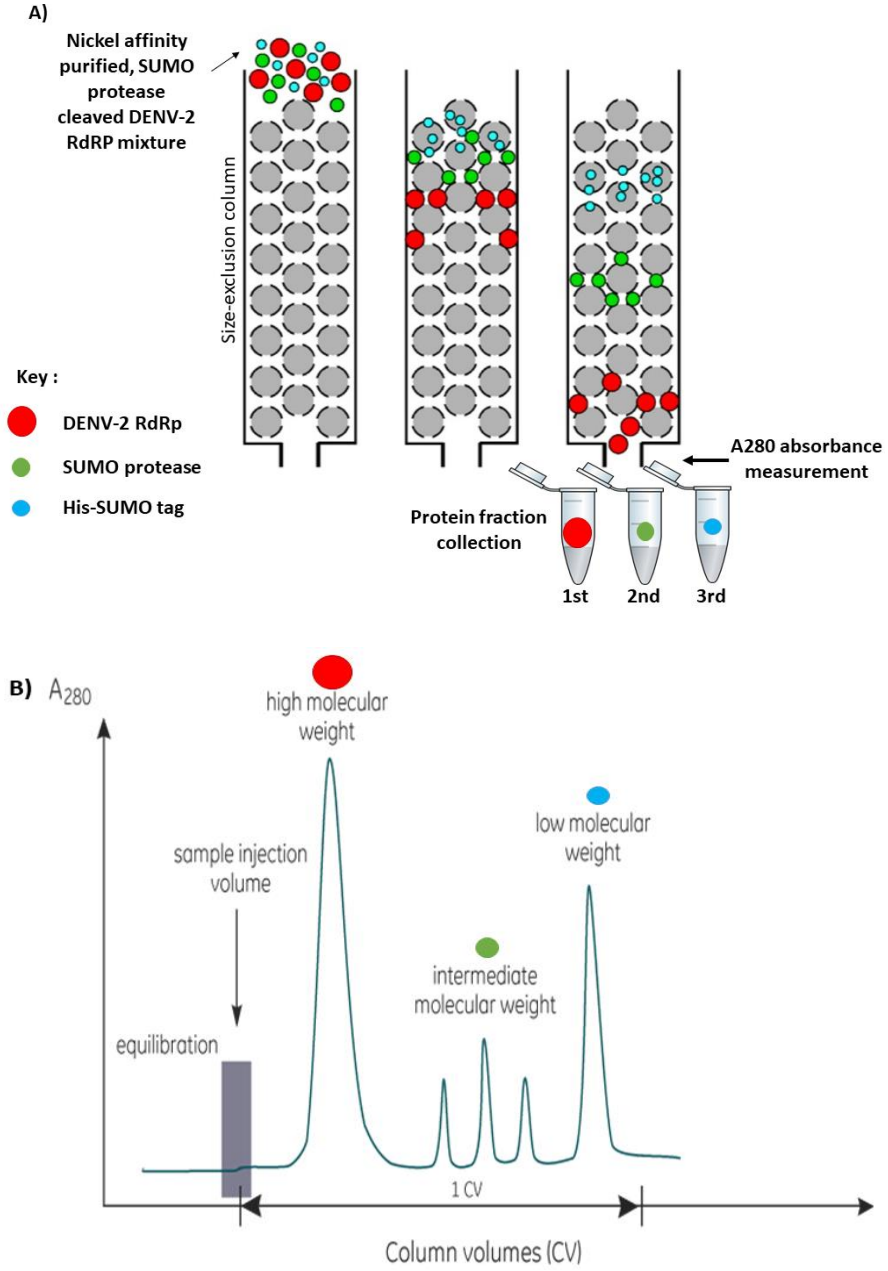



**Figure 32 Size-exclusion chromatography:**

***A)*** *The principle of size exclusion chromatography utilised in the purification of DENV-2 RdRp.* ***B)*** *A280nm absorbance chromatogram for protein fraction selection. Adapted from (Creative Biostructure, 2018; GE Healthcare Life Sciences, 2018).*

### 3.1.3 Confirming the identity and activity of DENV-2 RdRp

To determine the structural influence of DENV-2 RdRp on DENV genomic RNA structure and for structural analysis by X-ray crystallography correctly folded, active protein, with the ability to bind DENV RNA, and of confirmed identity was required. Consequently, we identified and analysed the folding capability and activity of the protein using multiple methods as described below:

The identity of DENV-2 RdRp was confirmed following purification by western blot. Western blot is an immunological technique whereby the protein of interest is identified following SDS-PAGE gel electrophoresis with subsequent transfer onto a membrane. The membrane is then probed with an antibody directed against the target protein, in this case the C' terminal portion of DENV-2 NS5. The bound primary antibody is then further detected using a secondary fluorescently tagged antibody directed against the species of the primary antibody. The secondary antibody and therefore the identity of the target protein can then be confirmed in comparison to a standard protein size ladder.

Circular dichroism (CD) was carried out to investigate the secondary structure of purified DENV-2 RdRp. The technique measures the difference in absorption of left and right circularly polarised light. This phenomenon arises due the chirality of proteins as determined by the presence of α- helices, β-sheets or random coils, each with unique absorbance spectre. The differential absorbance for different protein structures allows for both the type and extent of secondary structure formation of a protein to be determined. This was used to confirm the correctly folded structure of DENV-2 RdRp in comparison to other published serotype structures.

The *de novo* RNA synthesis capability of DENV-2 RdRp was confirmed using an optimised polymerase activity assay. Following filter purification, the incorporation of α-$^{32}$P ATP into the nascent RNA chain was quantified using scintillation counting.

The DENV genomic RNA binding capability of purified DENV-2 RdRp was confirmed using both Fluorescence Anisotropy (FA) and Electromobility Shift Assay (EMSAs) at both mammalian and insect temperatures. The principle of FA

is that when a fluorescently labelled molecule, in this case DENV RNA, is excited by polarised light, the fluorophore emits light with a degree of polarisation that is inversely proportional to the 'tumbling rate' of the labelled molecule. A small unbound labelled RNA will have a fast tumbling rate which will result in depolarisation of the emitted light during the lifetime of its excitation (~4 nanoseconds for fluorescein, used here). However, if the small labelled RNA binds a larger target protein molecule, i.e. DENV-2 RdRp, the RNA-protein complex will have a slower rotation and the emitted light will be polarised to an extent (Fig. 33). Binding of DENV-2 RdRp to different lengths of 3´ fluorescein-labelled RNAs was assessed in this manner, with higher anisotropy values indicative of binding. Additionally, FA was further utilised to select suitable DENV genomic binding partners with suitable $EC_{50}$s for DENV-2 RdRp – RNA co-crystallisation studies.

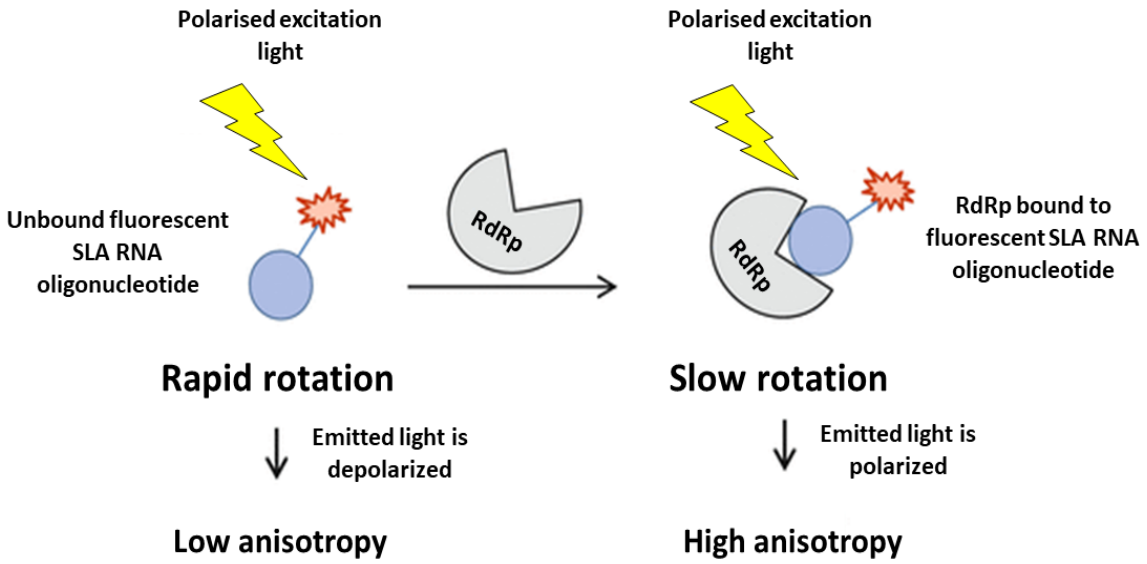



**Figure 33 Principle of fluorescence anisotropy:**

*A schematic of the principle of fluorescence anisotropy.* ***Left:*** *Small 3´-fluorescently tagged RNAs, for example small SLA RNA oligonucleotides, have rapid rotation in solution when unbound and therefore tumble rapidly. Fluorescence is subsequently depolarized by the rapid motion resulting in reduced anisotropy.* ***Right:*** *Binding of a larger protein, for example DENV RdRp, to the fluorescently labelled RNA reduces the rate of tumbling. The emitted fluorescence of the bound protein-RNA complex is subsequently depolarised less resulting in a greater anisotropy. Adapted from* (Du, 2015).

EMSAs were used to further confirm the binding capability of DENV-2 RdRp to DENV genomic RNA at both mammalian and insect physiological temperature (37 °C and 28 °C). A mobility shift assay is the electrophoretic separation of a protein-DNA or, in this case, protein-RNA mixture on a native polyacrylamide gel. The target RNA is radioactively labelled to allow for detection. Radioactively labelled RNA of the same concentration is subsequently incubated with an increasing concentration of DENV-2 RdRp, at the desired physiological temperature and analysed following native polyacrylamide gel electrophoresis and exposure to photographic film. If the protein is able to bind to the target labelled RNA, a larger complex is formed and as a result the complex 'shifts' up on the gel due to the slower migration of the bound protein-RNA complex. The shift is proportional to the concentration of bound protein, i.e. the greater the concentration of protein the greater the shift. Furthermore, the affinity and specificity of the protein for its target RNA can be interrogated using competition EMSAs, whereby the labelled target RNA is out-competed using different RNA transcripts. We exploited this technique to further analyse the requirement of DENV-2 RdRp binding to 5´SLA. Therefore, we utilised EMSAs and competition EMSAs to further confirm the RNA binding capability of DENV-2 RdRp and to interrogate the affinity of the protein for 5´SLA RNA.

### 3.1.4 Determination of the crystal structure of DENV-2 RdRp and RdRp-RNA co-crystallisation studies

The structural study of DENV full-length N55 and the RdRp domain has largely focussed on DENV serotype 3 (Yap *et al.*, 2007; Lim *et al.*, 2013; Noble *et al.*, 2013). Although a crystal structure has previously been solved of the DENV-2 RdRp bound to compound 27 (Lim *et al.*, 2016) the unbound DENV-2 RdRp has yet to be published. Moreover, neither full-length NS5 or RdRp crystal structures of any serotype have been solved bound to DENV-SLA RNA. Consequently, we sought to study this critical protein-RNA interaction to gain knowledge at the structural level of the change in the structure of the protein upon the initiation of DENV genome replication – a significant anti-viral target.

Approximately 80% of the macromolecular structure data in the Protein Databank (PDB) was obtained using the technique of X-ray crystallography, highlighting the

importance of this technique in structural research (RCSB PDB, 2018). During an X-ray crystallography experiment, X-rays are diffracted by the electrons contained within a protein crystal. Collection of the diffracted data allows the researcher to calculate an electron density map into which the atomic model of the protein can be fitted. However, structure solution poses a multitude of problems. Some of these difficulties are fundamental, for example the 'phase problem' and other difficulties are practical, such as crystal growth and harvesting. However, despite limitations, X-ray crystallography remains the preferred option for high-resolution structure solution with the ability to achieve atomic resolutions of typically 1-3Å (RCSB PDB, 2018).

#### 3.1.4.1 An introduction to crystal growth

Once a suitable protocol has been established for pure protein production the next step is crystal growth. X-ray crystallography requires well-ordered, and well diffracting protein crystals. Additionally, experimentation often requires highly automated robotic systems for setting up crystal screens and making buffers allowing for crystallisation conditions to be reproduced more accurately. Yet despite robotic advances and standardisation, crystal growth remains a fundamental hurdle of the crystallography process.

The growth of protein crystals is required because the intensity of diffracted X-rays from a single protein molecule is weak and undetectable. Protein crystals contain many copies of a protein in an ordered lattice that is utilised to amplify the signal of diffracted X-rays enabling their detection.

Crystallisation requires bringing the protein of interest to a supersaturated state whereby it can either crystallise or precipitate. Introducing the sample to a precipitating reagent, such as a salt or polymer (e.g. polyethylene glycol), can promote the nucleation of protein crystals in solution. Vapour diffusion is a widely used method for protein crystallisation whereby a drop containing a mixture of precipitant (mother liquor) and protein solution is sealed in a chamber containing pure precipitant. Water vapour then diffuses out of the drop until the osmolarity of the drop and precipitant reaches equilibrium. The resultant dehydration of the drop causes a slow increase in the concentration of both the protein and precipitant solution. If the appropriate precipitant condition (mother liquor) has been used for

a given protein then crystals will grow in the drop. This can be performed in either a 'sitting drop' or 'hanging drop' manner, depicted in figure 34. A schematic representation of crystal nuclei formation and growth is shown by the two-dimensional solubility diagram in figure 34. The time required for crystal formation varies for different proteins and this can take anywhere from hours to months.

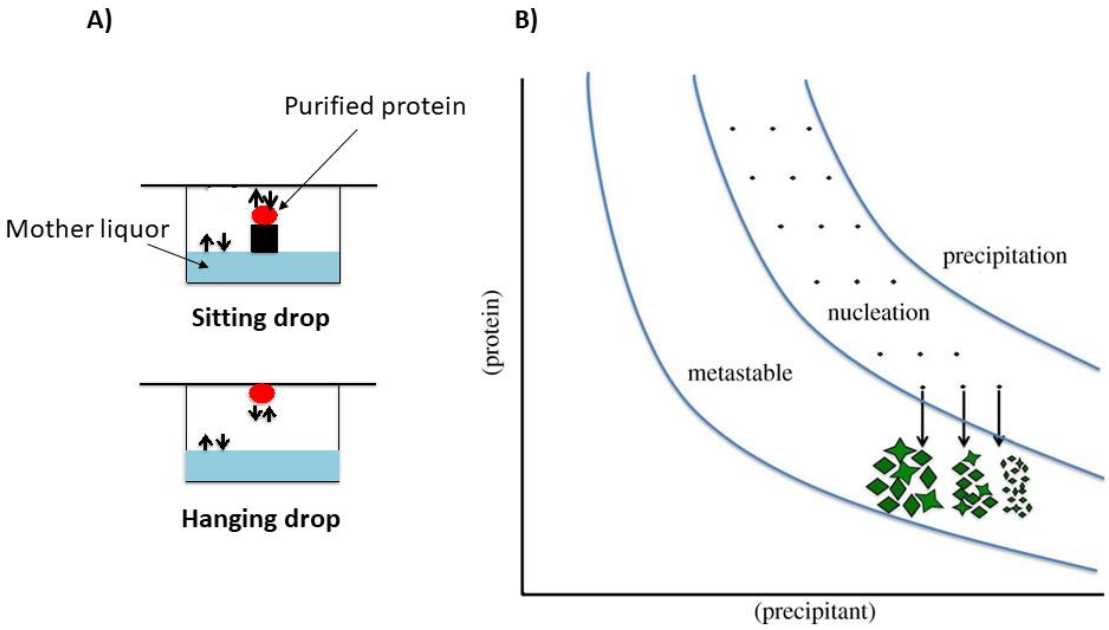



**Figure 34 Techniques and principles behind crystal growth:**

***A)*** *Sitting drop and hanging drop techniques used for crystal growth**.** Black arrows represent vapour diffusion.* ***B)*** *Crystal growth phase diagram. Taken from* (Kupitz *et al.*, 2014).

#### 3.1.4.1.1 Streak seeding for crystal growth

Streak seeding is a technique where small nuclei or micro-crystals are transferred to a mother liquor solution that is optimized for crystal growth, to induce the formation of larger singular protein crystals. Nucleation requires a higher concentration of precipitant than crystal growth therefore providing nuclei can aid crystal growth. Streak seeding is performed following the mechanical breakage of small protein crystals (often using a cat whisker or horse hair), and then diluting the micro-crystals through a mother liquor droplet containing fresh purified protein. This action increases the likelihood that crystals will grow in the metastable zone

(Bergfors, 2003) (Fig. 34). This technique was used throughout our experimentation to improve crystal yield and growth.

#### 3.1.4.1.2 Co-crystallisation and crystal soaking

Co-crystallisation is a technique whereby a specific binding partner is introduced during or following purification and the protein crystallises bound to its partner molecule. Co-crystallisation is used widely in the drug discovery process to aid in the understanding of how drug candidates bind to and affect their protein target. Much effort has been made to fully automate this process in order to enhance the drug discovery process (Stewart, Clark and Behnke, 2002). Crystal soaking is another method of protein structure determination in the presence of a bound ligand. Grown protein crystals are soaked at the harvesting stage in mother liquor containing an excess of ligand. Crystal soaking for 10 - 30 minutes is often sufficient to populate the crystal if the binding site is accessible in the crystal lattice (Schlichting, 2005). Both methods were utilised in our DENV-2 RdRp and SLA RNA co-crystallisation studies.

A co-crystal is defined as a crystal that is built up of two or more organic compounds (Technobis Crystallisation systems, 2018). Co-crystallisation X-ray crystallography studies are commonly used in the process of drug design, identifying the binding site and interactions of the compound of interest with the target protein. Due to the fact that protein-RNA interactions play an important role in a variety of fundamental biological activities protein-RNA co-crystallisation is an excellent way to study such interactions. For example, to gain insight into the mechanism of HCV RNA replication and its inhibition by nucleotide analogue inhibitors, atomic resolution ternary co-crystal structures of NS5B in both primed initiation and elongation states were determined, defining the structural requirements for HCV genomic replication (Appleby *et al.*, 2015). However, co-crystallisation studies are often difficult and require significant optimisation of the crystallisation conditions. Proteins that crystallised and diffracted in the absence of a ligand may cease to do so in the presence of a ligand for any number of reasons, for example as a result of crystal packing rendering the binding site inaccessible. Nevertheless, protein-RNA co-crystallography remains a powerful structural tool.

#### 3.1.4.2 Crystal harvesting

Prior to data collection, protein crystals are often harvested from their mother liquor solution into a cryoprotectant solution using a suitably sized cryo-loop. A suitable cryoprotectant solution for the crystal requires optimisation in order to minimise any damage to the crystal in order to obtain the best crystal diffraction. Cryoprotectant solutions prevent the formation of ice crystals within the crystal lattice that culminate in poor X-ray diffraction. Cryoprotectants often consist of the mother liquor solution with an additive such as glycerol, polyethylene glycol or sucrose. Additionally, crystals can be harvested using oil as a cryoprotectant, such as Paratone N, displacing the surrounding water. In the presence of Paratone N, the protein crystal and its thin layer of surrounding mother liquor will form an amorphous glass in which the crystal lattice suffers minimal damage. Crystals are frozen in cryoprotectant solution as data collection occurs in a gaseous nitrogen stream at 100 K and for long term storage.

#### 3.1.4.3 Crystal diffraction, data collection and structural solution

A harvested, flash-frozen protein crystal is placed in the path of a beam of monochromatic (single wavelength) X-rays to collect X-ray diffraction data. Crystallographic data necessitates the use of X-rays because the wavelength of X-rays is in the same order of magnitude as the interatomic distances that are resolved during an X-ray crystallography experiment (1 Å = 0.1 nm). Upon placing a crystal in the path of an X-ray beam, most of the radiation passes through. Occasionally a photon from the X-ray beam interacts with the electrons around the atoms and is scattered from the crystal. The resultant scattered X-ray waves reflect from the crystal with different phases and scattering angles. Constructive interference, where the waves are in phase, may occur which results in waves of higher amplitudes or destructive interference, where the waves are out of phase, which results in waves of lower amplitudes, or in certain cases no signal (Fig. 35).

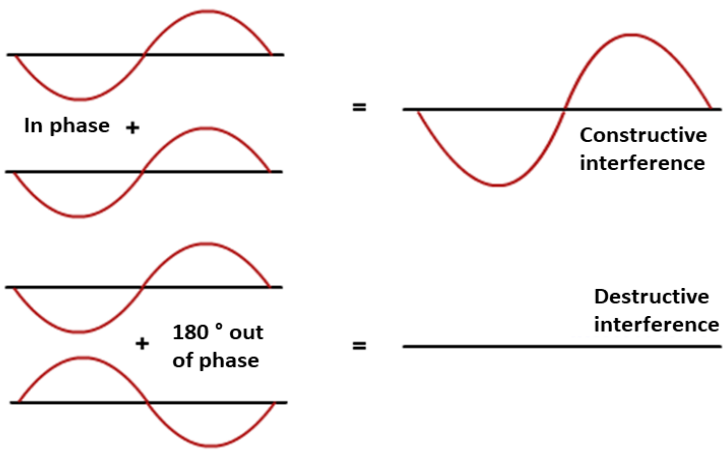



**Figure 35 Constructive and destructive interference:**

*A schematic of constructive and destructive interference. Adapted from* (TutorCircle, 2018).

When these waves are scattered from the protein crystal and hit a detector a pattern and position of intensities is recorded termed reflections. The diffraction pattern is a product of these reflections and constitutes the raw data of an X-ray diffraction experiment.

Bragg's law was invented by Sir William Henry Bragg and his son at the University of Leeds winning a Nobel prize in 1915 - founding crystal structure analysis using X-ray diffraction. Bragg's law explains the relationship between an incident X-ray onto and its scattering from the crystal surface. The law states that when the X-ray is incident onto a crystal surface, its angle of incidence, $\theta$, reflects back at the same angle of scattering, $\theta$. When the path difference, $d$, is equal to a whole number, $n$, of wavelength a constructive interference will occur. Bragg's Law is defined as $n\lambda = 2dsin\theta$ and is schematically represented in figure 36. If Bragg's law is satisfied, constructive interference will occur and the crystal will appear to have reflected the X-ray radiation. If the law is not satisfied, destructive interference will occur and data cannot be collected.

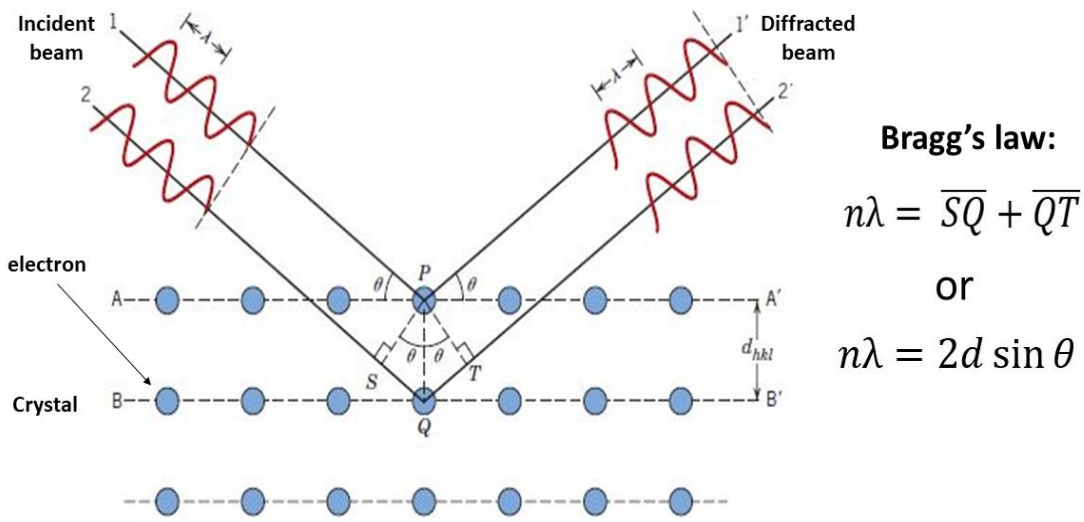



**Figure 36 Bragg's law:**

*A crystal is made up of parallel planes of atoms with interplanar spacing d. When a parallel beam of X-rays of wavelength λ is directed on these planes at an angle of θ constructive or destructive interference will occur. Two such X-rays, labelled as 1 and 2, are scattered by the atoms P and Q. The scattered waves 1' and 2' will constructively interfere if they are in phase. Bragg's law is met and a reflection is generated contributing to the diffraction pattern. Adapted from* (Callister and Rethwisch, 2011).

The resultant constructive wave that generates a reflection, which can be visualised on the detector, has contributions from each diffracting electron in every unit cell throughout the crystal (Fig. 37).

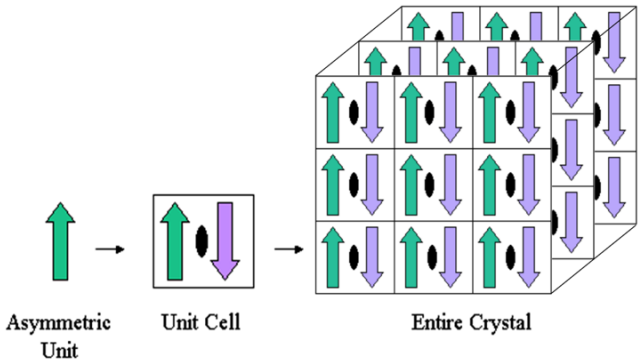



**Figure 37 A schematic representation of the crystal lattice and relationship with the unit cell and asymmetric unit:**

*The asymmetric unit is defined as the smallest section of a crystal structure in which symmetry operations can be utilised to generate an entire unit cell. The unit cell is the crystal repeating unit and, when translated in three dimensions, makes up the entire crystal. Taken from* (PDB, 2018).

Bragg's planes (defined as a plane in reciprocal space which bisects a reciprocal lattice vector) intersect each unit cell in the same manner therefore each unit cell contributes equivalent constructive interference to the resultant wave and thus amplifies the signal that is then detected. The way Bragg's planes intersect each unit cell are termed Miller indices (*hkl*) and these are assigned to each plane that contributes to the diffraction pattern. Each reflection in the diffraction pattern can be assigned corresponding Miller indices during indexing of the diffraction data. Corresponding waves across the whole crystal that culminate in a single reflection can be added together in the structure factor equation which results in the structure factor $F_{hkl}$ being assigned per reflection. The structure factor amplitude $|F_{hkl}|$ is therefore proportional to the square root of the reflection intensity (Fig. 38).

$$F(hkl) = \sum_{n=1}^{N} f_n e^{2\pi i(hu+kv+lw)}$$

**Figure 38 The structure factor equation:**

*hkl = indices of the considered diffraction plane, uvw = co-ordinates of the atoms in the lattice, N = number of atoms, fn = scattering factor of a particular atom.*

Each diffraction pattern provides information about the crystal from a single direction, consequently the crystal must be rotated in the beam to generate as much information as possible and collect a complete dataset. As the crystal is rotated a different set of resultant waves will be detected at each angle thus a different set of planes will satisfy Bragg's law at each angle.

Many scattering electrons do not lie on Bragg's planes in the unit cell but between the planes at fractional values of $d$. The X- ray waves scattered by these electrons are out of phase and therefore the resultant waves will culminate in a degree of destructive interference with the X-ray waves that obey Bragg's law, resulting in a decrease in the intensity of recorded reflection spots. Consequently, evidence as to the arrangement of atoms in the unit cell relative to Bragg's planes is therefore provided by the relative intensities of the reflection spots. Equally, the arrangement of the asymmetric units (AU) (the smallest repeating unit of the crystal lattice), and

how they are related to each other, for example the space group, is provided in the symmetry of the reflections in the diffraction patterns.

The diffraction pattern displays the intensity of reflections and their symmetry therefore providing much of the information required to generate an electron density map and determine the structure of the protein following data integration, scaling and reduction. However, importantly the phase information is lost from the resultant waves that generate the reflection spots observed in the diffraction pattern. To produce an electron density map from the structure factors, a Fourier transform is required.

Fourier transform is a way of describing a wave as a sum of all its component waves and a necessary data transformation to know the relative phase of these waves to ensure the correct addition. The lack of phase information due to the way the data is recorded is termed the ‘phase problem’ and must be solved in all crystallography experiments. One such way of solving the phase problem is the utilisation of molecular replacement.

#### 3.1.4.3.1 Molecular replacement

Molecular replacement (MR) involves the movement of a model molecule into the most overlapping position and orientation in the unit cell as the unknown molecule. If the model is satisfactorily similar to the unknown molecule, the phases from the model can be used along with the experimental structure factor amplitudes in the determination of the crystal structure of the unknown molecule. A Fourier transform can then be applied which generates an electron density map from the unknown experimental data. The determination of the correct orientation and position of the phasing model compared to the unknown structure in the AU is a fundamental requirement because the phases of the atomic structure factors, and thus the molecular structure factors, depend on the location of the atoms in the unit cell. If the phasing model is reasonably accurate the generated electron density map will display features absent in the original model that can subsequently be built into the crystal structure of the molecule to be determined as the model is solved through a series of refinement and fitting stages.

MR can only be used to solve the phase problem if there is a good structural model available for a large portion of the unknown structure in the crystal. Sequence identity and structural similarities correlate well and therefore if the proteins share > 30% sequence identity the model is probably suitable for MR of the unknown molecule. Solving the phase problem by MR becomes more and more difficult as the sequence identity decreases and other techniques are required to solve the phase problem such as heavy metal soaking using multiple isomorphous replacement.

Other techniques for macromolecular structural analysis, such as Nuclear Magnetic Resonance (NMR) and Cryogenic Electron Microscopy (Cryo-EM) have their advantages and disadvantages (RCSB PDB, 2018). Both techniques are suitable for the analysis of protein and protein complex dynamics (Kovermann, Rogne and Wolf-Watz, 2016; Murata and Wolf, 2018). Additionally, NMR spectroscopy can record signal for disordered and flexible protein regions and the atomic structure of macromolecules in solution but is an extremely costly experimental technique achieving resolutions of up to ~5 Å. Cryo-EM does not require a large protein concentration or crystallisation, and protein structures can be flash-frozen in a multitude of conformations allowing biological mechanisms to be deduced, achieving resolutions of up to ~ 2.5 Å. However, Cryo-EM demands protein samples of high molecular weight and costly, high maintenance equipment.

Additionally, techniques are being developed to overcome some of the common limitations of X-ray crystallography such as in plate data collection, eliminating the process of crystal harvesting, and X-ray free-electron lasers (X-FEL). X-FEL produce flashes of X-ray light with angstrom-level wavelengths and a billion times brighter than those produced before. Importantly, the flashes of light are fentoseconds short with the ability to record functions and forms of proteins such as protein folding and the splitting of a chemical bond (Marx, 2017). However, techniques such as X-FEL are not currently in widespread use.

This chapter therefore describes the expression, purification and activity of DENV-2 RdRp, for use in *in vitro* SHAPE analysis and X-ray crystallography, before presenting the crystal structure of DENV-2 RdRp at 2.2 Å resolution prior to considering our DENV-2 RdRp-SLA RNA co-crystallisation efforts.

## 3.2 Results

### 3.2.1 DENV-2 RdRp expression, purification and activity

#### 3.2.1.1 Generation of the pET-28a expression plasmid

A parental pET-28a expression vector was *Bam* HI and *Xho* I digested and ligated with the DENV-2 RdRp or full-length DENV-NS5 insert, in order to generate a his-SUMO-DENV-2 RdRp or his-SUMO-FL-DENV-NS5 fusion protein for use in downstream purification procedures. *E.coli* DH5α cells were subsequently transformed with the pET28a his-SUMO-DENV-2 RdRp plasmid and grown on selective media in order to isolate successfully transformed colonies. Plasmid DNA was extracted, purified and the construct confirmed as correct via sequencing.

#### 3.2.1.2 Identification of an effective bacterial expression vector

In order to identify the most appropriate strain of *E.coli* BL21 for the expression of full-length DENV-2 NS5 and DENV-2 RdRp, the expression properties of a panel of different *E.coli* BL21 strains were screened on a small scale level. BL21 bacterial strains were transformed with each expression plasmid and single colonies cultured. Test cultures were grown in 10 ml volumes until an O.D of 0.6 was reached and primed cultures subsequently induced with 1 M IPTG. Following induction and growth, the bacterial cells were pelleted, lysed and soluble and insoluble fractions separated and analysed for expression using SDS-PAGE gel electrophoresis and Coomassie stain.

The test expression of full-length DENV-NS5 and DENV-RdRp in BL21 Codon Plus cells is depicted in comparison to BL21 Rosetta 2 and BL21 gold expression cells respectively in figure 39. The overall expression level and solubility of both constructs was found to be higher in BL21 Codon Plus cells and was thus chosen as the expression strain for further expression purification procedures (Fig. 39).

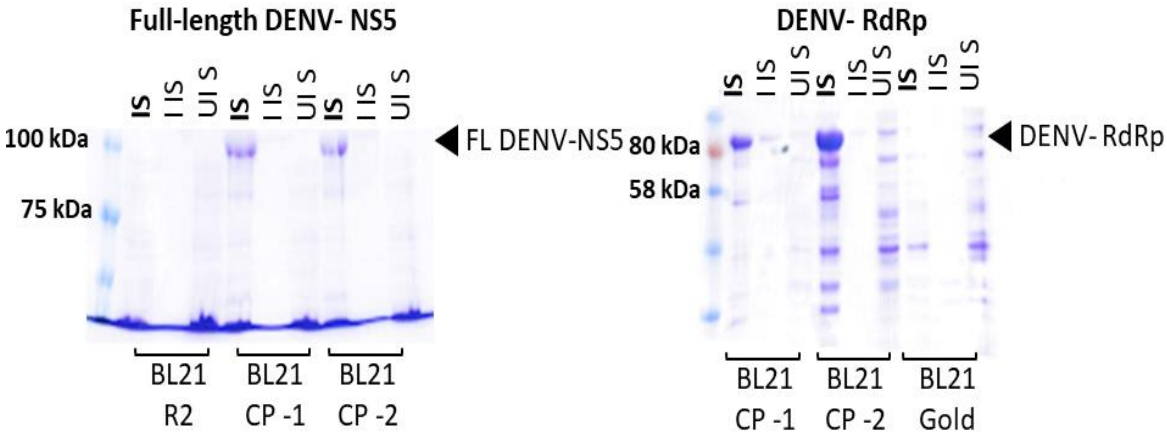



**Figure 39 *E.coli* expression strain testing:**

*Coomassise stain of SDS-PAGE gels showing the expression profile of his-SUMO-FL-DENV-NS5 (left) and his-SUMO-DENV-2 RdRp (right) in BL21 Codon Plus, Rosetta 2 and Gold expression strains. IS = IPTG induced soluble, I IS = IPTG induced insoluble, UI S = un- IPTG induced soluble. His-SUMO-FL-DENV-NS5 = ~100 kDa and his-SUMO-DENV-2 RdRp = ~90 kDa. Size compared to Colour protein standard, broad range (right) or Blue protein standard, broad range (left) (NEB).*

### 3.2.1.3 Optimisation of IPTG induction for protein expression

It is often the case that the solubility and yield of *E.coli* expressed proteins can be improved via optimisation of both the concentration and temperature at which IPTG induction is carried out. IPTG induction at high concentration, at a high temperature (37 °C) results in the fast production of expressed protein, ideal for extremely stable, soluble proteins. However, IPTG induction at low concentrations, at low temperature, over a long time-scale, often results in the slow production of better folded expressed protein of increased solubility. We therefore sought to optimise IPTG induction conditions in the aim of improving protein yield prior to large scale protein expression.

A small-scale culture was grown until an OD of 0.6 was reached, divided and IPTG induced at varying concentrations and temperatures. Following induction and growth, the bacterial cells were pelleted and protein expression levels analysed via SDS-PAGE gel electrophoresis and western blot. The best IPTG induction condition, producing the highest yield of soluble protein, was identified as 0.2 mM at 15 °C for 16 hours and was thus the condition used for large scale bacterial expression.

Figure 40 describes the optimisation of IPTG induction conditions for expression of full-length DENV-2 NS5. The same condition was suitably high-yielding for DENV-2 RdRp expression.

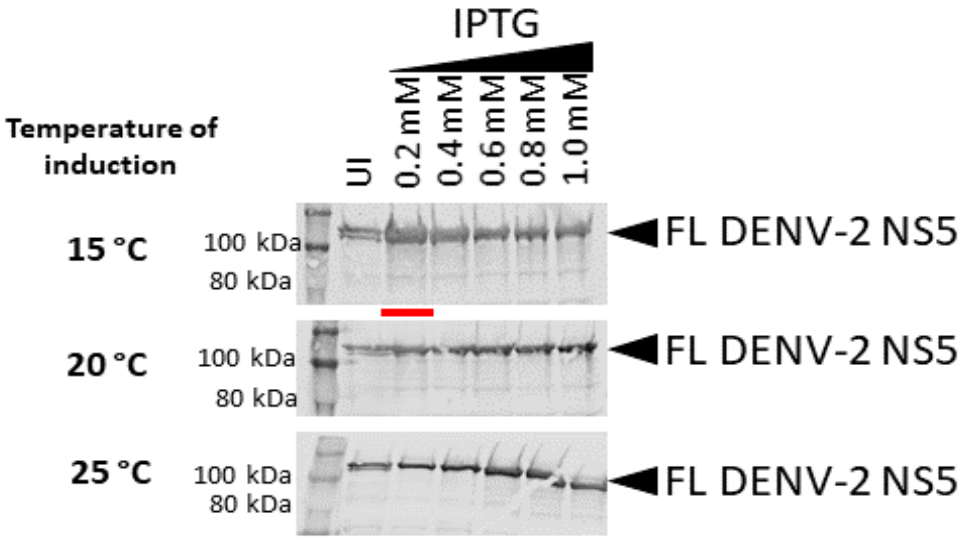



**Figure 40 Optimisation of IPTG induction conditions for protein expression:**

*The optimisation of IPTG induction conditions for the expression of full-length DENV-2 NS5. Size comparison ladder = Colour protein standard, broad range (NEB). The red line indicates the most optimal IPTG induction condition. UI = un-induced control.*

#### 3.2.1.4 Nickel affinity purification of DENV-2 RdRp

Nickel (Ni) affinity chromatography was utilized as the first purification step of DENV-2 RdRp. The purification procedure was optimised to produce a greater final yield of DENV-2 RdRp for both X-ray crystallography and SHAPE mapping procedures. For example, a 2 M NaCl wash was included to remove any contaminating RNA from the protein preparation. Additionally, the purified protein was eluted in a final volume of 5 ml, preventing a yield costly additional concentration step, usually required prior to SEC, due to the utilisation of a 5 ml injection loop prior to loading on to the GE Life Science Superdex S200 26/60 column.

Figure 41 describes the purification of his-SUMO-DENV-2 RdRp using Ni affinity chromatography. Eluted fractions were analysed by SDS-PAGE gel electrophoresis followed by Coomassie stain.

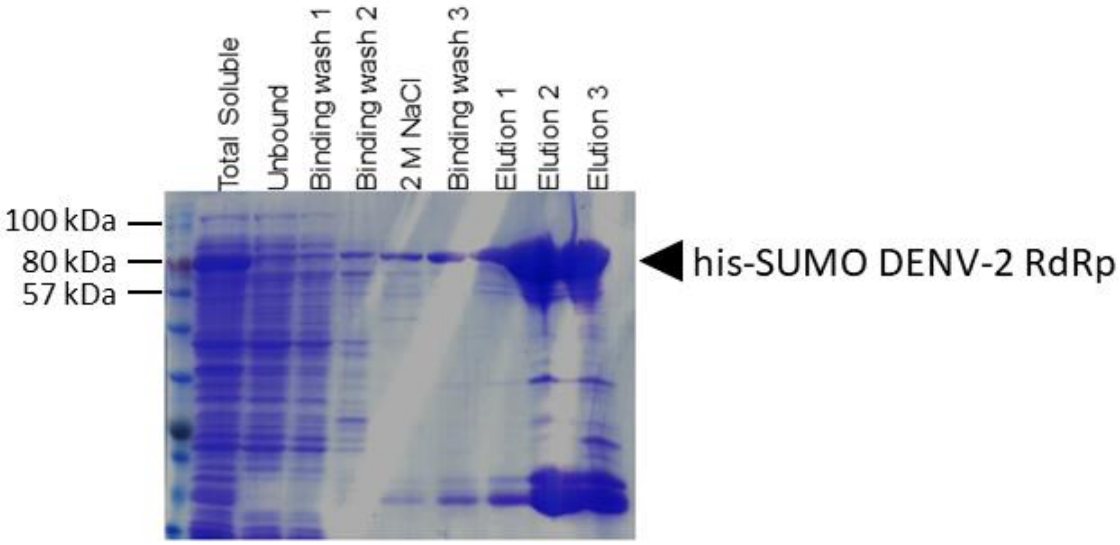



**Figure 41 Nickel affinity purification of his-SUMO-DENV-2 RdRp:**

*The purification of his-SUMO-DENV-2 RdRp by Ni affinity chromatography. Fractions analysed by SDS-PAGE gel electrophoresis and Coomassie stain. Binding wash 1 contained 20 mM Imidazole, Binding Wash 2 contained 50 mM Imidazole, Elution 1 contained 120 mM Imidazole, Elution 2 contained 300 mM Imidazole and Elution 3 contained 500 mM Imidazole. Size compared to Colour protein standard, broad range (NEB).*

#### 3.2.1.5 SUMO protease cleavage of his-SUMO tag removal

Following Ni affinity purification, the his-SUMO tag was cleaved during dialysis. The reasons for tag removal during dialysis were two-fold. Firstly, tag removal prevented its subsequent crystallisation during X-ray crystallography procedures. Secondly, tag removal was carried out for 16 hours at 4°C during dialysis of the protein preparation to ensure efficient removal of the tag but also to remove high-levels of imidazole from the purified protein preparation. A high level of imidazole is known to inhibit ionic interactions and affect protein concentration determination procedures. Therefore, the his-SUMO tag was removed during dialysis of the purified protein preparation prior to SEC.

#### 3.2.1.6 Purification of DENV-2 RdRp via size exclusion chromatography

Following Ni affinity purification, SEC was utilised to purify SUMO protease cleaved DENV-2 RdRp away from contaminant proteins, including the SUMO protease itself and the poly-histidine tag. Cleaved DENV-2 RdRp was loaded into a 5 ml injection loop prior to loading onto a GE Life Science Superdex 26/60 S200

column. The gel filtration system was run to completion at a flow rate of 0.3 ml/min with 3 ml fractions collected. A UV 280 nm chromatogram was used to identify fractions containing protein. Peak fractions were collected and analysed by SDS-PAGE gel electrophoresis and Coomassie stain to determine protein purity (Fig. 42).

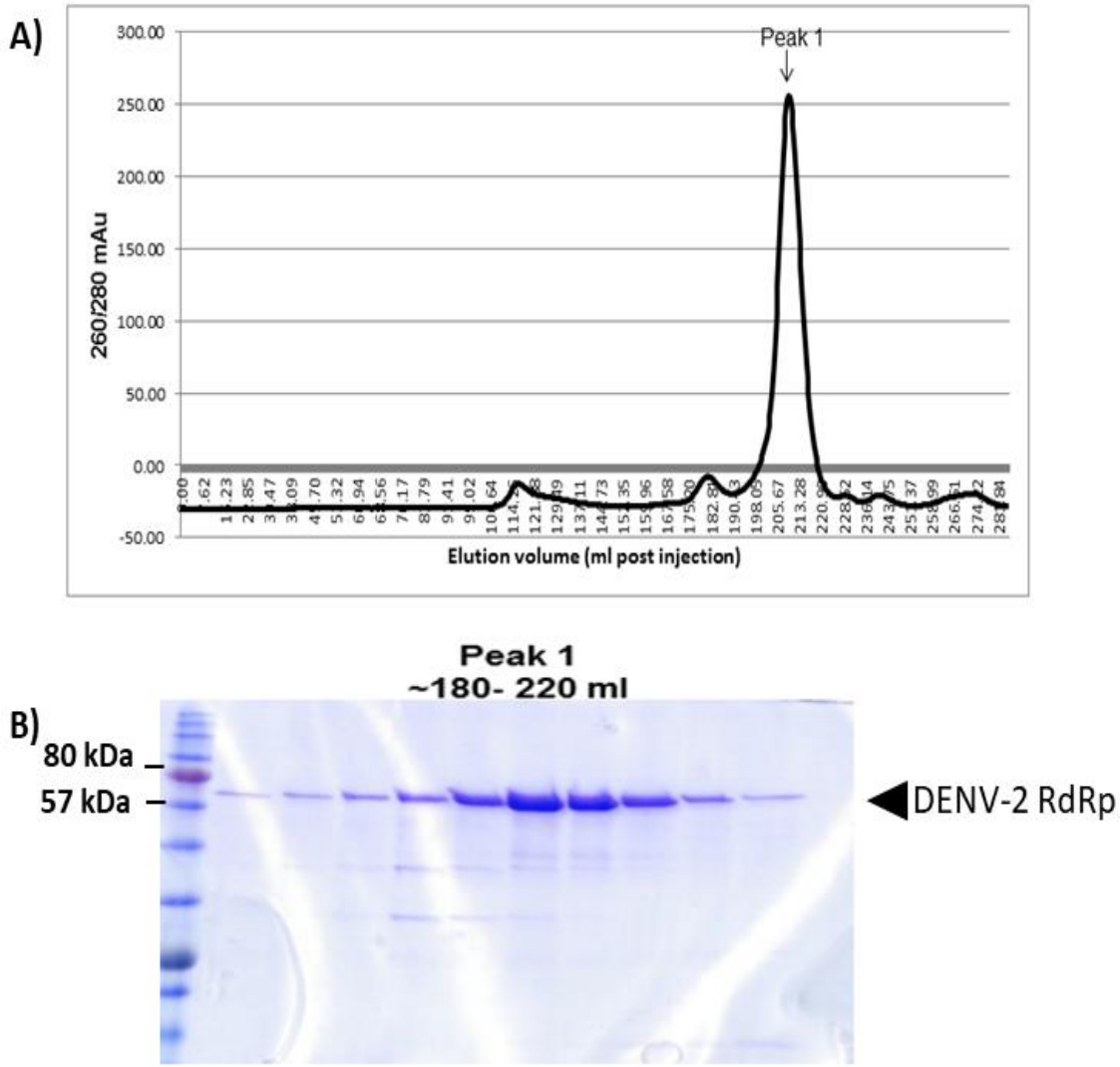



**Figure 42 The use of size exclusion chromatography in the purification of DENV-2 RdRp:**

***A)*** *UV 280 nm chromatogram trace. Peak 1 corresponds to monomeric DENV-2 RdRp using a S200 AKTA SEC column.* ***B)*** *Fractions corresponding to Peak 1 analysed by SDS-PAGE gel electrophoresis and Coomassie stain. Size compared to Colour protein standard, broad range (NEB).*

### 3.2.1.7 Identification of DENV-2 RdRp by western blot

Following Ni affinity chromatography, SEC and subsequent concentration, purified DENV-2 RdRp was identified using western blot, as seen in figure 43.

**Figure 43 The identification of purified DENV-2 RdRp by western blot:**

*Identification by western blot of purified DENV-2 RdRp using a DENV-2 NS5 antibody. DENV-2 RdRp is ~ 73 kDa. Size compared to Colour protein standard, broad range (NEB) ladder.*

### 3.2.1.8 Analysis of the purity of DENV-2 RdRp

X-ray crystallographic studies demand protein of a high concentration and of ~ 95% purity. We therefore determined the purity of purified DENV-2 RdRp, prior to X-ray crystallographic studies, via SDS-PAGE gel electrophoresis and Coomassie stain. Densitometric analysis was performed using ImageJ which determined a preparation purity of ~95% (Fig. 44).

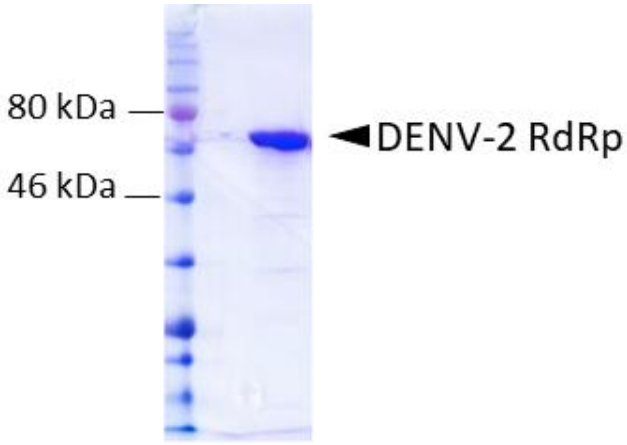



**Figure 44 Analysis of the purity of DENV-2 RdRp:**

*The purity purified DENV-2 RdRp was determined via SDS- PAGE gel electrophoresis, Coomassie stain and densitometric analysis using ImageJ. DENV-2 RdRp is ~73 kDa and size compared to a Colour protein standard, broad range (NEB) ladder.*

### 3.2.1.9 Circular Dichroism of DENV-2 RdRp

Circular dichroism (CD) was performed to confirm the secondary structure of DENV-2 RdRp. The differential absorbance for different protein structures allows for both the type and extent of secondary structure formation of a protein to be determined. Analysis of the CD spectrum for monomeric DENV-2 RdRp displayed characteristic troughs at 222 nm and 208 nm, confirming that the purified DENV-2 RdRp was folded correctly and was predominantly α-helical, the primary conformation displayed by other DENV RdRp serotype structures (Fig. 45).

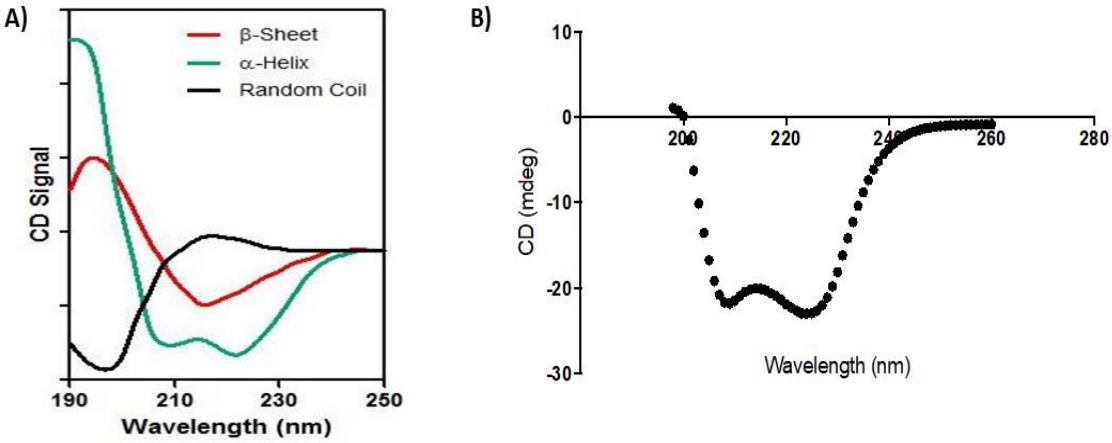



**Figure 45 Circular Dichroism spectrum of monomeric DENV-2 RdRp:**

***A)*** *Characteristic CD spectra. Taken from* (Da Yu Protein Sciences, 2018). ***B)*** *CD spectrum of monomeric DENV-2 RdRp protein at 0.2 mg/ml, troughs at 222 and 208 nm indicate a predominately α-helical conformation.*

### 3.2.1.10 Determination of DENV-2 RdRp polymerase activity

To ensure active protein was used for both *in vitro* SHAPE and crystallographic experimentation, the *de novo* RNA synthesis capability of DENV-2 RdRp was confirmed, in comparison to the activity of HCV NS5B and NS5B GND. GND refers to an active site mutant of the HCV NS5B rendering the polymerase unable to replicate. RNA was filter purified and α-$^{32}$P-ATP incorporation into the nascent RNA chain quantified using scintillation counting. In comparison to the negative control, DENV-2 RdRp was confirmed as active (Fig. 46).

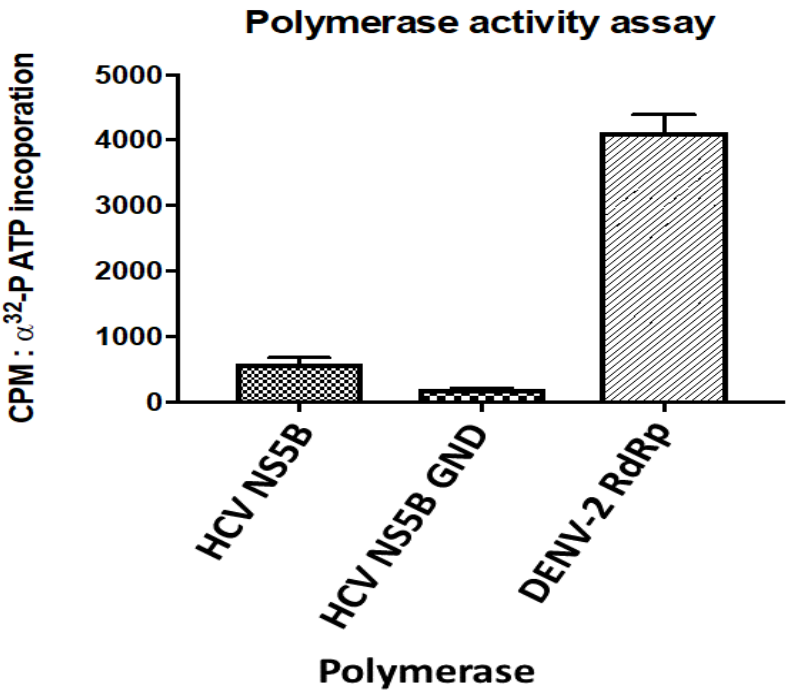
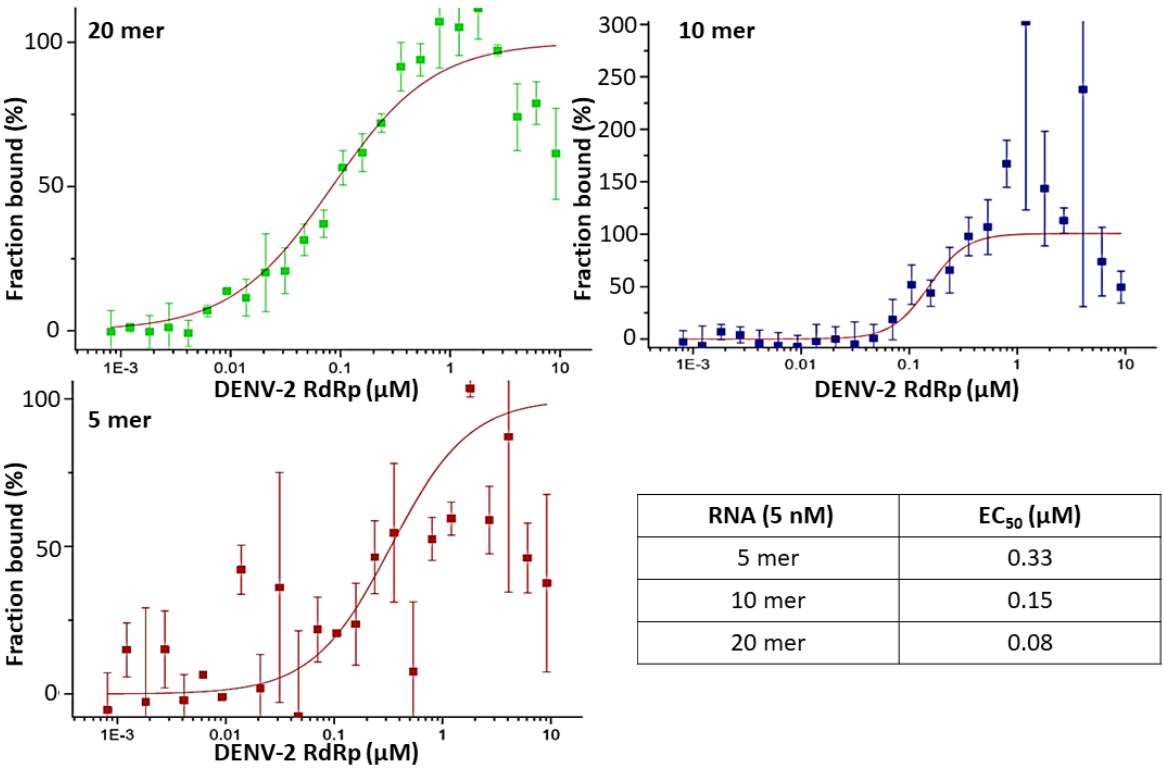



**Figure 46 DENV-2 RdRp Polymerase activity assay:**

*De novo RNA synthesis capability of DENV-2 RdRp, in comparison to HCV NS5B and NS5B GND (a HCV NS5B mutant unable to replicate). Activity determined following filter RNA purification and α$^{32}$P-ATP incorporation quantified using scintillation counting.*

### 3.2.1.11 Analysis of the RNA binding capability of DENV-2 RdRp by Fluorescence Anisotropy

The technique of Fluorescence Anisotropy (FA) was utilised to confirm the RNA-binding capability of DENV-2 RdRp and to select suitable, binding, co-crystallisation partners for X-ray crystallographic studies. A panel of short 3´FI tagged DENV-2 SLA RNAs, corresponding to different lengths of the confirmed DENV RdRp binding site on SLA, were selected for testing. All three of the oligonucleotides bound RdRp with the 20 mer and 10 mer displaying the lowest $EC_{50}$s of 0.08 and 0.15 µM respectively. All three RNAs were subsequently used for RdRp-SLA co-crystallisation experimental procedures. $EC_{50}$ was calculated from the concentration of protein required for 50% RNA to be bound (Fig. 47).




| RNA (5 nM) | $EC_{50}$ (µM) |
|---|---|
| 5 mer | 0.33 |
| 10 mer | 0.15 |
| 20 mer | 0.08 |

**Figure 47 The RNA binding capability of DENV-2 RdRp analysed by FA:**

*Binding isotherms and $EC_{50}$s for the panel of 3´FI tagged DENV-2 SLA RNAs. All graphs were plotted using OriginPro 9.1 (Origin Lab) according to:* $y = \frac{A1-A2}{1+\left(\frac{x}{x^0}\right)p} + A2$; *y is the fraction of RNA bound,* $x$ *is the protein concentration, A1 is the initial value of y, A2 is the final value of y,* $x^0 0$ *is the dissociation constant ($K_d$) and p is the Hill coefficient.*

#### 3.2.1.12 Analysis of DENV-2 RdRp-RNA interactions by native Electromobility shift assay

To further confirm the RNA binding capability of DENV-2 RdRp and analyse the specificity of DENV-2 RdRp binding to DENV-2 RNA templates with or without the 5´ promoter SLA, we utilised native Electromobility shift assays (EMSAs). The RNA binding capability and specificity of DENV-2 RdRp for radioactively labelled (hot) SLA RNA template at mammalian temperature was tested by competition against a panel of non-radioactively labelled (cold) DENV-2 RNAs, all of the same length (~160 nts). Two cold RNAs corresponded to the first 160 nts of the DENV RNA genome with SLA (+SLA) and without SLA (ΔSLA) and the third RNA corresponded to the first 160 nts of the C terminus of DENV genomic RNA (3´ UTR). Using EMSAs, we further confirmed that DENV-2 RdRp was able to bind to the + SLA transcript with the greatest affinity therefore indicating specific binding to SLA, greater than that of the 3´ UTR, and also that DENV-2 RdRp was able to bind to +SLA RNA at insect physiological temperature (Fig. 48).

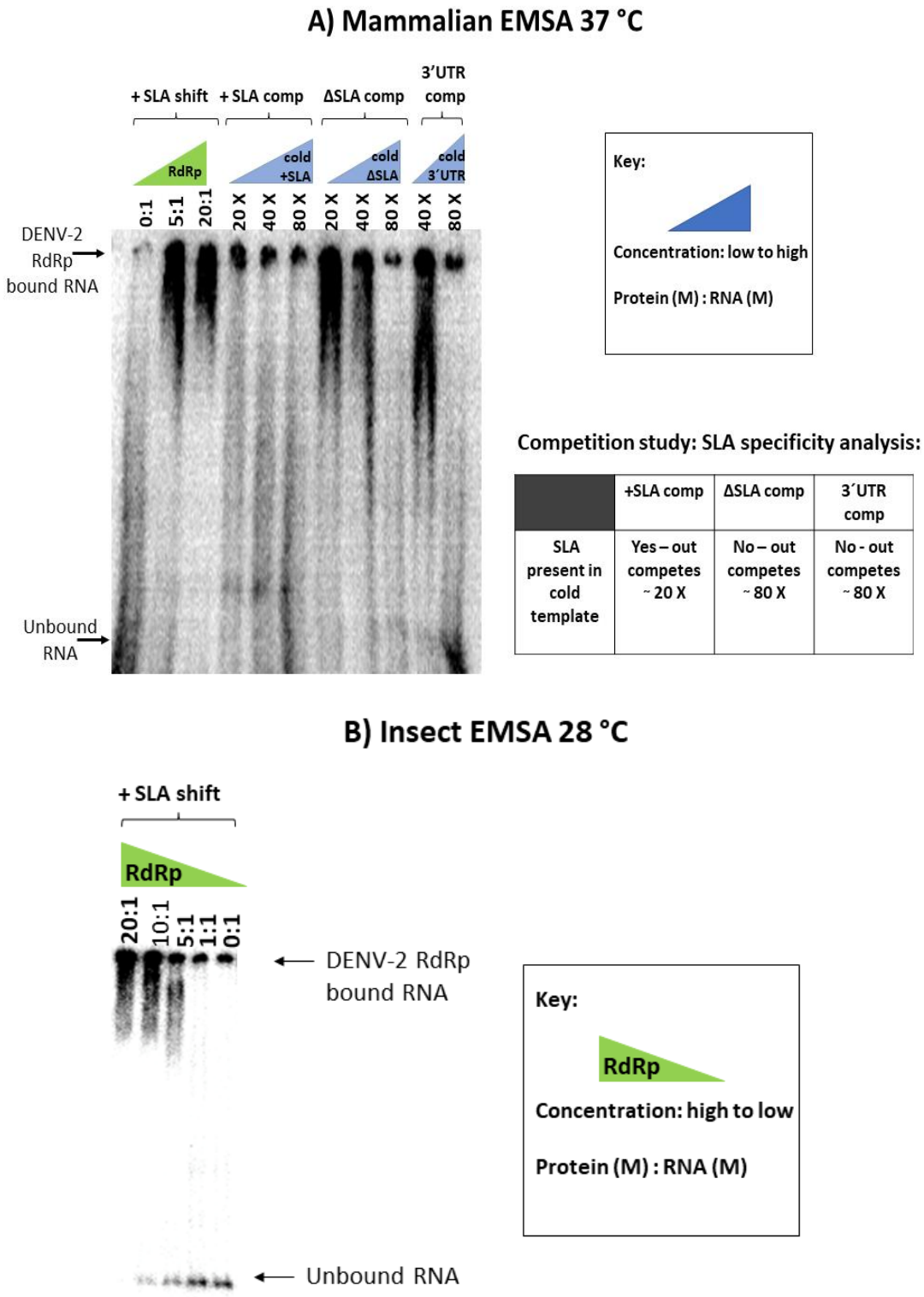



**Figure 48 Analysis of DENV-2 RdRp-RNA binding by native Electromobility Shift assay:**

*Autoradiographs and analysis of DENV-2 RdRp binding and specificity to SLA by native Electromobility shift assay (EMSA) at mammalian* ***(A)*** *and insect* ***(B)*** *physiological temperatures. DENV-2 RdRp binds most specifically to the +SLA RNA transcript containing the 5´ promoter SLA. Comp = competition. X = fold molarity over radioactively labelled +SLA transcript.*

### 3.2.2 X-ray crystallographic study of unbound DENV-2 RdRp

Prior to attempting to determine the SLA RNA-DENV-2 RdRp crystal structure we sought to determine the crystal structure of SLA unbound DENV-2 RdRp.

#### 3.2.2.1 Crystal growth

Purified, active DENV-2 RdRp was produced (described in section 3.2.1) and screened for initial crystallisation conditions using the commercially available Joint Centre for Structural Genomics Core (JCSG) suite screens (I-IV). The JCSG core suite screens were designed to test the most successful and reproducible conditions from the JCSG. The use of these screens samples a wide-range of precipitants, salts, buffers, pH, polymers and organic molecules in a high-throughput 96 well format, allowing for the testing of 384 different conditions. These screens were also compatible with an automated robot system which was utilised throughout experimentation to efficiently and reproducibly set up sitting-drops with various protein:mother liquor ratios across all conditions.

Several conditions resulted in initial ‘hits’ that appeared anywhere between 12 hours to 1 week and ranged in size and number within a drop. The most promising condition (based on visual appearance) contained 1 M Na3 Cit and 0.1 M CHES pH 9.5, produced the largest crystals with few splits and sharp edges and was therefore selected for further experimental optimisation (Fig. 49).

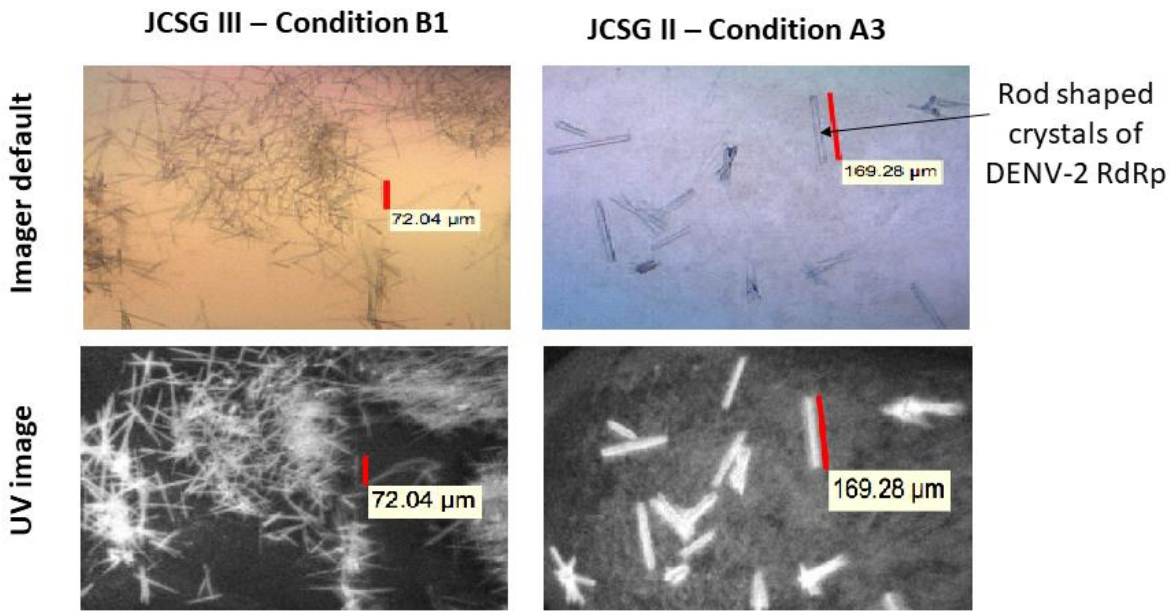



**Figure 49 Crystals grown in mother liquor droplets during JCSG sparse matrix screening:**

*Crystals monitored for growth using the Rockmaker imager® under default and UV settings confirming the crystallisation of DENV-2 RdRp protein.*

Crystal growth optimisation was performed in 3 drop 96 well format and was conducted using the sitting drop crystallisation technique. The molarity of Na3 Cit was varied alongside the pH of CHES and additionally the temperature at which the crystals grew (4°C and room-temperature). Different droplet sizes and also different protein:mother liquor ratios were utilised. Sitting droplets of 0.4 µl volume, 50:50 protein: mother liquor ratio generally yielded diffracting crystals, suitable for X-ray crystallography experiments. Due to the success of the initial optimisation screens further optimisation using additive screens was not utilised in this case.

#### 3.2.2.1.1 Streak seeding

To increase the number of crystals obtained for data collection and soaking procedures streak seeding was utilised. Streak seeding is a method used to introduce pre-formed crystal nuclei into a fresh protein:mother liquor drop to control nucleation, encouraging better crystal formation within the drop (described further in section 3.1.4.1.1). This technique was successfully utilised to produce a greater number of crystals for data collection (Fig. 50).

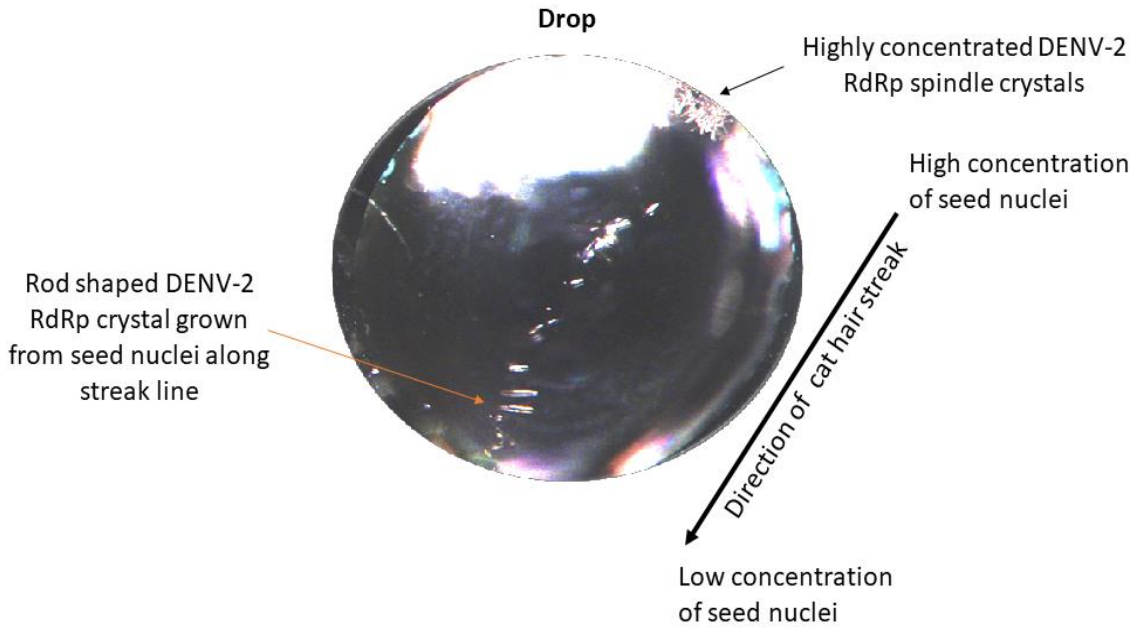



**Figure 50 DENV-2 RdRp streak seeding:**

*Streak seeding is a method used to introduce pre-formed crystal nuclei into a fresh protein:mother liquor hanging drop to control nucleation. A sub-optimal crystal is first mechanically sheared or stroked using a small fibre (Cat hair) and the hair is then streaked through the fresh protein:mother liquor drop. This process induces nucleation along the streak line, which is subsequently diluted across the drop. Crystal growth was monitored periodically using light microscopy.*

#### 3.2.2.2 Crystal harvesting

Once single crystals were obtained (>20 µm with minimal splits of cracks) they were picked by manipulating them out of their protein:mother liquor drops using small loops utilizing paratone N oil as the cryoprotectant (Fig. 51).

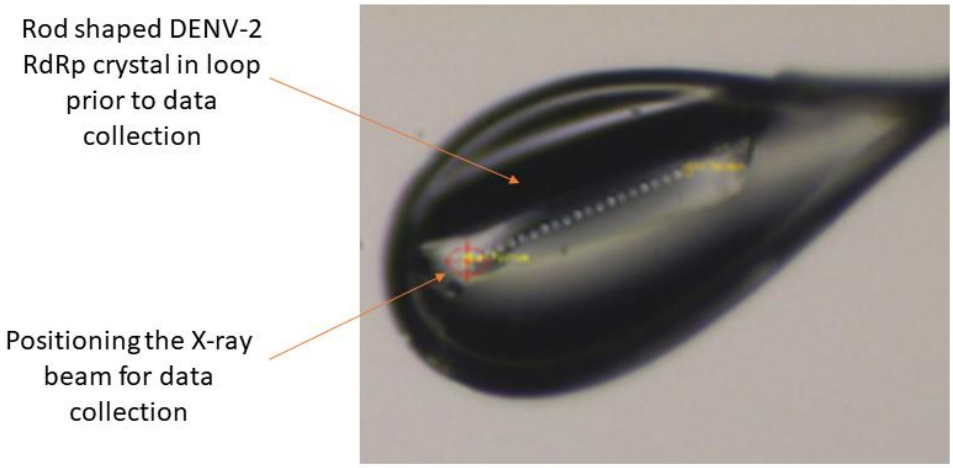



**Figure 51 Loop containing a rod-shaped DENV-2 RdRp crystal at the synchrotron:** *The crystal was picked in Paratone N oil and cryocooled in liquid nitrogen and maintained at 100 K.*

#### 3.2.2.3 Data collection

An example of a diffraction pattern collected from an unbound DENV-2 RdRp crystal is shown in figure 52. The diffraction pattern is the raw data of an X-ray crystallography experiment.

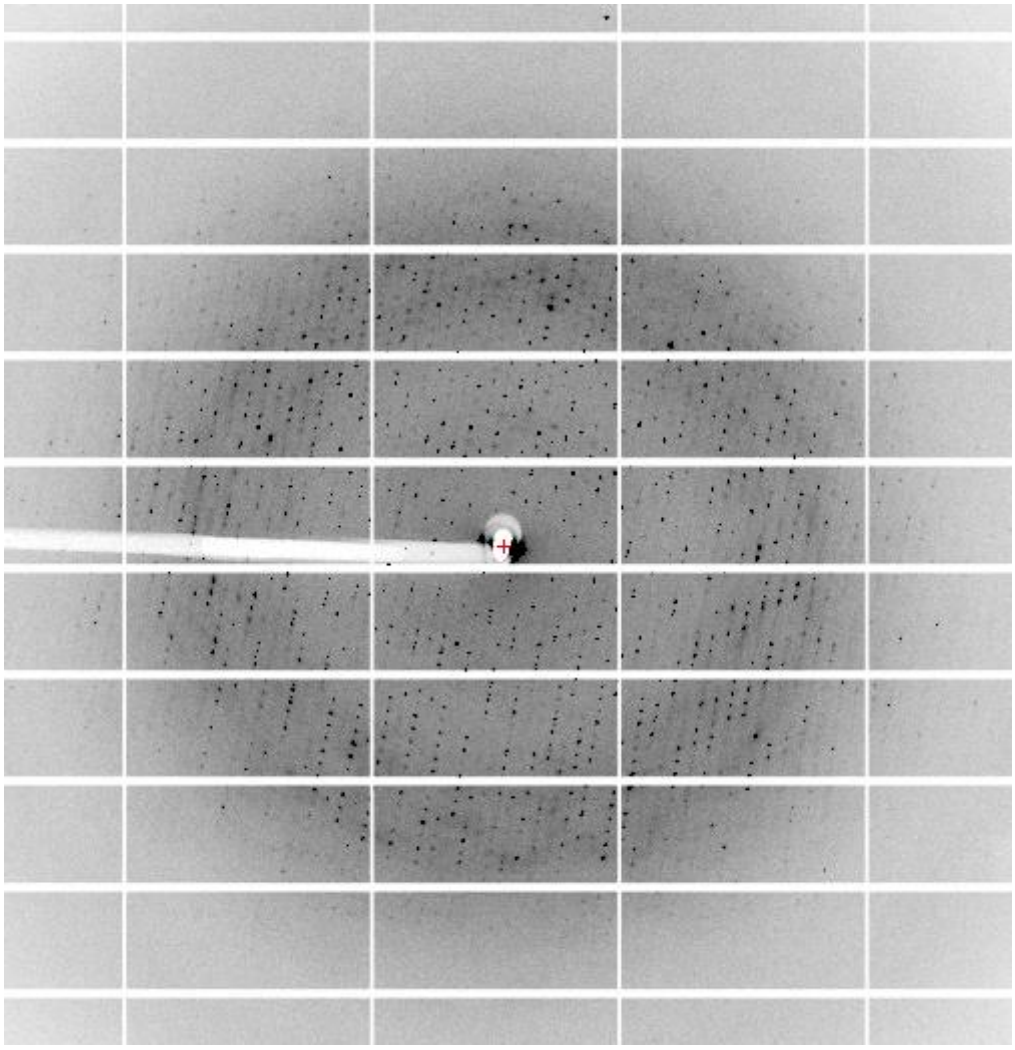


**Figure 52 Diffraction pattern of RNA unbound DENV-2 RdRp:**

*Diffraction pattern of RNA unbound DENV-2 RdRp. The orderliness of the reflection spots together with their intensities provides the information needed to determine the X-ray structure of DENV-2 RdRp. However, the phase information cannot be obtained experimentally from the diffraction pattern.*

#### 3.2.2.3.1 Data processing

Data processing (integration, scaling, merging) and subsequent structure determination by molecular replacement was carried out by Dr. Chi Trinh and the statistics of the refined unbound DENV-2 RdRp model are shown in table 3.

**Table 3 Data processing and scaling statistics of DENV-2 RdRp:**

| Dataset | **DENV-2 RdRP** |
|---|---|
| Source | Diamond Light Source |
| Beamline | I03 |
| Wavelength (Å) | 0.976 |
| Resolution range (Å) | 90.38 – 2.20 (2.27 – 2.20) |
| Space Group | *I* 2 2 2 |
| Unit–Cell parameters (Å) | A = 208.78, b = 57.77, c = 65.94, α = β = γ = 90.00 |
| Completeness (%) | 99.6 (99.9) |
| Total reflections | 228607 |
| Unique reflections | 33757 |
| Redundancy | 6.4 (6.8) |
| $I/\sigma(I)$ | 17.2 (2.9) |
| $R_{merge}$ (%) | 4.9 (64.0) |
| $R_{pim}$ (%) | 3.2 (39.8) |
| $CC_{1/2}$ | 0.99 (0.88) |
| **Refinement Statistics** | |
| Resolution range (Å) | 71.91 – 2.2 |
| *R* factor (%) | 22 |
| $R_{free}$ (%) † | 27 |
| No. of protein non-H atoms | 4577 |
| No. of water molecules | 13 |
| No. of Zn ions | 2 |
| R.m.s.d bond lengths (Å) | 0.010 |
| R.m.s.d bond angles (°) | 1.267 |
| Molprobidity sore | 1.56 (98th percentile) |
| **Ramachandran plot ((%statistics (%)** | |
| Favoured region (%) | 95.19 |
| Outliers (%) | 0.37 |

#### 3.2.2.4 The crystal structure of DENV-2 RdRp

The crystal structure of DENV-2 RdRp unbound to SLA promoter RNA was determined at 2.2 Å resolution, to our knowledge the first unbound crystal structure of the DENV-2 RdRp domain.

DENV-2 RdRp adopts a canonical right-handed polymerase fold, similar to that of DENV-3 RdRp, comprised of fingers, palm and thumb subdomains (Figs. 53 and 54). Since the polymerase is unbound to SLA RNA it is in the 'closed conformation' whereby the priming loop in the thumb subdomain is extended towards the active site of the polymerase occluding access to the template-binding channel. As seen in figure 53, the structure of the priming loop is incomplete. Due to the flexibility of the priming loop the electron density map for this region was too poor of quality to enable the modelling of the protein structure in this region.

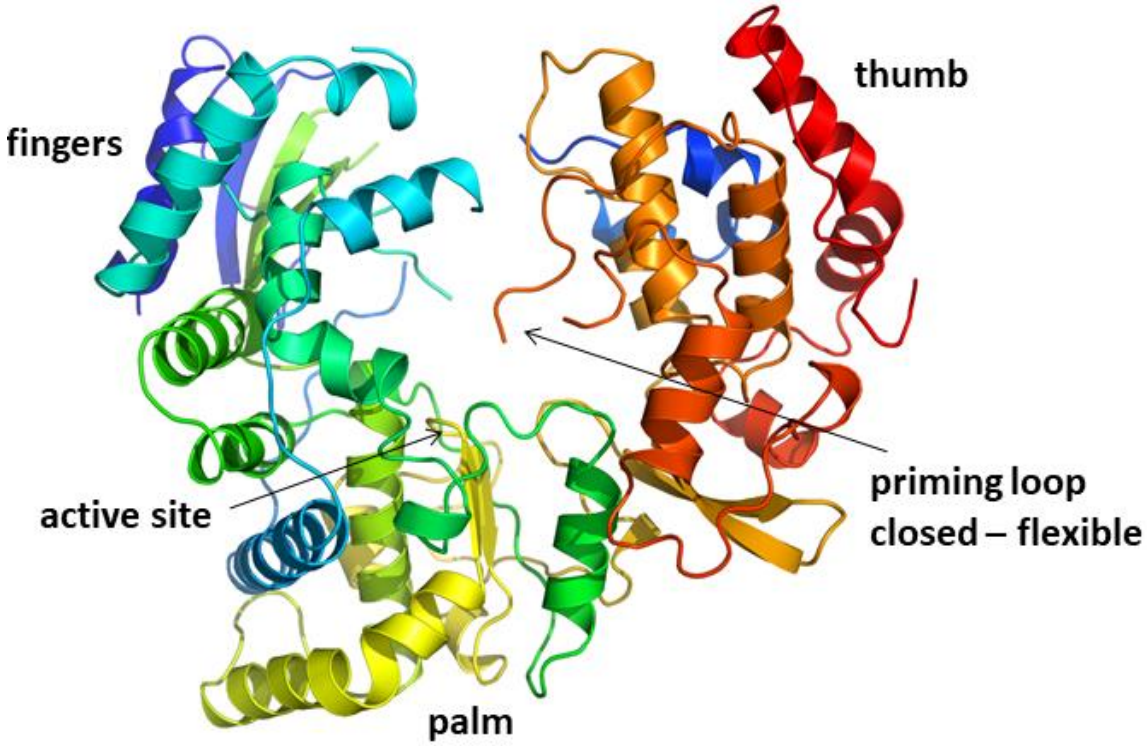



**Figure 53 DENV-2 RdRp crystal structure at 2.2 Å resolution:**

*The annotated crystal structure of DENV-2 RdRp at 2.2 Å resolution. The DENV-2 RdRp is of typical right handed viral polymerase structure and is in the closed conformation as indicated by the inward position of the priming loop.*

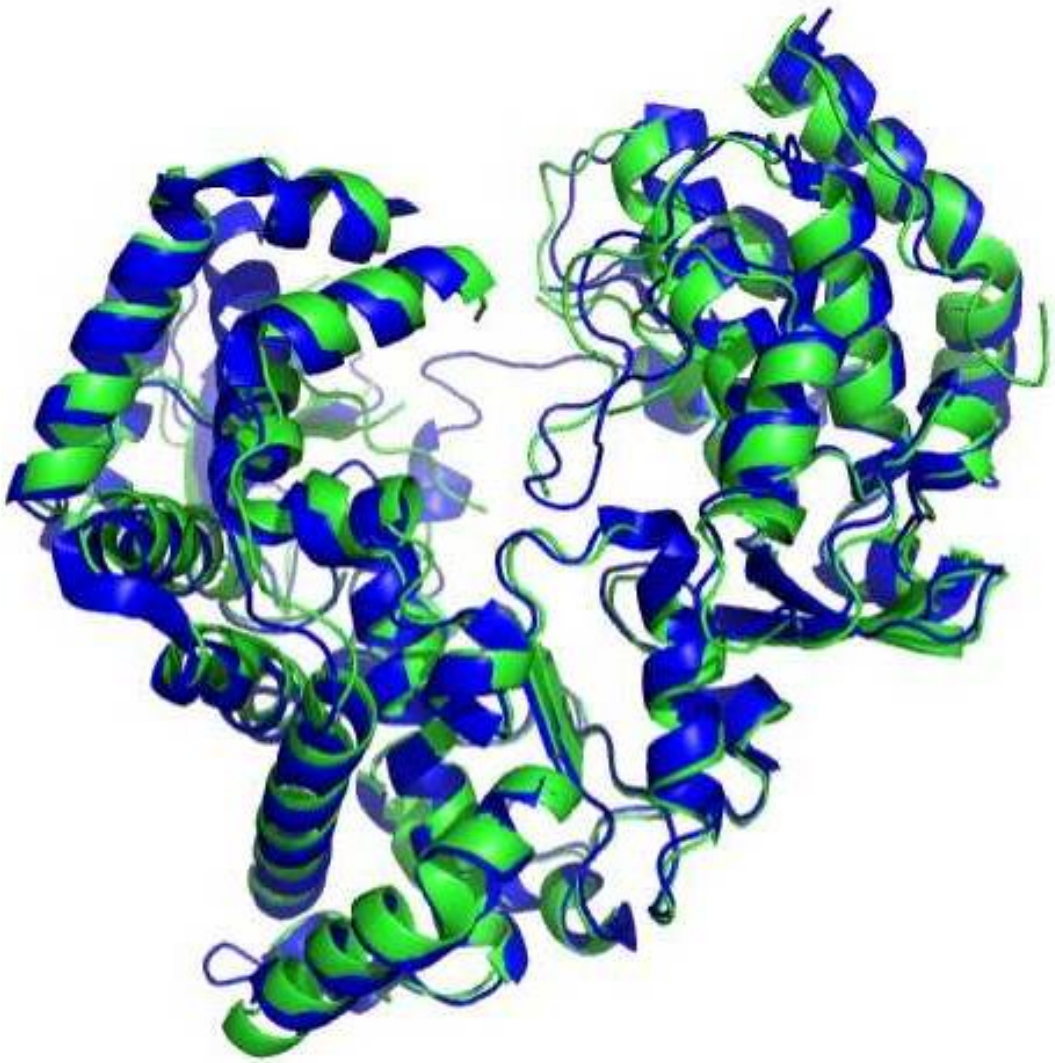


**Figure 54 DENV-2 RdRp superposed onto the crystal structure of DENV-3 RdRp:**

*The determined crystal structure of DENV-2 RdRp superposed onto the crystal structure of the DENV-3 RdRp domain (PDB: 2J7U). The two proteins overlay in a similar manner.*

### 3.2.3 Co-crystallisation of DENV-2 RdRp with SLA RNA

It is predicted that following RdRp binding to SLA and completion of the rate limiting replication initiation steps, the RdRp is suggested to transition from a 'closed conformation' to an 'open conformation' to accommodate the growing dsRNA chain and replicate the viral genome. Structural interrogation of this protein-RNA interaction would, for the first time, reveal the site of RdRp interaction with SLA RNA and critically examine RdRp conformational change during the initiation of DENV genome replication – a significant anti-viral target.

To achieve this goal, we sought to co-crystallise DENV-2 RdRp with SLA RNA. Despite significant effort utilising both co-crystallisation and soaking procedures of different lengths of SLA RNA we were unable to obtain a crystal structure of DENV-2 RdRp bound to SLA RNA. Following co-crystallisation procedures, the produced crystals diffracted poorly or not at all, suggestive of crystal lattice disruption. Furthermore, upon crystallisation the DENV-2 RdRp active site could be packed, due to neighbouring molecules within the lattice, in such a way that renders the active site inaccessible to RNA and subsequently the RNA molecules tested were unable to bind.

## 3.3 Discussion

### 3.3.1 Expression and purification of DENV-2 RdRp

As described in this chapter, the expression and purification of DENV-2 RdRp was successfully and reproducibly achieved to sufficient purity and concentration for the downstream applications of *in vitro* SHAPE mapping and X-ray crystallography.

### 3.3.2 The activity of DENV-2 RdRp

We determined purified DENV-2 RdRp to be correctly folded and active via several means; CD, polymerase activity assay, FA and EMSA.

We determined DENV-2 RdRp to be correctly folded using CD. As predicted, the folded secondary structure of purified DENV-2 was predominantly alpha helical in conformation. An extension of this study would be to combine the purified protein with SLA RNA and analyse whether the protein differentially folds in the presence of SLA RNA.

The purified DENV-2 RdRp was capable of *de novo* RNA synthesis as determined using a polymerase activity assay in comparison to both an active HCV NS5B polymerase and an inactive GND HCV NS5B polymerase.

We determined the RNA binding capability of purified DENV-2 RdRp using both FA and EMSA. FA was utilised to analyse the RNA-binding capability of DENV-2 RdRp to a range of SLA RNAs and to select optimal binding RNA partners for downstream co-crystallisation procedures. All three SLA oligonucleotides tested were able to bind RdRp. We found the oligonucleotides of the longest length (20 mer) to have the most optimal $EC_{50}$ of 0.08 µM followed by the 10 mer with an $EC_{50}$ of 0.15 µM. It is notable that DENV-2 RdRp binds to longer RNAs more effectively than shorter oligonucleotides. It may be that longer oligonucleotides fit better into the binding cleft of the protein, forming more protein-RNA interactions, resulting in better binding and a lower $EC_{50}$. Since FA was conducted at room-temperature, an extension of this study would be to consider both mammalian and insect temperature and its effect on protein-SLA RNA binding. Further experimentation should consider the requirement of SLA RNA and the effect of random oligonucleotides on the specificity of protein-RNA binding.

EMSAs were used to examine the binding requirements of DENV-2 RdRp to DENV-2 RNA, with and without SLA and to the 3´UTR, taking into consideration the mammalian and insect physiological environment. In line with published EMSA experiments, we determined through competition analysis, that DENV-2 RdRp binds with greater specificity to the DENV-2 RNA template containing SLA over the DENV-2 RNA template without SLA and the 3´UTR at mammalian temperature (Iglesias, Filomatori and Gamarnik, 2011). This analysis further confirms the requirement of SLA for efficient RdRp binding as a competition molarity of >20-fold was required to outcompete the bound SLA RNA with unlabelled RNA template containing the SLA promoter. Since the DENV-2 RdRp is an RNA binding protein, it does bind both the DENV-2 RNA template without SLA and the 3´UTR but to a significantly lesser extent. This was evidenced as it was harder to out compete the DENV-2 RNA template containing SLA RNA, requiring unlabelled competition template molarities of > 40-fold over the bound SLA containing template. We further extended our studies to examine the binding of DENV-2 RdRp to radiolabelled DENV-2 RNA containing SLA for the first time considering insect physiological temperature (28 °C). We have determined that DENV-2 RdRp does bind a DENV-2 RNA template containing SLA at insect physiological temperature. Our protein-RNA EMSA analysis could be extended to analyse the binding of mutant DENV-2 RdRp on DENV-2 RNA templates containing SLA promoter RNA interrogating the requirements of the protein to bind SLA template RNA.

Therefore, we determined the purified DENV-2 RdRp as correctly folded and active via several methods and consequently viable for *in vitro* SHAPE mapping and X-ray crystallographic procedures.

### 3.3.3 The crystal structure of DENV-2 RdRp at 2.2 Å resolution and co-crystallisation studies

To date and to our knowledge, we present the first crystal structure of the unbound DENV-2 RdRp domain, solved by molecular replacement, at a resolution of 2.2 Å.

The structure of the RNA unbound NS5 polymerase domains of several Flaviviruses , such as DENV, JEV and WNV, have previously been determined (Malet *et al.*, 2007; Yap *et al.*, 2007; Lu and Gong, 2013). Like other Flaviviral RdRps, the DENV-2 RdRp adopts a canonical right-hand structure comprised of

fingers, thumb and palm sub-domains. The active site is completely encircled and found at the centre of the molecule at the end of the RNA template and NTP entry tunnels. Additionally, the priming loop extends from the thumb subdomain towards the active site. The priming loop has been shown to be essential for *de novo* RNA synthesis in a primer-independent manner. As evidenced by the position of the priming loop, the DENV-2 RdRp is in the 'closed conformation' prior to RNA template elongation.

Despite our efforts, we were not able to obtain a crystal structure of the DENV-2 RdRp bound to SLA template RNA. We attempted both co-crystallisation and soaking X-ray crystallographic studies to no avail. Upon data collection, sufficient diffraction was not achieved or even at all, suggestive of crystal lattice interference. The way DENV-2 RdRp packs within the crystal lattice may make the active site inaccessible to the input RNA and may not be able to crystallise in this manner. Therefore, a structure of the DENV RdRp in the replicative elongation state remains to be elucidated.

To our knowledge the HCV NS5B is the only *Flaviviridae* polymerase to have been structurally interrogated using X-ray crystallography in the elongation state (Appleby *et al.*, 2015). Similarly, traditional soaking and co-crystallisation procedures failed to yield bound elongation state structures. Therefore, multiple stalled enzyme-ternary complexes were prepared in a step-wise manner. Firstly, HCV NS5B from the JFH1 genotype 2a isolate was used for cloning and manipulation procedures, with the benefit that the polymerase from this isolate was extraordinarily efficient at RNA synthesis. Secondly, a conformational stabilisation strategy was employed whereby NS5B resistance was conferred to a guanosine analogue inhibitor, following selective pressure and the subsequent cloning of a triple mutant, which displayed 1.5-fold the initiation activity of wild-type. Therefore, it was reasoned that this mutant polymerase might stabilise a specific conformational state along the initiation pathway. This was indeed the case and a crystal structure of the HCV NS5B was determined displaying a substantial structural rearrangement of the polymerase. During further co-crystallisation procedures nucleotide di-phosphates were used acting to stall the polymerase, allowing for structural determination in a catalytically relevant state. Additionally, it

was also found that increased $Mn^{2+}$ in solution lowered the Michaelis constant ($k_m$) of the initiating nucleotide and increased the activity of NS5B 20-fold relative to $Mg^{2+}$ in solution. Subsequently, the inclusion of $Mn^{2+}$ during experimentation acted to stabilise the incoming nucleotide allowing for soaking crystallography experiments targeting distinct polymerase elongation assemblies.

Moving forward, the strategies harnessed to determine the crystal structure of the HCV NS5B polymerase in the elongation state should be translated to the structural analysis of the elongation state of the DENV-2 RdRp. An isolate of DENV with increased replicative ability should be selected and an RdRp with an acquired resistance to a suitable nucleoside analogue should be expressed and purified. The activity of the resistant DENV RdRp should be considered in different solution conditions, for example the inclusion of $Mn^{2+}$, to determine the most activity optimal solution condition prior to X-ray crystallographic study of the DENV-2 RdRp. Also, the co-crystallisation of the DENV-2 RdRp with nucleotide di-phosphates should be considered as an alternative to SLA RNA to be able to stall the DENV-2 RdRp in the 'open conformation' or the elongation state allowing for structural interrogation.

In summary, we present the first crystal structure of the DENV-2 RdRp, solved by molecular replacement, to a resolution of 2.2 Å. The strategies employed to solve the structure of the HCV NS5B polymerase in the elongation state should be translated to DENV RdRp X-ray crystallographic study. However, the determination of the crystal structure of RNA unbound DENV-2 RdRp has increased our knowledge of a protein critical to DENV replication in an understudied serotype and broadened our horizons in the potential for structure-based drug design, desperately required in the treatment of DENV.

# Chapter 4 *In vitro* SHAPE mapping of DENV-2 genomic RNA

## 4.1 Introduction

RNA structure is critical to the DENV life cycle, clearly evidenced by genomic cyclisation, and the range of fundamental *cis*-acting replication elements located throughout the DENV genome (Clyde and Harris, 2006; Clyde, Barrera and Harris, 2008; Lodeiro and Filomatori, 2009; Villordo, Alvarez and Gamarnik, 2010; Wang *et al.*, 2017).

Previous work on DENV genomic RNA structure, has made use of various *in silico* methods, for example Mfold (Zuker, 2003), and *in vitro* biochemical probing techniques utilising truncated DENV transcripts of both the 5´ and 3´ UTRs or mini-genomes to make RNA structural predictions in the absence of *trans*-activating factors (You and Padmanabhan, 1999; Sztuba-Solinska *et al.*, 2013; Sztuba-Solinska and Le Grice, 2014; de Borba *et al.*, 2015; Villordo *et al.*, 2015; Liu *et al.*, 2016). Published work has not focused on the full-length DENV genome or considered the influence of *trans*-activating factors on RNA structure. Recent RNA structural research has focused on RNA transcripts of DENV serotypes 1, 3 and 4 (Sasmono *et al.*, 2015; Liu *et al.*, 2016). Consequently, this study focuses on the RNA structure of the 5´ extremity of DENV-2 genomic RNA in the context of the full-length genome in the absence and presence of *trans*-activating factors at physiological temperature.

We chose to focus our structural analysis on the first 145 nts of the DENV-2 genome, in the context of the full-length genome and a truncated 5´ 500 nt transcript, for multiple reasons. Firstly, and most crucially, using this nucleotide region one can determine the linear or cyclised genomic conformation of the RNA via the identification of specific structural motifs, as discussed in section 4.2.2. Secondly, a fundamental requirement of DENV replication is the binding of DENV RdRp to SLA, the first stem-loop structure within the 145 nt region chosen for structural analysis (further described in section 1.3.2). Thirdly, the 5´ 500 nt transcript was utilised as a negative control, unable to cyclise and to mimic published DENV SHAPE experiments. Therefore, following the binding of DENV RdRp or other *trans*-activating factors, to genomic RNA one can determine direct downstream structural modifications to *cis*-acting RNA structures within this 145 nt

region and whether or not *trans*-activating factors can induce the linear to circular, or vice versa, genomic conformational switch *in vitro*.

RNA structure, i.e determining whether a nucleotide is base-paired or single-stranded, has been investigated for many years by measuring how individual RNA nucleotides react with enzymatic or chemical probes. For example, digestion of radioactively end-labelled RNA with Ribonucleases (RNases) such as RNAase TI, which cleaves selectively at single-stranded nucleotide positions, followed by detection and analysis using denaturing polyacrylamide gels can be used to determine RNA structure (Nilsen, 2013). Conversely, the RNA of interest can be probed chemically. Ethylnitrosourea alkylates phosphate oxygens that are not involved in tertiary structure interactions. Following alkaline treatment, the phosphotriester hydrolyses resulting in RNA strand cleavage allowing for RNA structure determination analysis following radioactive end-labelling and gel electrophoresis (Ziehler and Engelke, 2001). However, these traditional RNA structure determination techniques have their limitations. For example, the use of denaturing polyacrylamide sequencing gels and radioactive source handling demands a high level of technical expertise. Extensive RNA degradation or strong secondary structure induced terminations may obscure cleavages or modifications at several nucleotides of interest. Due to these limitations, SHAPE mapping was the experimental technique of choice for the analysis of DENV-2 genomic RNA structure.

Mapping RNA structure with SHAPE chemistry is a versatile method for characterising RNA structure. SHAPE reagents, such as N-Methylisatoic anhydride (NMIA) or 2-methylnicotinic acid imidazolide (NAI), act to covalently modify RNA creating adducts at the 2´OH- group of the ribose backbone at nucleotides that are structurally flexible and therefore single-stranded (Fig. 55).

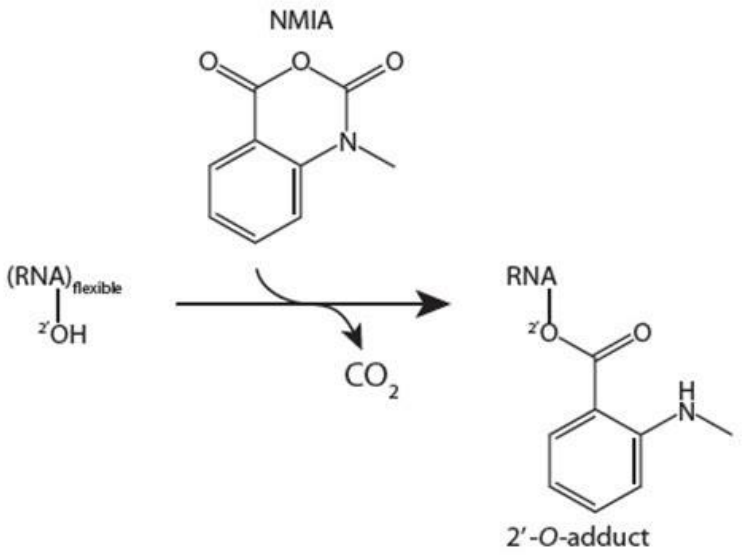



**Figure 55 RNA modification by NMIA:**

*NMIA reacts with the 2´-hydroxyl group in flexible, unpaired nucleotides forming a stable 2´-O-adduct. Because the modifications correlate with RNA structure, the ends of the resulting pool of cDNAs map to the areas of single-stranded nucleotides within the RNA. Taken from* (IDT, 2018)

The positions of these modified nucleotides are detected using reverse transcription, which terminates 3´ of the modified nucleotide. This action results in a pool of cDNAs whose lengths reflect the position, and therefore nucleotide, of the SHAPE modification if analysed alongside a sequencing ladder. Quantification of the cDNA pools is used to estimate the 'reactivity' of each nucleotide in an RNA molecule to the reagent utilised. A nucleotide displaying reactivity over 0.2 has partial reactivity, and high reactivity (over 0.5) indicates that the nucleotide is structurally flexible and single-stranded. Conversely, a nucleotide displaying low reactivity (below 0.2) indicates that the nucleotide is inflexible and therefore base-paired (Figs. 55 and 56).

SHAPE mapping for RNA structure determination has many advantages over traditional enzymatic and chemical probing procedures, the most significant being that long RNA transcripts can be analysed and that every base within the RNA molecule is reacted with, and therefore the structure determined, within one single reaction which can be quantified. The development of high-throughput SHAPE (hSHAPE) resulted in a paradigm shift in SHAPE data analysis and experimental procedure whereby NMIA probed RNA was no longer reverse transcribed using

radiolabelled primers and analysed utilizing technically challenging sequencing denaturing polyacrylamide gels but fluorescently labelled and analysed by capillary electrophoresis. Analysis by capillary electrophoresis allows for the consecutive investigation of up to 96 samples, enabling serial biological repeats and increasing experimental scope for quantification (Fig. 56).

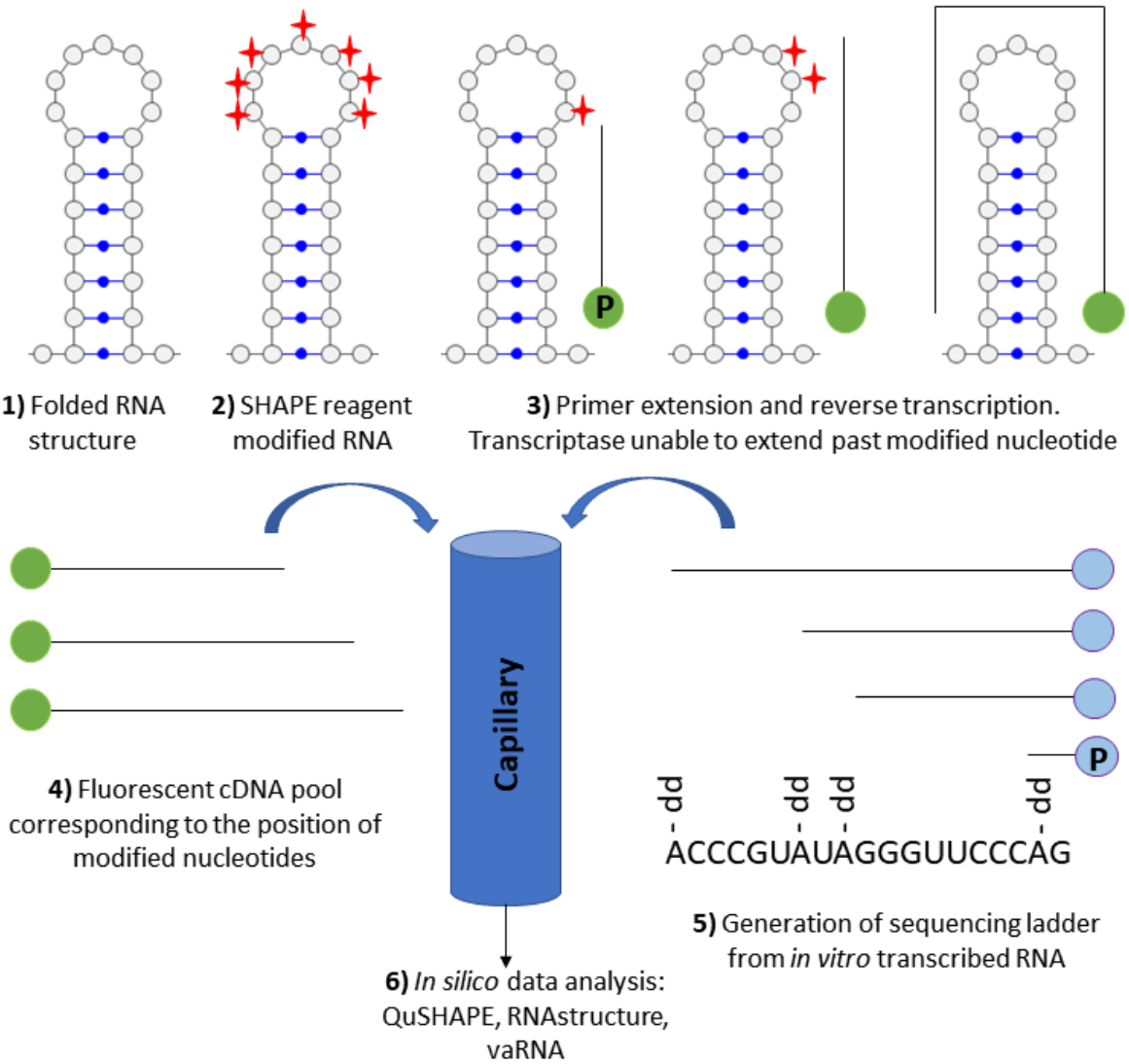



**Figure 56 Schematic representation of NMIA treatment, primer extension, reverse transcription and fluorescent capillary electrophoresis:**

*The SHAPE experimental procedure. NMIA is depicted in red and fluorescently labelled primers are shown in green and purple. RT extensions are represented by black lines.*

Experimentally determined SHAPE reactivities can be used to infer RNA structures by restraining an input RNA sequence with the predicted SHAPE reactivities utilising *in silico* algorithms such as Restrained sample (Rsample). Rsample, an algorithm incorporated into RNAstructure, models RNA secondary structure using thermodynamics guided by RNA secondary structure mapping data. Rsample addresses two common limitations in comparison to non-thermodynamically guided structural predictions. Firstly, Rsample considers that multiple copies of the same sequence can simultaneously fold into different structures. Secondly, it analyses the agreement between experimental mapping data and *in silico* predicted mapping data by sampling RNA structure models. This technique therefore provides a comprehensive approach in the prediction of RNA structure via the integration of thermodynamic prediction with mapping data (Spasic *et al.*, 2018).

The inclusion of thermodynamic data in RNA structure prediction has been shown to be extremely accurate. For example, the use of SHAPE data as a 'soft' pseudo-free energy constraint increased the prediction accuracy of the 16S *E. coli* RNA structure from 72%, as determined by traditional chemical probing, to approximately 95% (Deigan *et al.*, 2009). The inclusion of SHAPE data continuously results in base-pair prediction accuracies of approximately >90% (Deigan *et al.*, 2009; Hajdin *et al.*, 2013; Rice, Leonard and Weeks, 2014). Additionally, simple overlays of SHAPE reactivities on to previously thermodynamically predicted RNA structures provide a visually impactful method of displaying structural data and investigating the dynamics of RNA structural changes. Consequently, the recent advances in SHAPE chemistry and analysis of RNA structure using experimentally determined SHAPE data has culminated in an extremely powerful and relatively high-throughput accurate method to determine RNA structure (Kenyon, Prestwood and Lever, 2014).

Furthermore, SHAPE chemistry is being increasingly utilised to determine the influence of interacting partners, or *trans*-activating factors, introduced into the reaction, following folding and prior to the addition of SHAPE reagent. This utilisation of the SHAPE methodology is depicted in figure 57.

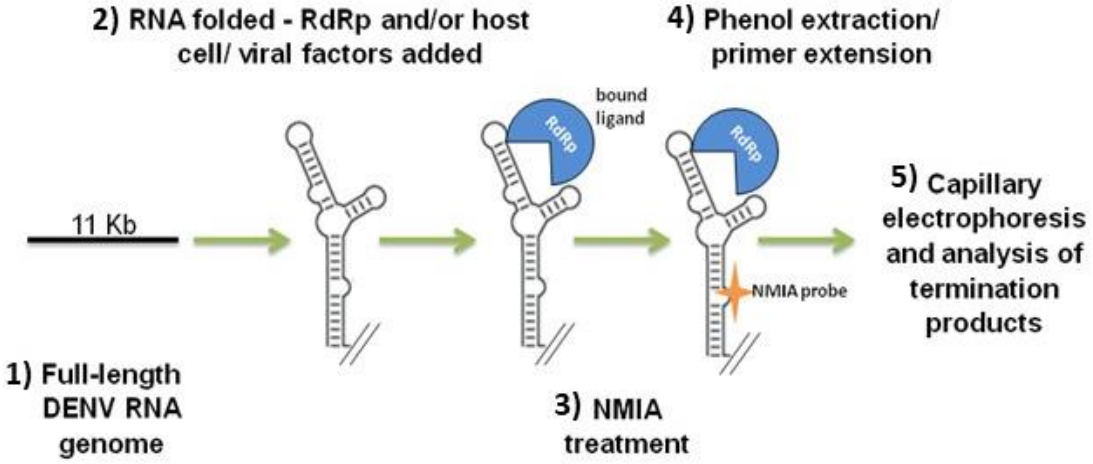



**Figure 57 A schematic diagram of the *in vitro* SHAPE procedure in the presence of *trans*-activating factors:**

***1)*** *The full-length or 5´500 nt DENV genome is in vitro transcribed.* ***2)*** *The RNA is folded and trans-activating factors added.* ***3)*** *Folded RNA is probed with NMIA, modifying the RNA at single-stranded nucleotides.* ***4)*** *The position of NMIA modified nucleotides are detected following reverse transcription, creating a cDNA pool reflecting the position of the modified nucleotides.* ***5)*** *Fluorescent capillary electrophoresis is used to quantify and align the cDNA pool, in comparison to a sequencing ladder, and in silico analysis using QuSHAPE, RNAstructure and VARNA is used to accurately determine RNA structure* (Reuter and Mathews, 2010; Karabiber *et al.*, 2013).

This method was utilised to predict eIF3 interactions with HCV SLIIIb and similarly the 40S ribosome with SLIIId (Sun *et al.*, 2013; Angulo *et al.*, 2016). However, this experimental study requires extra experimental optimisation, demanding careful consideration of RNase contamination - as a constitutive element of cellular extracts or purified protein preparations. However, following consideration, and in the case of purified protein - determination of RNA binding characteristics, this utilisation of SHAPE experimentation provided a novel avenue for this project. Exploring DENV-2 genomic RNA structure under the influence of *trans*-activating factors, such as DENV-2 RdRp and cellular proteins extracted from stably expressing DENV HEK293T replicon cells, at physiological temperature considers the impact of the physiological environment on DENV genomic RNA structure and viral replication *in vitro*.

The aims of our DENV-2 *in vitro* SHAPE experimentation are detailed below:

1. Determine the genomic conformation, either linear or circular, of full-length DENV-2 RNA in the absence/presence of *trans*-activating factors, in comparison to a truncated DENV genomic transcript at physiological temperature
2. Determine the influence of *trans*-activating factors on *cis*-acting replication elements in the 5´ extremity of the DENV genome at physiological temperature

This chapter begins by introducing the *in vitro* SHAPE experimental technique and the structural motifs used for diagnosis of the linear or circular genomic conformation. The experimental findings of *in vitro* SHAPE are then described utilizing both the truncated 5´ 500 nt transcript and full-length DENV genomic RNA in the absence and presence of *trans*-activating factors and at both mammalian and insect physiological temperature. This study therefore presents a novel take on the *in vitro* SHAPE experimental technique, considering the structural influence of *trans*-activating factors on DENV-2 genomic RNA in the context of viral replication.

## 4.2 Results

### 4.2.1 *In vitro* SHAPE mapping in the absence/presence of *trans*-activating factors

SHAPE fragment analysis was conducted using capillary electrophoresis by DNA Sequencing and Services, University of Dundee. Data analysis was subsequently carried out using QuSHAPE (Karabiber *et al.*, 2013) and reactivities overlaid onto the Mfold predicted structure of the DENV-2 5´extremity RNA (1 – 145 nts) at the appropriate physiological temperature. The QuSHAPE data analysis pipeline is described in figure 58. Example thermodynamically SHAPE data restrained DENV-2 RNA structures can be viewed in Appendix II.

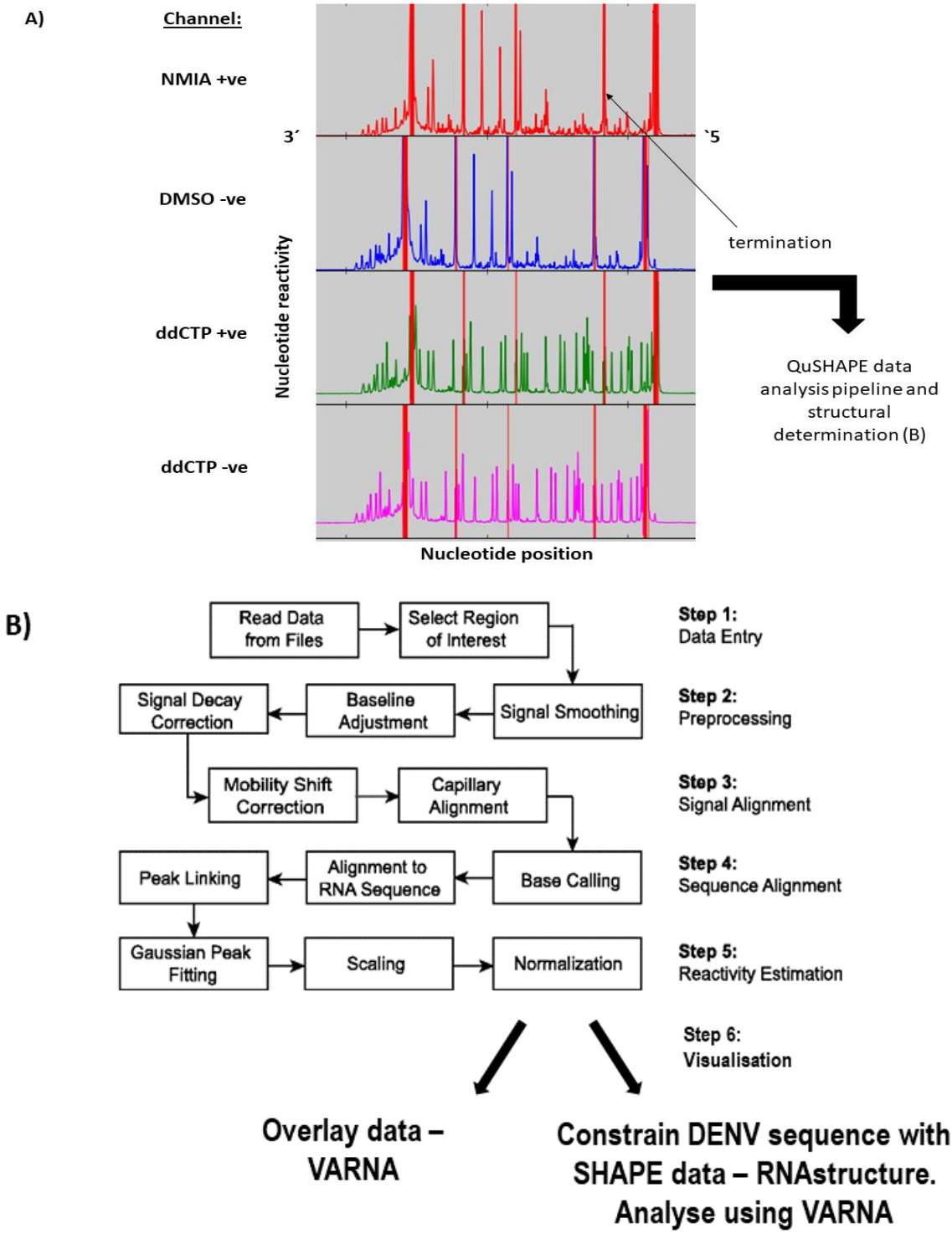



**Figure 58 QuSHAPE data analysis pipeline and RNA structural prediction:**

***A)*** *Example QuSHAPE NMIA reactivity profile in comparison to a sequencing ladder.* ***B)*** *The QuSHAPE analysis pipeline and analysis of RNA structure using RNAstructure and the visualisation applet VARNA* (Reuter and Mathews, 2010; Karabiber *et al.*, 2013).

High quantity, high quality 5´500 nt and full-length DENV-2 genomic RNA was used to reproducibly obtain high quality SHAPE reactivity profiles spanning the first 145 nucleotides of the DENV-2 genome (Figs.  59 and 60).

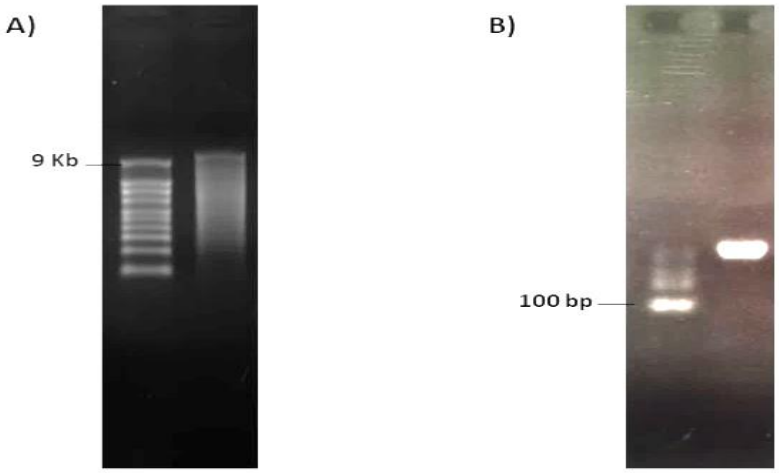



**Figure 59 Analysis of RNA quality as templates for *in vitro* SHAPE mapping:**

*Quality of in vitro transcribed RNA used for in vitro SHAPE mapping procedures.* ***A)*** *the full-length DENV-2 transcript in comparison to the Ambion Millenium RNA ladder.* ***B)*** *5´ 500 nt DENV-2 transcript in comparison to the Ambion Century RNA size ladder.*

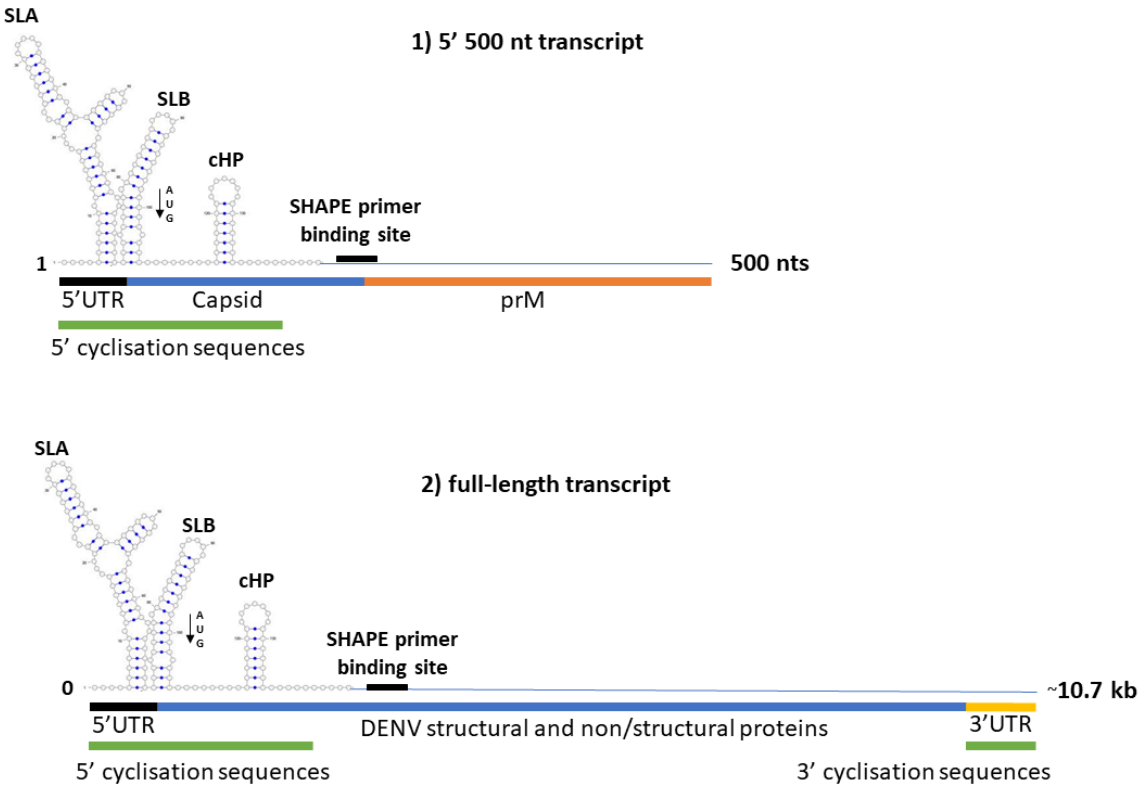



**Figure 60 Schematic of 5´500 nt and full-length DENV genomic transcripts used throughout *in vitro* SHAPE experimentation:**

*A schematic diagram of in vitro transcribed RNA transcripts used throughout in vitro SHAPE experimentation.* ***1)*** *The 5´ 500 nt transcript was used to mimic published SHAPE experimentation using DENV-2 RNA and also as a control RNA, unable to cyclise due to the lack of complementary 3´ cyclisation sequences.* ***2)*** *The full-length DENV-2 genomic RNA transcript.*

### 4.2.2 Identification of circular or linear structural genomic form

Upon analysing SHAPE data, it was vital to be able to identify the linear or cyclised structural form of the DENV genome. We utilised structural motifs at the 5´ region of DENV RNA for this. In the linear genomic form SLB is present, as indicated by a top loop reactive to SHAPE reagent and a base paired stem. Additionally, the 5´CS is un base-paired and therefore SHAPE reagent reactive. In the cyclised form, the SLB structure is unfolded due to base-pairing of the 5´-3´ UAR and the 5´CS and DAR become base-paired upon binding to the 3´CS and DAR during genomic cyclisation. The structural motifs and predicted rearrangements are displayed in figure 61 and utilised throughout SHAPE experimentation to identify genomic conformation.

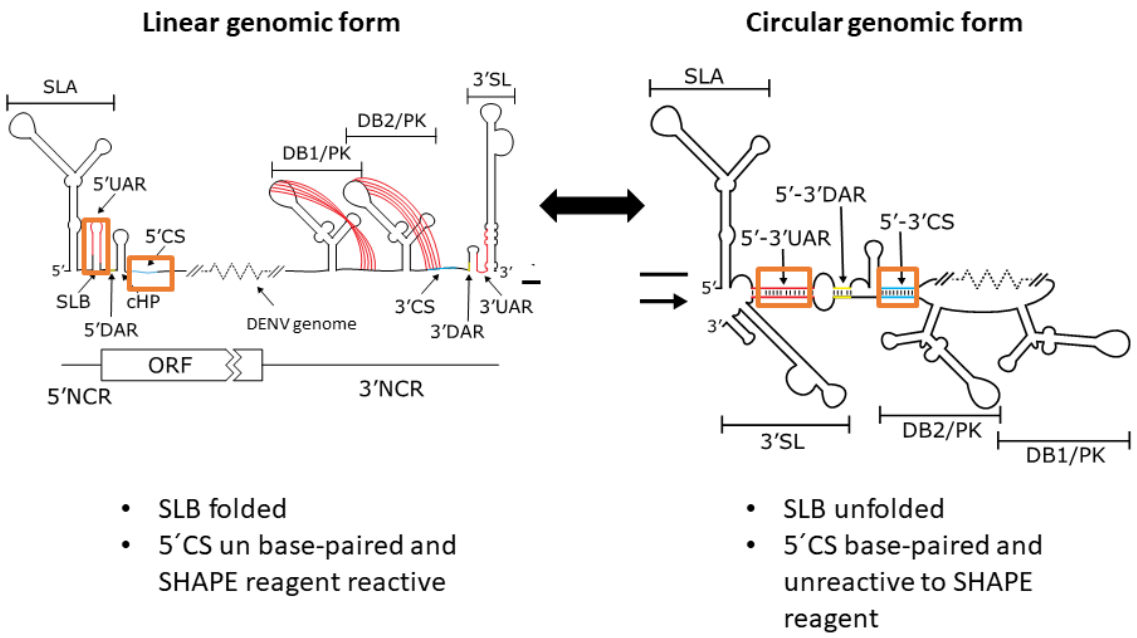



**Figure 61 Diagnosis of the linear of circular genomic form for data analysis:**

*Schematic of structural motifs used in to identify the circular or linear DENV genome conformation throughout SHAPE experimentation. Linear genomic conformation: SLB structure is present, 5´CS is un base-paired. Circular genomic conformation: SLB structure is unfolded due to the base pairing of the 5´-`3 UAR interaction. The 5´CS hybridises to the 3´CS and is therefore unreactive to NMIA/NAI. The structural motifs used for genomic conformation identification are boxed in orange. DB = Dumbbell, PK= Pseudoknot, UAR= Upstream of AUG region, DAR = Downstream of AUG region and CS= Cyclisation sequence. Adapted from* (Tuplin, 2015).

### 4.2.3 *In vitro* SHAPE mapping of DENV-2 RNA in the absence of *trans*-activating factors at 37 °C

To date and to our knowledge, published DENV RNA structure analysis using SHAPE mapping data has utilised short *in vitro* transcribed RNAs of approximately 500 nts in length, generated from mini-genome DENV templates, or as a combination of the 5´ and 3´UTR of DENV genomic RNA (You and Padmanabhan, 1999). Published DENV full-length SHAPE data has focussed on the influence of the 3´UTR (Villordo *et al.*, 2015). Therefore, DENV-2 genomic RNA has been understudied in this manner. Prior to the analysis of the structure of the DENV-2 5´ region RNA in the context of the full-length genome we mimicked published experiments, mapping the structure of the 5´ extremity RNA in the context of a truncated 5´500 nt transcript. We compared this data with SHAPE data obtained following experimentation with the full-length DENV-2 genome.

The normalised nucleotide reactivities to NMIA are compared, in the context of the 5´500 nt truncated transcript and the full-length DENV-2 genome, before being overlaid onto the predicted structure of the DENV-2 5´extremity RNA at 37°C in figure 62.

As expected, the 5´500 nt transcript was in the linear conformation as 3´ UTR cyclisation sequences were unavailable. This was displayed by NMIA reactivities above 0.2 corresponding with the top-loop of SLB, a base-paired stem and the 5´CS predicting these regions to be single-stranded. Analysis of NMIA reactivities produced utilising the full-length DENV-2 transcript displayed that this RNA was also in the linear conformation as NMIA reactivities were above 0.2 in regions corresponding to the top-loop of SLB, a base paired stem and the 5´CS indicating these regions to be single-stranded and therefore in the linear conformation. The circular genomic form is predicted to be the most thermodynamically stable form of the DENV genome however, *in vitro* in the absence of *trans*-activating factors the full-length DENV-2 genome forms the linear conformation.

The stabilisation of the UFS, at the base of SLB, is suggested to be fundamental for the binding of DENV-2 RdRp in the linear context of the genome and in controlling the switch between the linear and cyclic genomic form. In the absence

of *trans*-activating factors, utilising both transcripts, the UFS is unstable, as determined by the high accessibility to NMIA in this region, indicating this region to be predominantly single-stranded and therefore the absence of the UFS structure (Fig. 62).

The structure of cHP is suggested to be important in the regulation of viral replication. In the context of the 5´500 nt transcript, in the absence of *trans*-activating factors the structure ceases to form. This is indicated by a lack of NMIA reactivity displayed by nucleotide bases corresponding to the top-loop of the structure being pre-dominantly base-paired. Additionally, the base of cHP ceases to form indicated by the nucleotides corresponding to the side of the base of the stem displaying a high-reactivity and are therefore predominantly single-stranded. This observation is also echoed in the context of the full-length genome whereby the structure ceases to form as evidenced by high NMIA nucleotide reactivities corresponding to the side of the stem-loop base indicating this region to be predominantly single-stranded (Fig. 62)

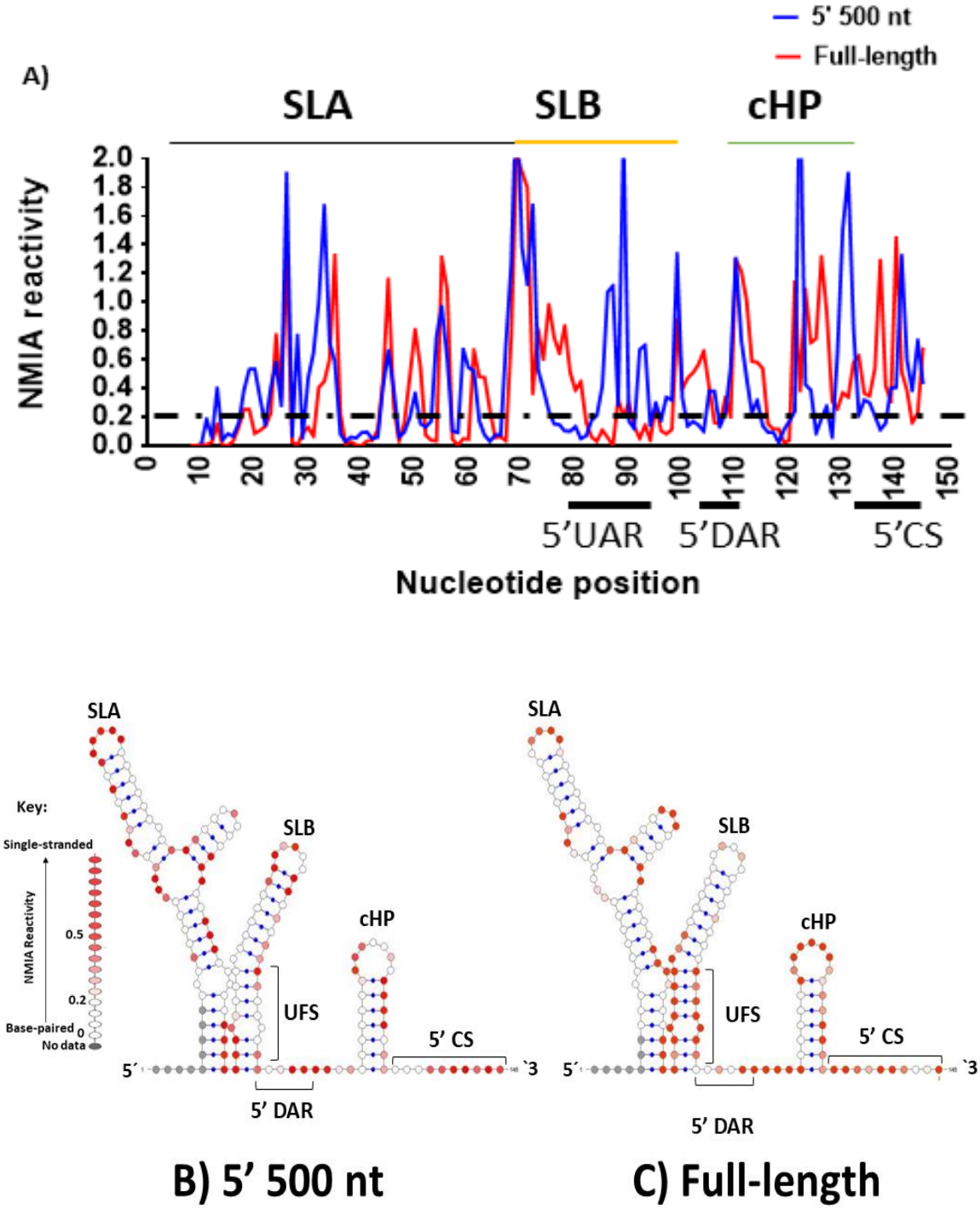

A)
5' 500 nt
Full-length
SLA
SLB
cHP
NMIA reactivity
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
5'UAR
5'DAR
5'CS
Nucleotide position
SLA
SLB
cHP
Key:
Single-stranded
NMIA Reactivity
0.5
0.2
Base-paired
0
No data
UFS
5' CS
5'
`3
5' DAR
B) 5' 500 nt
SLA
SLB
cHP
UFS
5' CS
5'
`3
5' DAR
C) Full-length

**Figure 62 *In vitro* SHAPE mapping of DENV-2 5´500 nt and full-length genomic RNA transcripts:**

**A)** *NMIA reactivities corresponding to the 5´ 500 nt and full-length transcript are shown in blue and red respectively. The dashed line corresponds to a NMIA reactivity of 0.2. Reactivities above this line are partially reactive and above 0.5 are highly reactive to NMIA and therefore single-stranded. Nucleotide positions of RNA structures and cyclisation sequences are indicated.* ***B)*** *5´500 nts NMIA reactivities overlaid onto the first 145 nts of DENV-2 genomic RNA.* ***C)*** *Full-length DENV-2 NMIA reactivities overlaid onto the first 145 nts of DENV-2 genomic RNA. Overlays of NMIA reactivities conducted using VARNA.*

### 4.2.4 *In vitro* SHAPE mapping of DENV-2 RNA in the presence of DENV-2 RdRp at 37 °C

Both DENV-2 5´500 nt and full-length RNA transcripts form the linear conformation, displaying an unstable UFS, *in vitro* in the absence of *trans*-activating factors. The next logical step was to consider the physiological environment of DENV infection *in vitro*, analysing the influence of the DENV-2 RdRp on DENV genomic RNA structure, with the aim of investigating what controls the switch between the linear and cyclized genomic forms. It was hypothesised that DENV-2 RdRp may cause the switch from the linear to cyclised form since it was suggested that the DENV UFS is required for the binding of DENV-2 RdRp to SLA, yet the cyclised genomic form is required to transfer the RdRp to the 3´ SL and therefore replicate the viral genome. Consequently, the binding of RdRp to SLA was hypothesised to induce the genomic structural change.

The normalised nucleotide reactivities to NMIA are compared, in the context of both the 5´500 nt and full-length genome transcript in the absence and presence of DENV-2 RdRp, before being overlaid onto the predicted structure of the DENV-2 5´ extremity RNA at 37°C in figures 63 and 67.

In the context of the 5´ 500 nt transcript a number of structural rearrangements occur upon the addition of DENV RdRp. The nucleotide region corresponding to the UFS becomes yet more unstable, as indicated by the increased nucleotide reactivities in this region in comparison to the SHAPE mapping data in the absence of *trans*-activating factors. The structure of cHP also ceases to form adequately, as indicated by the increased nucleotide reactivities to NMIA in regions corresponding to the stem of cHP and to the top-loop of cHP. Therefore, significant structural rearrangements occur in the context of the 5´500 nt DENV-2 RNA transcript upon the addition of DENV-2 RdRp (Fig. 63).

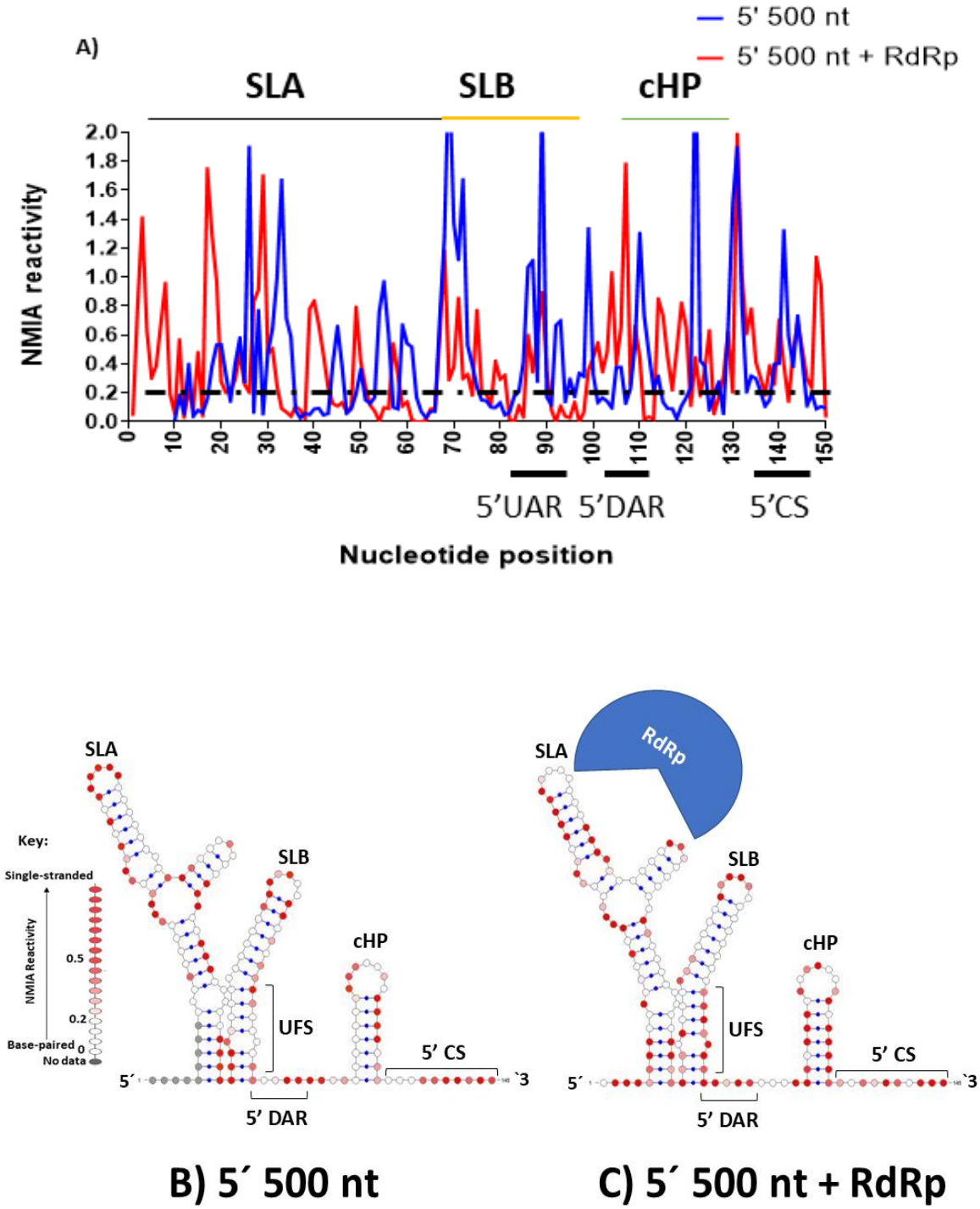
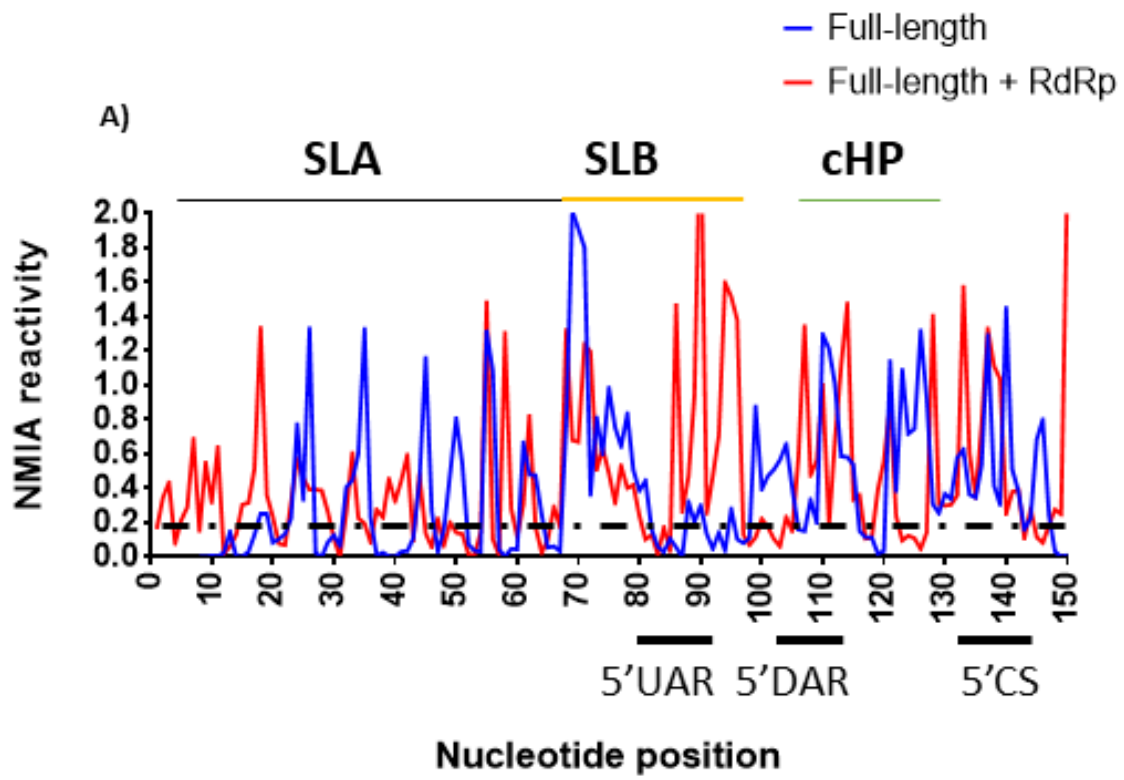

A)
5' 500 nt
5' 500 nt + RdRp
SLA
SLB
cHP
NMIA reactivity
5'UAR
5'DAR
5'CS
Nucleotide position
SLA
SLB
cHP
UFS
5' CS
5' DAR
Key:
Single-stranded
NMIA Reactivity
Base-paired
No data
RdRp
B) 5´ 500 nt
C) 5´ 500 nt + RdRp

**Figure 63 *In vitro* SHAPE mapping of DENV-2 5´500 nt transcript in the absence and presence of DENV-2 RdRp:**

**A)** *NMIA reactivities corresponding to the 5´ 500 nt in the absence and presence of DENV-2 RdRp in blue and red respectively. The dashed line corresponds to a NMIA reactivity of 0.2. Reactivities above this line are partially reactive and above 0.5 are highly reactive to NMIA and therefore single-stranded. Nucleotide positions of RNA structures and cyclisation sequences are indicated.* ***B)*** *5´500 nts NMIA reactivities overlaid onto the first 145 nts of DENV-2 genomic RNA.* ***C)*** *5´ 500 nt in the presence of DENV-2 RdRp NMIA reactivities overlaid onto the first 145 nts of DENV-2 genomic RNA. Overlays of NMIA reactivities conducted using VARNA.*

As previously discussed, in the context of the DENV full-length genome in the absence of *trans*-activating factors the genome forms the linear conformation, displaying a NMIA reactive top-loop of SLB and a reactive 5´DAR and 5´CS. Furthermore, the UFS is inherently unstable in this context. Upon the addition of DENV-2 RdRp the genome remains in the linear form, indicated by a NMIA reactive top-loop of SLB and reactive cyclisation sequences. The UFS does not form and the SLB structure melts as the stem becomes more reactive to NMIA. In addition, cHP also ceases to form as indicated by the loss of reactivity in regions corresponding the top-loop of the structure and the increased reactivity in regions corresponding to the stem/base of the structure (Fig. 64).


Full-length
Full-length + RdRp
A)
SLA
SLB
cHP
NMIA reactivity
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
5'UAR
5'DAR
5'CS
Nucleotide position

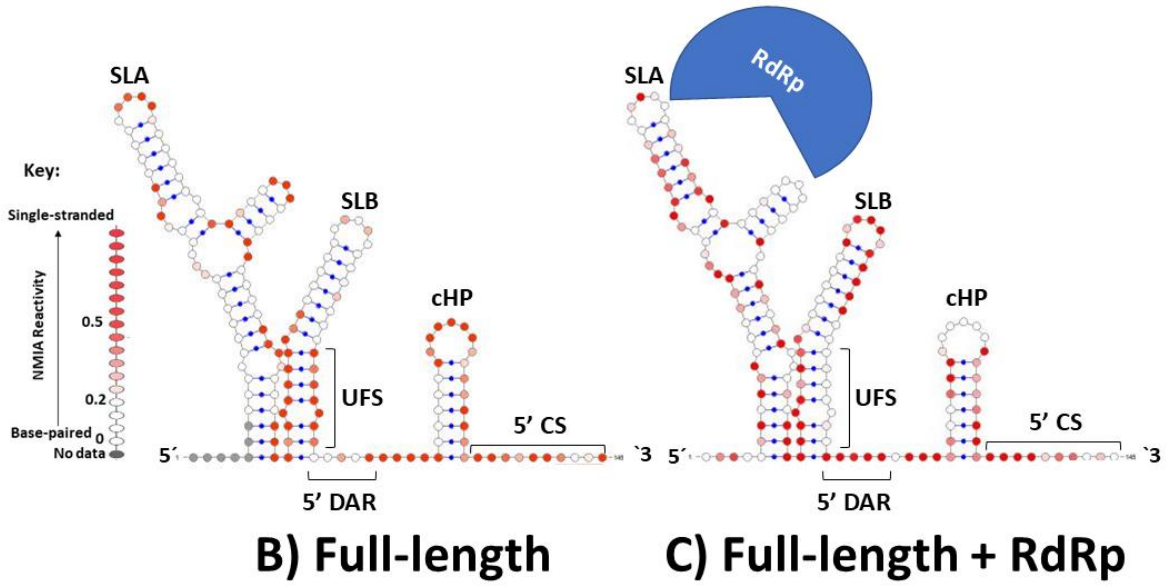

SLA
Key:
Single-stranded
NMIA Reactivity
0.5
0.2
Base-paired
0
No data
SLB
cHP
UFS
5' CS
5'
`3
5' DAR
B) Full-length
RdRp
SLA
SLB
cHP
UFS
5' CS
5'
`3
5' DAR
C) Full-length + RdRp

**Figure 64 *In vitro* SHAPE mapping of full-length DENV-2 genomic RNA in the absence and presence of DENV-2 RdRp:**

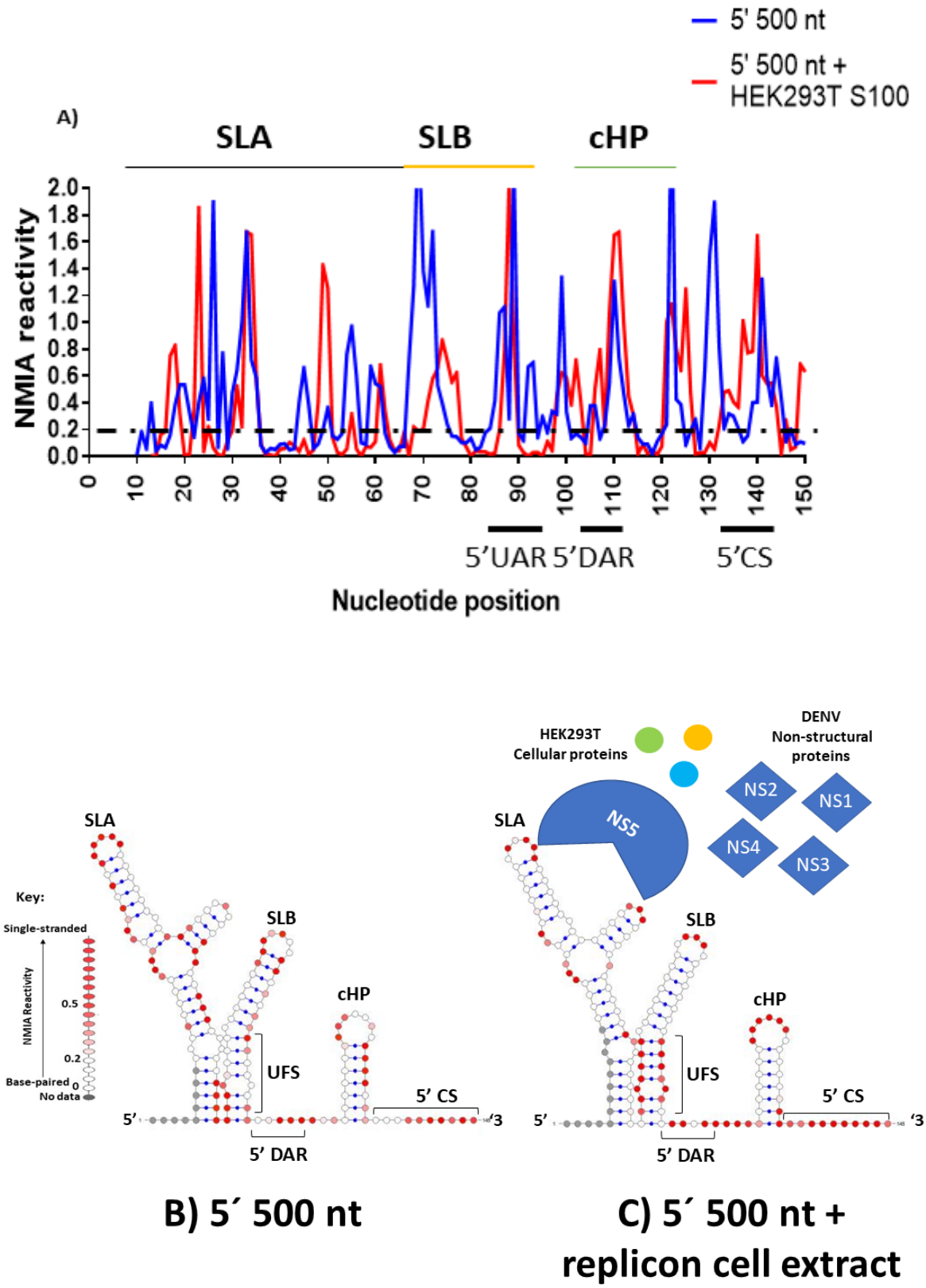
**A)** *NMIA reactivities corresponding to full-length DENV-2 genomic RNA in the absence and presence of DENV-2 RdRp in blue and red respectively. The dashed line corresponds to a NMIA reactivity of 0.2. Reactivities above this line are partially reactive and above 0.5 are highly reactive to NMIA and therefore single-stranded. Nucleotide positions of RNA structures and cyclisation sequences are indicated.* ***B)*** *Full-length DENV-2 genomic RNA NMIA reactivities overlaid onto the first 145 nts of DENV-2 genomic RNA.* ***C)*** *Full-length DENV-2 genomic RNA in the presence of DENV-2 RdRp NMIA reactivities overlaid onto the first 145 nts of DENV-2 genomic RNA. Overlays of NMIA reactivities conducted using VARNA.*

### 4.2.5 *In vitro* SHAPE mapping of DENV-2 RNA in the presence of HEK293T replicon extract at 37 °C

Following *in vitro* SHAPE mapping experiments in the presence of the *trans*-activating factor DENV-2 RdRp, the next logical step was to consider the structure of DENV RNA in the context of the cellular environment under a mimicked virally infected state. We extracted proteins from replication competent HEK293T cells, stably expressing DENV-2 non-structural proteins, and utilised them in *in vitro* SHAPE experimentation (for DENV replicon map and confirmation of the stable expression of DENV non-structural proteins refer to section 5.2.2). This aimed to further deduce the biological mechanisms behind the switch from the linear to circular genomic form of the genome, an act fundamental to viral replication, in a more physiologically relevant context.

The normalized nucleotide reactivities to NMIA are compared, in the context of both the 5´500 nt and full-length genomic transcript in the absence and presence of stably expressing DENV HEK293T replicon cellular extract, before being overlaid onto the predicted structure of the DENV-2 5´ extremity RNA at 37°C (figures 65 and 66).

In the context of the 5´ 500 nt truncated transcript, upon the addition of HEK293T replicon cellular extract the linear genomic form is maintained, as displayed by the reactivity of the top-loop of SLB to NMIA a base-paired stem, and a single-stranded 5´ CS. The UFS is further de-stabilized in this context as displayed by the increased NMIA reactivities in this region. Upon the addition of stably expressing DENV HEK293T replicon cellular extract the cHP structure is stabilised as indicated by the loss of NMIA reactivity corresponding to bases comprising the stem of cHP and an increase in reactivity corresponding to bases comprising the top-loop of cHP. Therefore, the addition of DENV HEK293T replicon cellular extract causes structural rearrangements in the context of the 5´500 nt DENV RNA transcript (Fig. 65).



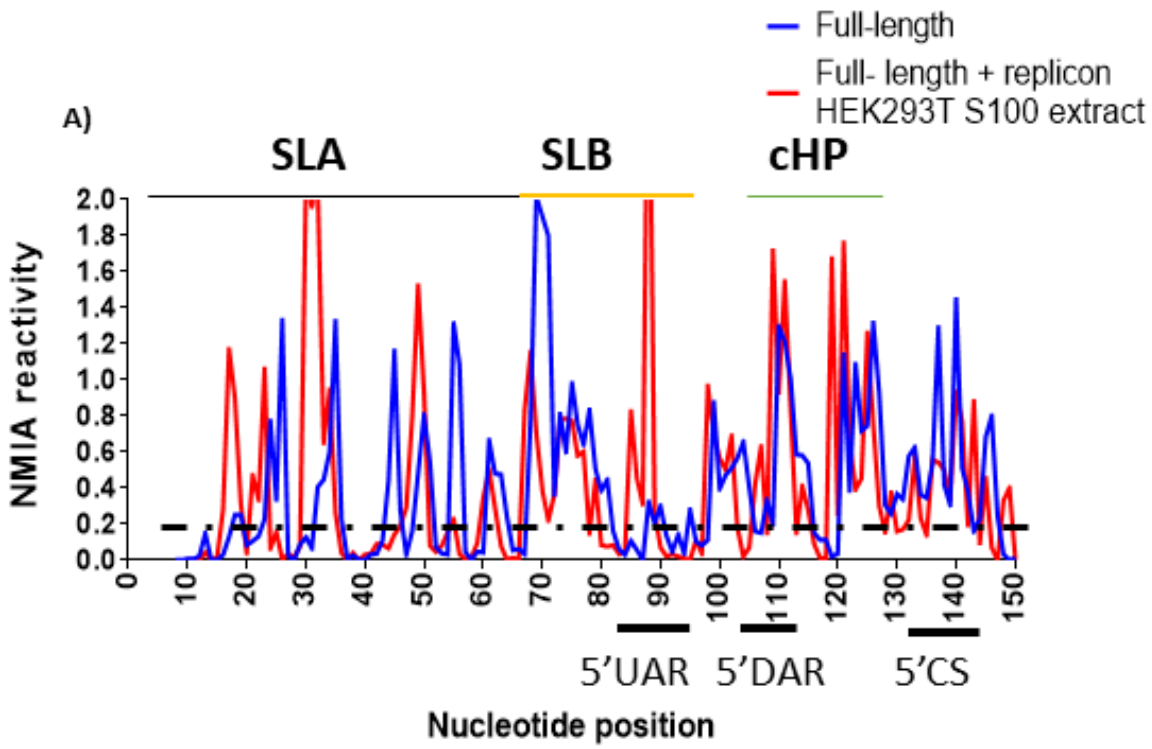

B) 5´ 500 nt

C) 5´ 500 nt + replicon cell extract

**Figure 65 *In vitro* SHAPE mapping of 5´500 nt DENV-2 genomic RNA in the absence and presence of stably expressing DENV HEK293T replicon cellular protein extract:**

**A)** *NMIA reactivities corresponding to 5´ 500 nt DENV-2 genomic RNA in the absence and presence of* HEK293T DENV replicon S100 protein extract *in blue and red respectively. The dashed line corresponds to a NMIA reactivity of 0.2. Reactivities above this line are partially reactive and above 0.5 are highly reactive to NMIA and therefore single-stranded. Nucleotide positions of RNA structures and cyclisation sequences are indicated.* ***B)*** *5´ 500 nt DENV-2 genomic RNA NMIA reactivities overlaid onto the first 145 nts of DENV-2 genomic RNA.* ***C)*** *5´ 500 nt DENV-2 genomic RNA in the presence of* HEK293T DENV replicon S100 protein extract *NMIA reactivities overlaid onto the first 145 nts of DENV-2 genomic RNA. Overlays of NMIA reactivities conducted using VARNA.*

Furthering our studies, we SHAPE mapped the full-length DENV genome *in vitro* in the presence of stably expressing DENV HEK293T replicon cell extract. As previously described, we have deduced that the full-length DENV-2 genome favours the linear conformation in the absence of *trans*-activating factors *in vitro*. Upon the addition of stably expressing DENV HEK293T replicon cellular extract the linear form of the genome is maintained, as indicated by nucleotides corresponding to the top-loop of SLB being consistently NMIA reactive and the reactivity of bases corresponding to the 5´CS indicating this region to be single-stranded. The UFS is unstable in this context as displayed by the increased NMIA reactivity in the nucleotide region corresponding to the base of SLB. The cHP structure became stabilized in this context, as indicated by the reduced reactivity to NMIA corresponding to bases in the region of the base of cHP indicating this region to be base-paired and the increased reactivity of bases corresponding to the top-loop of cHP inferring this region to be predominantly single-stranded (Fig. 66).


Full-length
Full- length + replicon HEK293T S100 extract
A)
SLA
SLB
cHP
NMIA reactivity
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
5'UAR
5'DAR
5'CS
Nucleotide position

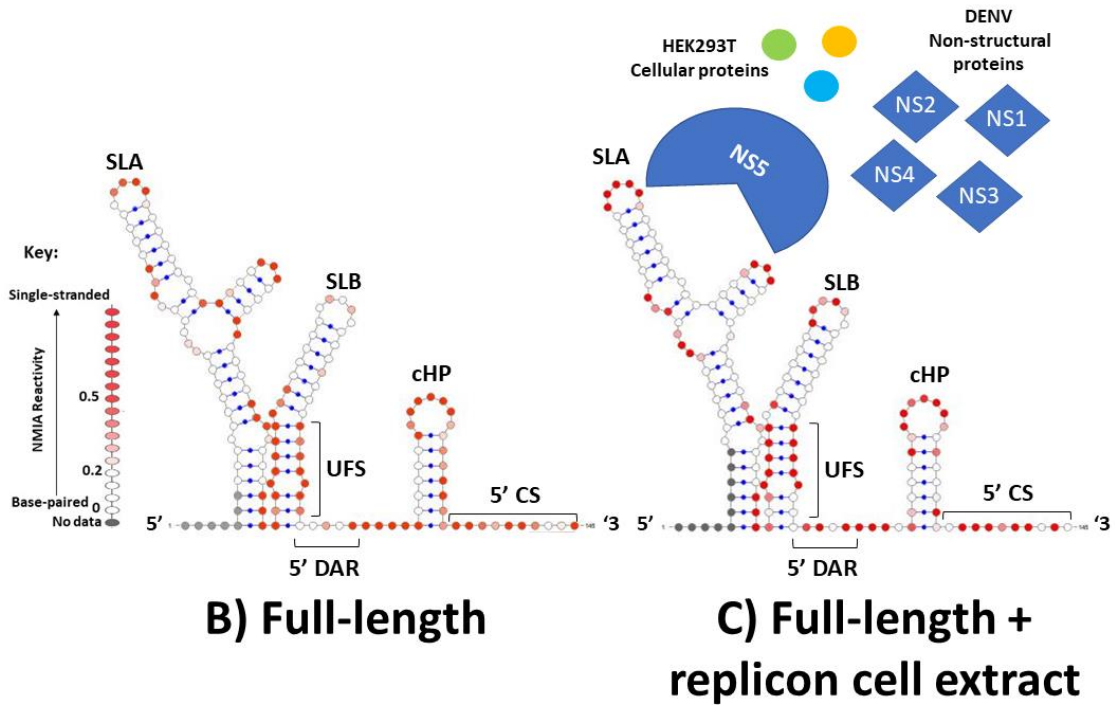

Key:
Single-stranded
NMIA Reactivity
0.5
0.2
Base-paired 0
No data
SLA
SLB
cHP
UFS
5' CS
5'
'3
5' DAR
B) Full-length
HEK293T Cellular proteins
DENV Non-structural proteins
NS5
NS2
NS1
NS4
NS3
SLA
SLB
cHP
UFS
5' CS
5'
'3
5' DAR
C) Full-length + replicon cell extract

**Figure 66 *In vitro* SHAPE mapping of full-length DENV-2 genomic RNA in the absence and presence of HEK293T DENV replicon S100 protein extract:**

**A)** *NMIA reactivities corresponding to full-length DENV-2 genomic RNA in the absence and presence of* HEK293T DENV replicon S100 protein extract *in blue and red respectively. The dashed line corresponds to a NMIA reactivity of 0.2. Reactivities above this line are partially reactive and above 0.5 are highly reactive to NMIA and therefore single-stranded. Nucleotide positions of RNA structures and cyclisation sequences are indicated.* ***B)*** *Full-length DENV-2 genomic RNA NMIA reactivities overlaid onto the first 145 nts of DENV-2 genomic RNA.* ***C)*** *Full-length DENV-2 genomic RNA in the presence of* HEK293T DENV replicon S100 protein extract *NMIA reactivities overlaid onto the first 145 nts of DENV-2 genomic RNA. Overlays of NMIA reactivities conducted using VARNA.*

### 4.2.6 *In vitro* SHAPE mapping of DENV-2 RNA in the absence and presence of DENV-2 RdRp at 28 °C

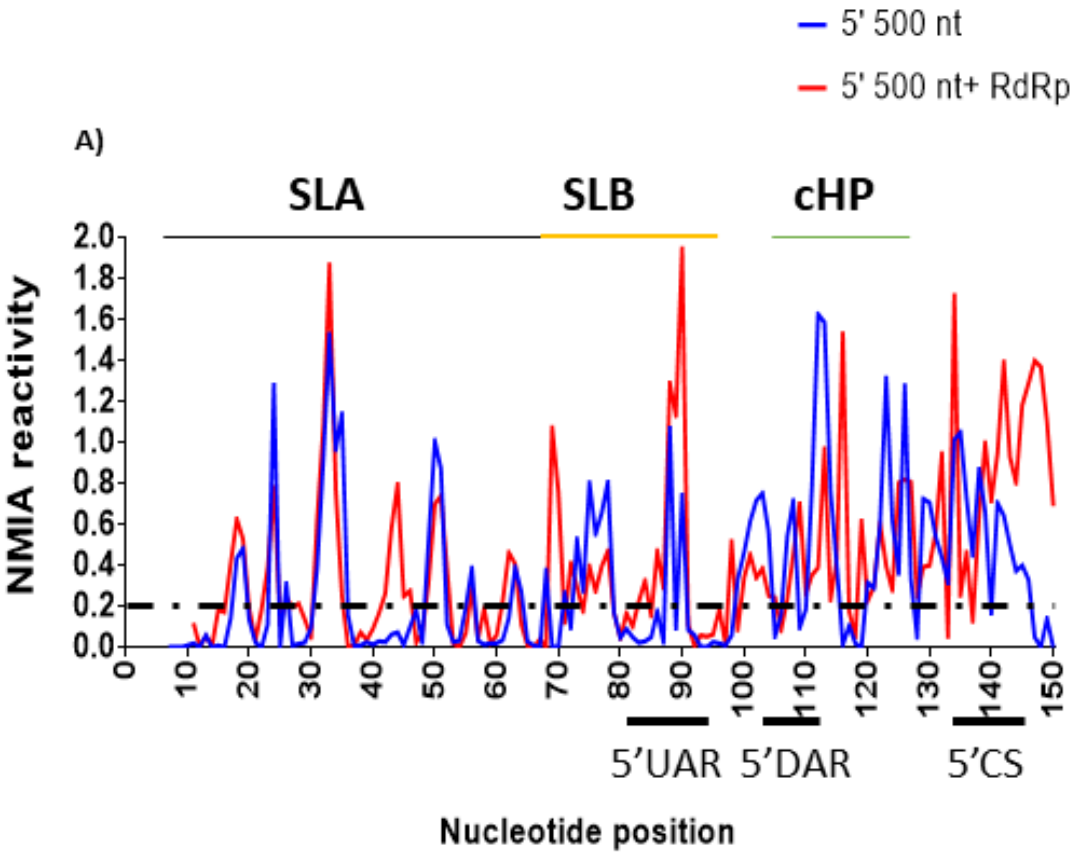
DENV cycles between mammalian and insect hosts and therefore must structurally adapt to the constraints of each environment to establish a robust infection. In contrast to mammalians, DENV infection of mosquito cells leads to a persistent infection without apparent pathological effect. Additionally, previous work has identified differential RNA sequence requirements for DENV replication between mosquito and human cells (Villordo and Gamarnik, 2013). However, the mechanistic basis of this host adaptation is not understood at the RNA structural level. We therefore sought to analyse the structure of the DENV-2 5´ extremity RNA at 28°C (insect physiological temperature) in the absence and presence of DENV-2 RdRp to further deduce the biological mechanisms of the switch between the linear and circular forms of the genome during replication, in an environment where a persistent infection is established.

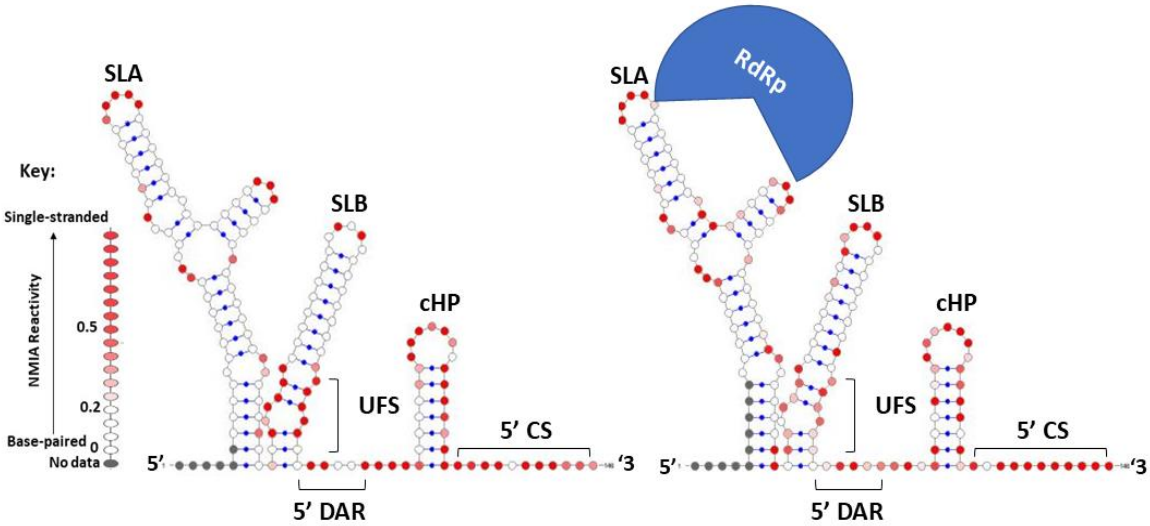
The normalized nucleotide reactivities to NMIA are compared, in the context of both the 5´500 nt and full-length genome transcript in the absence and presence of DENV-2 RdRp, before being overlaid onto the predicted structure of the DENV-2 5´ extremity RNA at 28°C (Figs 67 and 68).

As depicted in figure 67, the 5´500 nt DENV transcript used throughout this study as a DENV genomic RNA form unable to cyclise is, as expected linear, in the absence of DENV-2 RdRp at 28°C as determined by the increased reactivity to NMIA in regions corresponding to the top-loop of SLB and to nucleotides comprising the 5´CS indicating this region to be single-stranded. The UFS at the base of SLB is unstable in this context, with reactivities at the base of SLB indicating this region to be single-stranded. The cHP structure is destabilized as shown by the nucleotides corresponding to the base of cHP displaying a high reactivity to NMIA, inferring this region to be single-stranded.

Upon the addition of DENV-2 RdRp, in the context of the 5´ 500 nt transcript, the genomic conformation of the RNA remains the same, displaying the described diagnostic signs of the linear molecule. However, following the binding of DENV RdRp the UFS at the base of SLB remains unstable. The cHP structure remains

destabilized, as displayed by nucleotides corresponding to the base of cHP displaying a high reactivity profile to NMIA (Fig. 67).



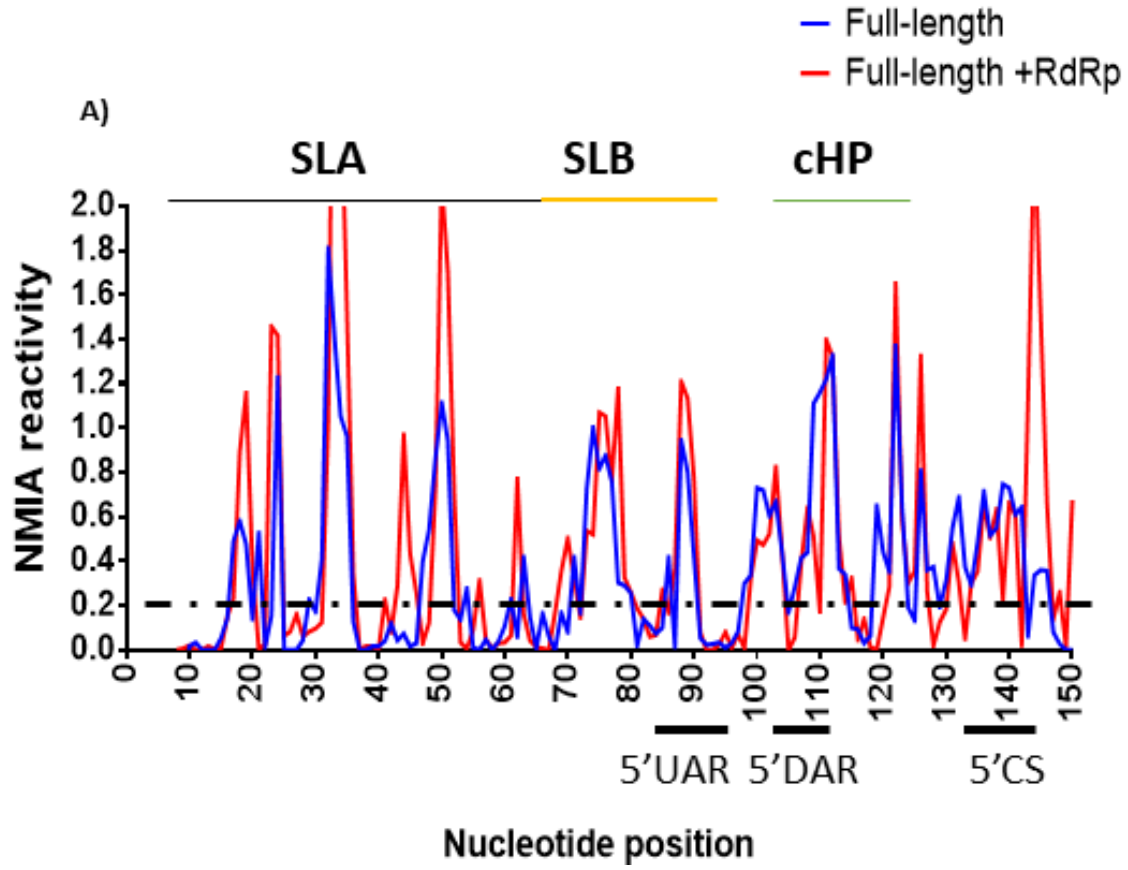




**B) 5´ 500 nt** **C) 5´ 500 nt + RdRp**

**Figure *67 In vitro SHAPE mapping* of 5´ 500 nt DENV-2 genomic RNA in the absence and presence of DENV-2 RdRp at insect physiological temperature:**

**A)** *NMIA reactivities corresponding to 5´ 500 nt DENV-2 genomic RNA in the absence and presence of DENV-2 RdRp in blue and red respectively. The dashed line corresponds to a NMIA reactivity of 0.2. Reactivities above this line are partially reactive and above 0.5 are highly reactive to NMIA and therefore single-stranded. Nucleotide positions of RNA structures and cyclisation sequences are indicated.* ***B)*** *5´ 500 nt DENV-2 genomic RNA NMIA reactivities overlaid onto the first 145 nts of DENV-2 genomic RNA.* ***C)*** *5´ 500 nt DENV-2 genomic RNA in the presence of DENV-2 RdRp NMIA reactivities overlaid onto the first 145 nts of DENV-2 genomic RNA. Overlays of NMIA reactivities conducted using VARNA.*

Furthering our studies, we SHAPE mapped the full-length DENV RNA transcript at insect physiological temperature in the absence and presence of DENV-2 RdRp. The full-length transcript in the absence of DENV-2 RdRp displayed the linear genomic form with NMIA reactivities present corresponding to the top-loop of SLB and to the 5´CS, indicating these regions to be single-stranded. Additionally, in this context the UFS at the base of SLB is reactive to NMIA indicating this region to be predominantly single-stranded. However, upon the addition of DENV-2 RdRp the linear genomic form is maintained, as indicated by the same characteristic structural motifs. Additionally, the UFS at the base of SLB remains de-stabilized. However, in this context the cHP stabilizes. This is indicated by the reduction in nucleotide reactivities corresponding to nucleotides comprising the stem/base of cHP. The top-loop of cHP is reactive to NMIA and is therefore single-stranded, indicating the restoration of the structure of this stem-loop (Fig. 68).


Full-length
Full-length +RdRp
A)
SLA
SLB
cHP
NMIA reactivity
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
5'UAR
5'DAR
5'CS
Nucleotide position

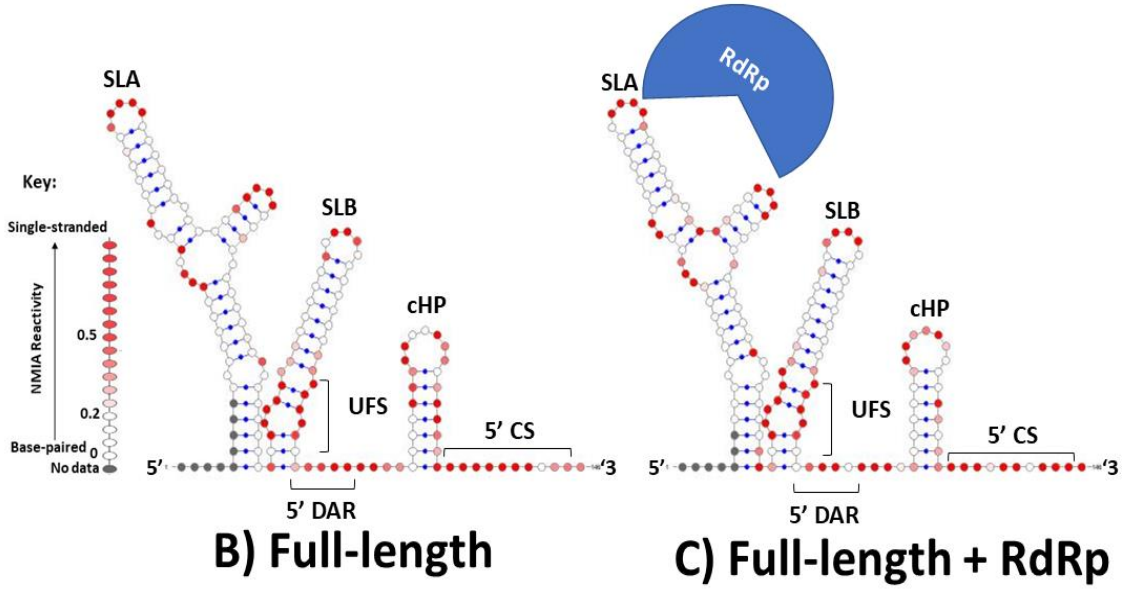

SLA
Key:
Single-stranded
NMIA Reactivity
0.5
0.2
Base-paired
0
No data
SLB
cHP
UFS
5' CS
5'
3'
5' DAR
B) Full-length
RdRp
SLA
SLB
cHP
UFS
5' CS
5'
3'
5' DAR
C) Full-length + RdRp

**Figure 68 *In vitro* SHAPE mapping of full-length DENV-2 genomic RNA in the absence and presence of DENV-2 RdRp at insect physiological temperature**

**A)** *NMIA reactivities corresponding to full-length DENV-2 genomic RNA in the absence and presence of DENV-2 RdRp in blue and red respectively. The dashed line corresponds to a NMIA reactivity of 0.2. Reactivities above this line are partially reactive and above 0.5 are highly reactive to NMIA and therefore single-stranded. Nucleotide positions of RNA structures and cyclisation sequences are indicated.* ***B)*** *Full-length DENV-2 genomic RNA NMIA reactivities overlaid onto the first 145 nts of DENV-2 genomic RNA.* ***C)*** *Full-length DENV-2 genomic RNA in the presence of DENV-2 RdRp NMIA reactivities overlaid onto the first 145 nts of DENV-2 genomic RNA. Overlays of NMIA reactivities conducted using VARNA.*

### 4.2.7 *In vitro* SHAPE mapping of DENV-2 genomic RNA results summary

Summary of *in vitro* SHAPE mapping experimental results at mammalian physiological temperature:

**Table 4 Summary of *in vitro* SHAPE mapping experimental results at mammalian physiological temperature**

FL = full-length

| | Transcript: | 500 nt | | | FL | | |
|---|---|---|---|---|---|---|---|
| | Fidelity of structural feature: | Linear | UFS | cHP | Linear | UFS | cHP |
| Condition | Absence *trans*-activating factors | ✓ | × | × | ✓ | × | × |
| | + DENV-2 RdRp | ✓ | × | × | ✓ | × | × |
| | + DENV HEK293T replicon cellular extract | ✓ | × | ✓ | ✓ | × | ✓ |

Summary of *in vitro* SHAPE mapping experimental results at insect physiological temperature:

**Table 5 Summary of *in vitro* SHAPE mapping experimental results at insect physiological temperature**

FL = full-length

| | Transcript: | 500 nt | | | FL | | |
|---|---|---|---|---|---|---|---|
| | Fidelity of structural feature: | Linear | UFS | cHP | Linear | UFS | cHP |
| Condition | Absence *trans*-activating factors | ✓ | × | × | ✓ | × | × |
| | + DENV-2 RdRp | ✓ | × | × | ✓ | × | ✓ |

**<u>Key:</u>**

✓ **=** presence of RNA structure

× **=** absence of RNA structure

## 4.3 Discussion

### 4.3.1 The linear DENV-2 genomic conformation is predominant in the absence of *trans*-activating factors at mammalian and insect physiological temperatures.

The plasticity of the positive-sense single-stranded RNA genome is fundamental to the replication cycle of DENV, the cyclic form vital to negative-sense RNA synthesis, repositioning the promoter-RdRp complex in proximity to the 3´SL (Filomatori *et al.*, 2006; Lodeiro and Filomatori, 2009). The cyclic form of the DENV-2 genome is suggested to be thermodynamically favoured, previously identified to form spontaneously in the absence of *trans*-activating factors in the context of the full-length DENV genome and utilising RNA constructs containing intact 5´-3´ cyclisation sequences (Alvarez, Lodeiro, *et al.*, 2005). However, despite this, a balance between the two structural forms is essential to efficient viral replication (Villordo, Alvarez and Gamarnik, 2010). It is likely that both structural forms of the genome are present throughout replication at both mammalian and insect physiological temperatures, establishing a controlled viral replication cycle.

In an environment where *trans*-activating factors are absent it was predicted that the full-length DENV-2 genome would spontaneously form the circular conformation, identified using specific structural motifs corresponding to the circular form of the genome and in comparison to a 5´500 nt truncated DENV-2 RNA transcript, unable to cyclise. It was expected that the circular genomic conformation would form at both mammalian and insect physiological temperatures, perhaps further stabilized at 28 °C owing to the ability of DENV to establish a persistent infection. However, our extensive studies at both mammalian and insect temperatures in the absence of *trans*-activating factors show that the linear conformation is the predominant genomic form. It may be that in the absence of *trans*-activating factors, and therefore an environment not conducive to replication, the DENV-2 genome does not require the cyclic conformation. It is possible that the linear conformation acts to inhibit minus strand RNA synthesis and therefore in an environment not conducive to viral replication it may be that

the DENV-2 linear genomic conformation is stabilised (Villordo, Alvarez and Gamarnik, 2010). Previous work has suggested that the DENV-2 genome is inherently unstable *in vitro* (Liu *et al.*, 2016). We further hypothesise, that the DENV-2 serotype used for our structural studies is inherently unstable and therefore predominantly remains in the linear conformation *in vitro* in the absence of *trans*-activating factors (Sasmono *et al.*, 2015).

### 4.3.2 The influence of DENV-RdRp binding on DENV RNA structure at mammalian and insect physiological temperatures

The binding of RdRp to 5´SLA is fundamental to viral replication. However, although it is known that the DENV-RdRp binds to SLA, genomic cyclisation is required to bring the RdRp bound 5´SLA into the proximity of the 3´SL, thereby promoting RdRp transfer to 3´SL, allowing for the initiation of negative-sense RNA synthesis and viral replication. Additionally, it has been proposed that the RdRp binds to SLA when the UFS is formed, as the presence of the UFS increases the affinity of the RdRp for SLA when in the linear conformation. It is suggested that the loss of the UFS structure results in a reduced affinity of RdRp for SLA and therefore promotes the transfer of the RdRp to the 3´ end of the genome (Liu *et al.*, 2016). Consequently, it was hypothesised that the binding of DENV-RdRp may cause the switch from the linear genomic conformation to the circular conformation.

Using both the 5´500 nts and full-length DENV RNA transcripts at mammalian and insect physiological temperatures, we analysed the influence of DENV-2 RdRp on DENV-2 RNA structure in the first 145 nts of the DENV genome. In comparison to the structure of DENV RNA in the absence of bound RdRp, we determine both the full-length and 5´500 nts transcripts to remain in the linear conformation when bound to RdRp *in vitro*. Both transcripts, in the presence of RdRp, displayed a NMIA reactivity profile indicative of the linear form whereby the top-loop of SLB and 5´CS were reactive and therefore predicted to be single-stranded. Consequently, the binding of RdRp alone is not sufficient to cause the cyclisation of the DENV full-length genome in an *in vitro* context, and therefore requires other factors of either viral or host cell origin in order to cyclise the genome.

However, upon RdRp binding to DENV genomic RNA some significant structural changes were observed, altering the conformation of downstream stem-loops SLB and cHP at both mammalian and insect temperatures. Notably, the DENV-2 UFS remains structurally unstable in the presence of DENV-RdRp.

The UFS has been suggested to be a critical regulatory structure in DENV replication, regulating the binding of the RdRp to SLA (Liu *et al.*, 2016). *In vitro,* in the absence and presence of DENV RdRp, we did not observe a stable UFS, but consistently an unstable, single-stranded UFS at both mammalian and insect physiological temperature. We hypothesise the inherent instability of the DENV-2 UFS is as a result of its inherent genomic instability and increased replication ability whereby the UFS structure is unfolded and stabilised frequently as a result of genomic cyclisation during replication (Sasmono *et al.*, 2015). Moreover, despite RdRp binding, the UFS ceases to stabilise and the genome remains in the linear conformation. We hypothesise that, upon RdRp binding to 5´SLA the UFS remains unstable, preparing the genome to cyclise as the UFS structure is unfolded during genomic cyclisation. Therefore, the binding of RdRp may be acting to further promote the instability of the DENV UFS, preparing the genome for genomic cyclisation a critical first step in DENV replication.

It has been suggested that the cHP element is a critical determinant of the DENV life-cycle, with regulatory roles associated with both viral replication and translation. In terms of viral replication, the cHP has been shown to operate independently of its sequence – indicating that it is likely to function in RNA synthesis as part of the overall topology of the 5´end of the genome. It is proposed that the cHP element acts to stall the ribosome at the correct AUG initiation site and in doing so allowing for the rearrangement of RNP complexes around the start codon or acting to recruit other factors to allow the genomic conformational shift from linear to circular necessary for RNA synthesis. Furthermore, if rounds of RNA synthesis are closely coupled to rounds of translation, the cHP element could be acting to enhance RNA synthesis by promoting virus translation during the viral life-cycle. It was also suggested that the formation of the cHP element may favour the cyclised genomic state (Clyde and Harris, 2006; Clyde, Barrera and Harris, 2008). However, the function of cHP during viral RNA synthesis remains unknown.

This information may be translated to our studies; in the absence of DENV RdRp or ribosomal initiation factors there is no need for the cHP structure to be maintained and therefore remains unstable in this context. In the context of the full-length genome bound to RdRp at 37 °C, the binding of the RdRp alone is not sufficient to cause genomic cyclisation and therefore the structure remains unstable. On the other hand, in the context of the full-length genome at 28 °C, when bound to DENV RdRp the cHP structure appears to re-form and stabilise whilst remaining in the linear conformation. We hypothesise that DENV genomic RNA is more stable when replicating at the lower temperature of 28 °C, evidenced by the ability to establish a persistent infection in mosquito cells. Therefore, upon the binding of RdRp the structural changes induced by RdRp are more prevalent and can be seen by the restoration of the cHP structure. This action caused by the binding of RdRp, whilst not inducing the cyclisation of the genome alone, may be acting to stabilise the cHP structure, preparing for replication complex recruitment or preparing to form the circular genomic conformation, or aiding start codon selection in mosquito cells (Clyde and Harris, 2006; Clyde, Barrera and Harris, 2008).

Consequently, our studies suggest that DENV RdRp binding to DENV genomic RNA *in vitro* acts to prepare the genome, via the structural rearrangement of down-stream *cis*-acting RNA structures, notably SLB, UFS and cHP, for genomic cyclisation. The binding of DENV-2 RdRp alone does not cause the linear to circular genomic switch but does act to prepare the genome for genomic cyclisation – a critical first stage in the DENV replication cycle.

### 4.3.3 The influence of HEK293T DENV replicon protein extract on DENV RNA structure at mammalian physiological temperature

To analyse DENV RNA structure in the context of a viral infection, we mapped the structure of the 5´ end of the DENV genome in the presence of stably expressing DENV HEK293T replicon protein extract at mammalian physiological temperature. In the context of the 5´500 nt transcript, unable to cyclise, upon the addition of stably expressing DENV HEK293T replicon protein extract the UFS is observed to become more unstable. However, the cHP structure stabilises upon the addition of the protein extract. These observations are also detected in the context of the full-

length genome and the full-length genome is observed to remain in the linear conformation.

The influence of cHP in viral replication is discussed in section 4.3.2. We further hypothesise, that in the context of a more representative host and viral physiological environment *in vitro*, DENV HEK293T replicon protein *trans*-activating factors are acting to induce structural changes to DENV genomic RNA, preparing the genome for cyclisation, evidenced by the stabilisation of cHP and the continuous instability of the DENV UFS. Also, in the context of the full-length DENV genome the circular genomic form is not observed in the presence of *trans*-activating factors. The interaction of the 5´UTR with translation initiation factors could act to de-stabilise the 5´-3´ ends of the viral genome and therefore induce the formation of the linear genomic form (Villordo, Alvarez and Gamarnik, 2010). On the other hand, we observe downstream genomic structural rearrangements and therefore viral and host *trans*-activating factors could be acting to prepare the genome for genomic cyclisation. We hypothesise that HEK293T DENV replicon *trans*-activating factors do not act to induce the formation of the circular DENV genome but act to further prepare DENV genomic RNA for genomic cyclisation and viral replication.

This study could be further extended via the analysis of DENV genomic RNA structure in the presence of DENV replicon protein extract of mosquito origin (C6/36 or Aag2) at mosquito physiological temperature, to further deduce the biological mechanisms behind the linear to circular genomic switch, considering the important influence of the insect host environment.

### 4.3.4 The limitations of *in vitro* SHAPE mapping and future directions

Due to a multitude of biochemical advances, *in vitro* SHAPE mapping is routinely used for RNA structural analysis. However, there are significant limitations involving its usage *in vitro*:

Primarily, *in vitro* transcription is used to generate the template RNA for structural analysis which is then unfolded and re-folded, in the absence or presence of *trans*-activating factors. Despite the template RNA routinely quality analysed, via

denaturing MOPS agarose gel electrophoresis, this action does not take into account the multitude of conformational forms or truncations present within the *in vitro* transcribed RNA sample. Therefore, this is not controlled in the *in vitro* SHAPE experiment and the experimental data samples all conformations at one time. One way to overcome this experimental limitation is the utilisation of in-gel SHAPE which can be used to map the structure of RNA molecules of the same structural conformation. A mixed structural population of RNAs is separated by non-denaturing agarose gel electrophoresis and the individual conformers probed, with SHAPE reagent, within the gel matrix before extraction and computational analysis of RNA structure. This technique was used to examine the RNA structure of the monomeric and dimeric species of the HIV-1 packaging signal RNA (Kenyon *et al.*, 2013). However, this technique is currently technically challenging, involving the use of non-denaturing agarose gels which require experienced handling.

In contrast, SHAPE mapping has been coupled to deep sequencing technology, termed SHAPE-seq, characterizing the structure of multiple RNAs within a single experiment at the transcriptome level. However, while providing higher resolution structural information, this technique is limited to relatively short *in vitro* transcribed RNA molecules of approximately 300 nucleotides and demands expert statistical analysis (Lucks *et al.*, 2011).

Furthermore, a common problem upon SHAPE data analysis is NMIA reactivity induced experimental variability due to RNA breathing in solution and the average sampling of multiple RNA conformations. This problem is overcome by the utilisation of multiple experimental repeats and the inclusion of average SHAPE reactivities in thermodynamic predictions.

Fundamentally, the main experimental limitation of *in vitro* SHAPE mapping is that the determined RNA structure may not be truly reflective of RNA structure intracellularly or *in vivo*, regardless of the experimental inclusion of *trans*-activating factors. To overcome this fundamental limitation, *in vivo* or intracellular SHAPE mapping is coming to the forefront of SHAPE RNA structural analysis. Intracellular SHAPE mapping utilises SHAPE reagents specifically designed to modify RNA within the cellular environment eliminating the fundamental flaw of *in vitro* SHAPE mapping experimentation and is consequently the subject of Chapter 5.

# Chapter 5 Intracellular SHAPE analysis of DENV-2 genomic RNA structure

## 5.1 Introduction

RNAs of interest have been extensively structurally probed and characterised using SHAPE chemistry *in vitro*. Furthermore, *in vitro* RNA structure probing experiments have improved the accuracy of *in silico* structural models, allowing for the inclusion of RNA structural data, adding extensively to our knowledge of RNA structural motifs. Previously, we utilised *in vitro* SHAPE to analyse the RNA structure of the 5´ extremity of the DENV-2 RNA genome – investigating the influence of *trans*-activating factors on genome conformation (Chapter 4). These results led us to propose that, in an *in vitro* context, *trans*-activating factors influence local RNA conformation, acting to prepare the DENV-2 RNA genome for cyclisation and subsequently genome replication following the initiation of genomic transcription. However, RNA structure in the intracellular context, during active viral replication, is likely to be more complex than can be observed *in vitro*. To investigate this further, we chose to probe DENV-2 genomic RNA structure within live cells during active virus replication. The aim of experimentation was to further elucidate the biological mechanisms influencing conformational changes within the Flavivirus genome associated with initiation of replication and switching between the linear and circular genomic forms.

RNA structure in cells is influenced by a multitude of factors for example; the rate of transcription, local solution conditions, RNA chaperones, the binding of *trans*-activating factors and viral infection, all of which are challenging to truly replicate *in vitro* (Schroeder *et al.*, 2002; Spitale *et al.*, 2013). For example, local solution conditions vary throughout the host cell. pH varies according to the functional requirements of the organelle, ranging from pH 6 to 6.7 through the golgi apparatus and therefore throughout viral encapsidation and release (Rivinoja *et al.*, 2011). Previous studies have demonstrated that RNA folding can differ intracellularly and *in vitro*. For example, human telomerase RNA forms a 5´ phylogenetically conserved pseudo-knot which is not observed in an intracellular context. This was identified following HeLa cell nuclear fractionation, RNA modification and subsequent radiolabelled primer extension in comparison to the treatment of *in vitro* transcribed telomerase RNA (Antal *et al.*, 2002). Additionally, exon mutants of the *Tetrahymena* pre-rRNA shows a 100-fold decrease in self-splicing *in vitro*.

However, self-splicing capability was fully rescued when the pre-rRNA was expressed in *E. coli* therefore suggesting that RNA folding is facilitated intracellularly, not accounted for in an *in vitro* context as determined by experimental differences in splicing rates (Nikolcheva and Woodson, 1999). Therefore, for our studies, it was important to consider the intracellular environment and its influence on RNA folding in the context of DENV-2 replication.

Intracellular SHAPE probes the structure of RNA molecules within living cells – thus investigating the effect of the intracellular environment on RNA conformation. SHAPE reagents, for example 2-methylnicotinic acid imidazolide (NAI), specifically designed and synthesised to have properties amenable for RNA labelling inside living cells, such as increased solubility and half-life, are introduced in cell culture (Lee *et al.*, 2017). The intracellular SHAPE reagent (NAI) acts to modify total cellular RNA at nucleotide positions that are single-stranded, and therefore un-restrained with a free 2' OH group. Double-stranded, restrained nucleotides remain unmodified (Fig. 69).

Following RNA extraction of total cellular RNA, NAI modified nucleotides are identified by primer extension and fragment size analysis, as described in detail previously for NMIA modifications during *in vitro* SHAPE (Chapter 4, section 4.1). Briefly, NAI was used at limiting concentrations, so that individual RNA molecules are only modified at a single position. Reactivity at individual nucleotides was then measured by capillary electrophoresis fragment size analysis of reverse transcription termination products (using DENV specific primers). This information can then be used to restrain the RNA sequence of interest, with corresponding SHAPE reactivities, or overlaid onto a predicted structural model producing an accurate RNA structural model representative of the intracellular environment (Spitale *et al.*, 2013; Watters *et al.*, 2016; Lee *et al.*, 2017). Statistics of SHAPE reactivity-based RNA structural prediction accuracy are discussed in Chapter 4 section 4.1.

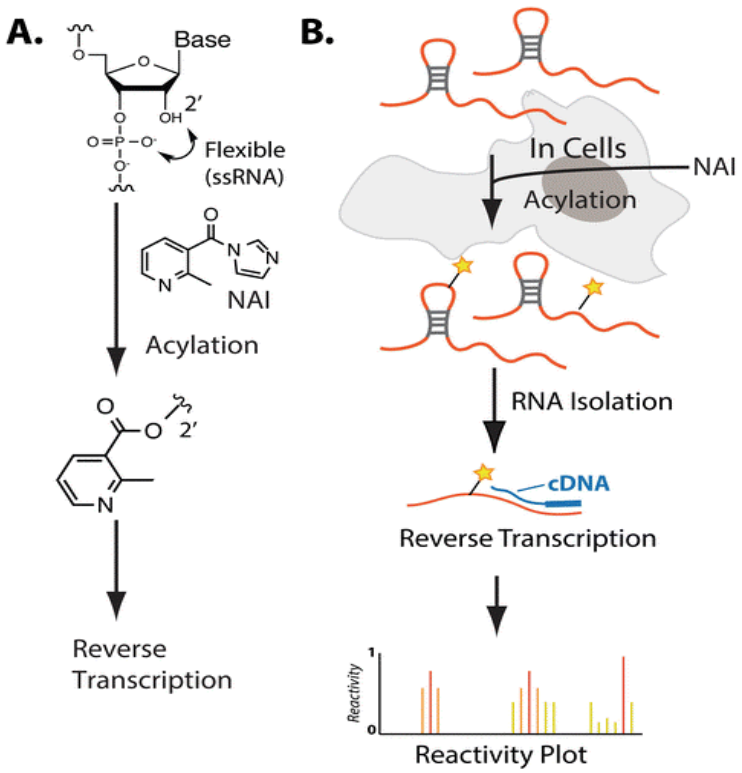



**Figure 69 Intracellular SHAPE modification by NAI:**

***A)*** *The modification of single-stranded RNA (ssRNA) nucleotides by NAI.* ***B)*** *Overview of the experimental process of intracellular SHAPE using NAI to determine RNA structure. Taken from* (Feng, Chan and Spitale, 2017).

To our knowledge, intracellular SHAPE mapping of viral RNA structure during active virus replication has not been previously published. Intracellular SHAPE was used to uncover dynamic RNA-RNA interactions between the UTR regions of Foot-and-Mouth disease virus, in the context of transfected RNA transcripts, un-representative of full virus infection due to the transfection procedure and non-replicating (Diaz-Toledano, Lozano and Martinez-Salas, 2017). Additionally, in virion SHAPE has been utilised, for example, following the purification of intact poliovirus virions to determine the structure of full-length poliovirus RNA (Burrill *et al.*, 2013). Therefore, investigation of DENV RNA structure during active virus replication/ infection posed a novel avenue to explore DENV RNA structure, in a more physiologically relevant context during active viral replication.

Virus entry, replication and assembly are dynamic, highly co-ordinated processes fundamentally requiring specific interactions with host cell components, often

occurring within sub-cellular compartments, concentrating these interactions to one location but acting to evade the host immune response, promoting their replication (Inoue and Tsai, 2013). Flavivirus replication is known to occur within a sub-cellular compartment termed the replication complex, formation of which is virally induced at the ER membrane. Combined immunoelectron microscopy and electron tomography studies have investigated the architecture of Flavivirus-induced membrane structures associated with functional replication complexes within mammalian and insect cells (Welsch *et al.*, 2009; Gillespie *et al.*, 2010; Offerdahl *et al.*, 2012; Junjhon *et al.*, 2014). Following infection, a complex set of DENV-modified structures are formed, including double-membrane vesicles enclosed within membrane packets, convoluted membranes and tubular structures were observed, covered by a single continuous ER membrane (Welsch *et al.*, 2009). Double membrane vesicles, were found to be comprised of viral RNA, replication proteins and cellular proteins forming the active viral replication complex where negative-sense RNA synthesis occurs and viral replication takes place (Mackenzie, 2005; Salonen, Ahola and Kaariainen, 2005; Miller and Krijnse-Locker, 2008; Welsch *et al.*, 2009).

As evidenced, DENV interacts with host cell membranes to induce the formation of the replication complex, the site of viral replication. We therefore sought to extend our RNA structural analysis utilising intracellular SHAPE to investigate the structure of viral RNA within an environment enriched for active viral replication complexes. Due to the containment requirements of DENV, we isolated the membrane fraction from stably expressing DENV HEK293T replicon cells, to determine the structure of DENV RNA within an environment enriched for active replication complexes thus providing vital insight into the conformation of DENV-2 RNA undergoing replication.

To date and to our knowledge, DENV-2 RNA structure has not been investigated intracellularly following infection or during active viral replication in the context of stably expressing DENV replicon cells. Consequently, the aims of our intracellular experimentation are detailed as follows:

1. Optimisation of DENV intracellular SHAPE

2. Compare and contrast *in vitro* SHAPE experimental results at both mammalian and insect physiological temperature with intracellular SHAPE experimental results following total RNA extraction of stably expressing DENV HEK293T replicon cells or following DENV-2 infection
3. Compare and contrast intracellular SHAPE results from DENV RNA extracted from the membrane fraction of stably expressing DENV HEK293T cells enriched for active viral replication complexes with intracellular structural results from total RNA extraction

This chapter therefore describes the optimisation of the intracellular SHAPE method utilised for our studies. We then present the structure of 5´ extremity DENV-2 RNA extracted from mammalian and insect cells following DENV-2 infection or extracted from DENV stably expressing HEK293T replicon cells. We then present the structure of DENV genomic RNA within DENV replication complex enriched fractions.

## 5.2 Results

### 5.2.1 Optimisation of DENV-2 Intracellular SHAPE

#### 5.2.1.1 Determination of NAI activity and optimum incubation time

NAI was synthesised following the protocol described in (Lee *et al.*, 2017). 1M stocks were stored at – 80°C and allowed to equilibrate to room temperature before use. We analysed the activity of synthesised NAI prior to use, considering incubation at 37°C or 28 °C, in comparison to the incubation of insect/mammalian cells in 10% DMSO, utilised as a negative, unreactive modification control as experimental NAI is resuspended in 10% DMSO. An optimal incubation time of 15 minutes allowed for the modification of total RNA and minimised the cellular toxicity caused by DMSO which prevented the extraction of high quality total RNA. We determined optimal NAI activity following a 15-minute incubation period, at either 37 or 28 °C, utilising *in vitro* transcribed DENV RNA following RNA extraction, reverse transcription and analysis using QuSHAPE (Karabiber *et al.*, 2013) (Fig. 70, QuSHAPE pipeline described in section 4.2.1).

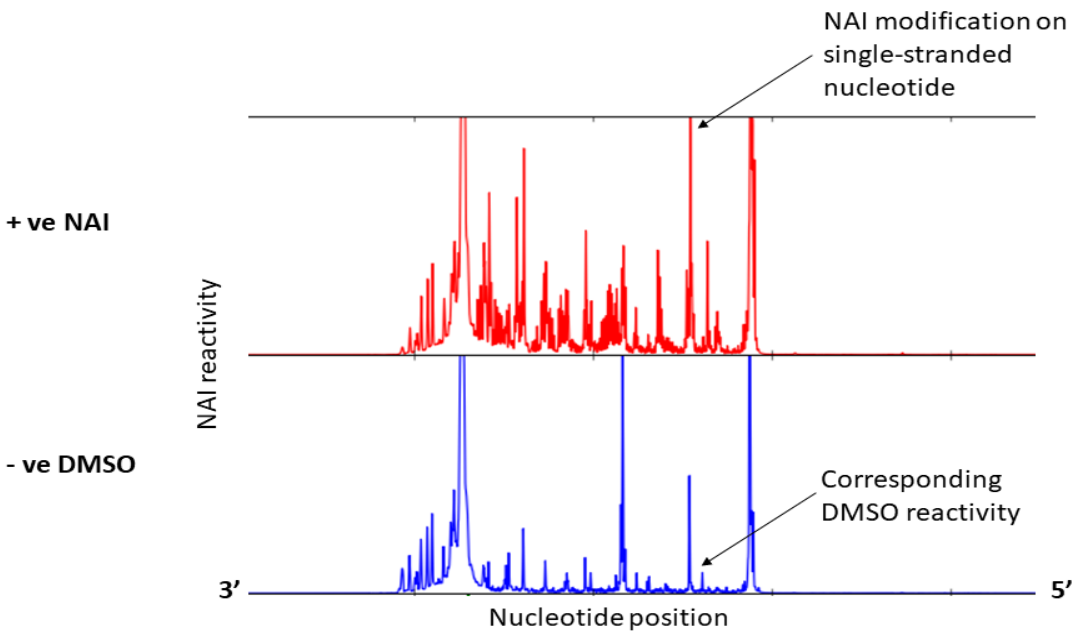



**Figure 70 The activity of NAI:**

*NAI reactivity chromatograms analysed using QuSHAPE* (Karabiber *et al.*, 2013)*. The red chromatogram corresponds to NAI treated RNA, the blue chromatogram corresponds to DMSO treated RNA.*

### 5.2.1.2 Extraction of total RNA from stably expressing DENV HEK293T replicon cells

Successful SHAPE experimentation is reliant on high quality RNA for structural analysis. Firstly, DENV-2 replicon RNA was transfected into HEK293T cells and maintained under puromycin selection (Fig. 71). The stable expression of GFP was confirmed via fluorescent microscopy analysis and DENV non-structural protein expression analysed by RT-PCR (Figs 71 and 76).

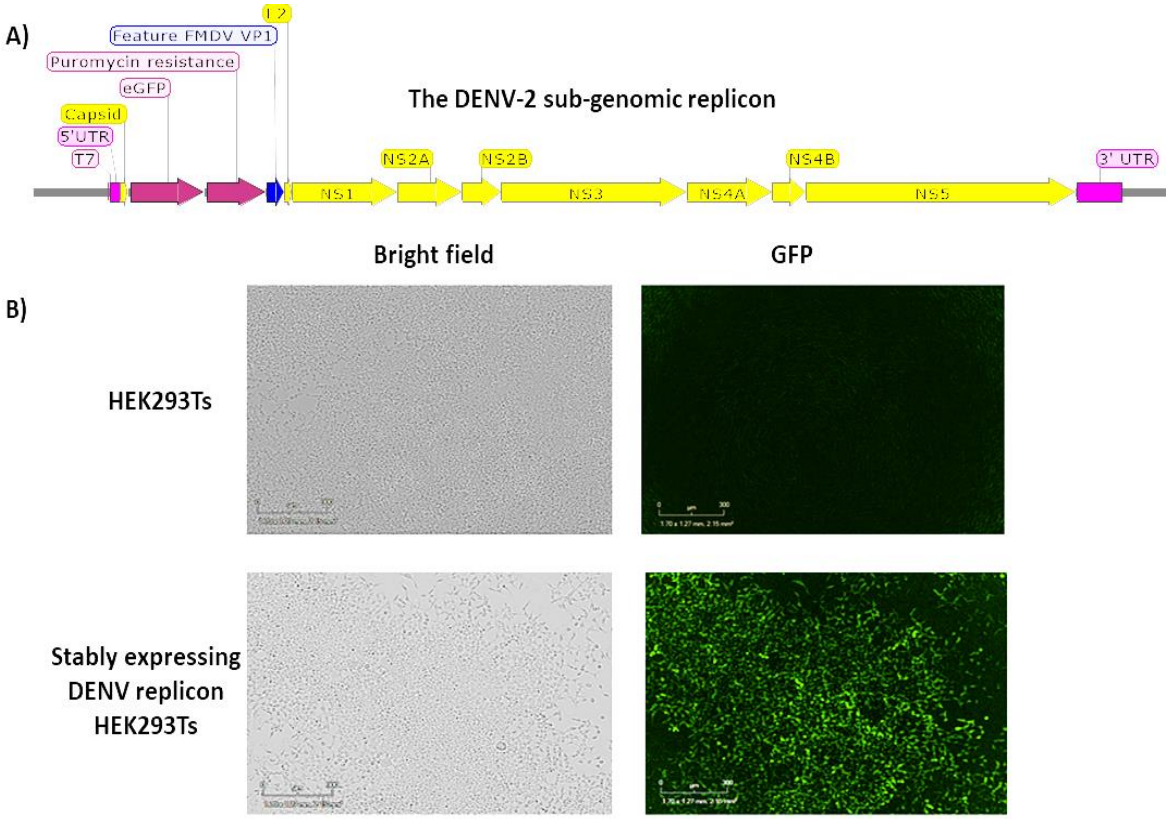



**Figure 71 Confirmation of the stable expression of DENV-2 non-structural proteins:**

***A)*** *DENV-2 sub-genomic replicon map. Clone kindly provided by Dr. Andrew Davidson (University of Bristol).* ***B)*** *Confirmation of GFP expression under puromycin selection and therefore the establishment of stably expressing DENV-2 HEK293T replicon cells. Imaged with Incucyte®*

Stably expressing DENV HEK293T replicon cells were treated with NAI or DMSO and total RNA extracted. The quality of total extracted RNA was visually confirmed by denaturing MOPS agarose gel electrophoresis. As seen in figure 72, the extracted RNA was of high quality since the 18S and 28S ribosomal subunits were clearly detected without evidence of degradation.

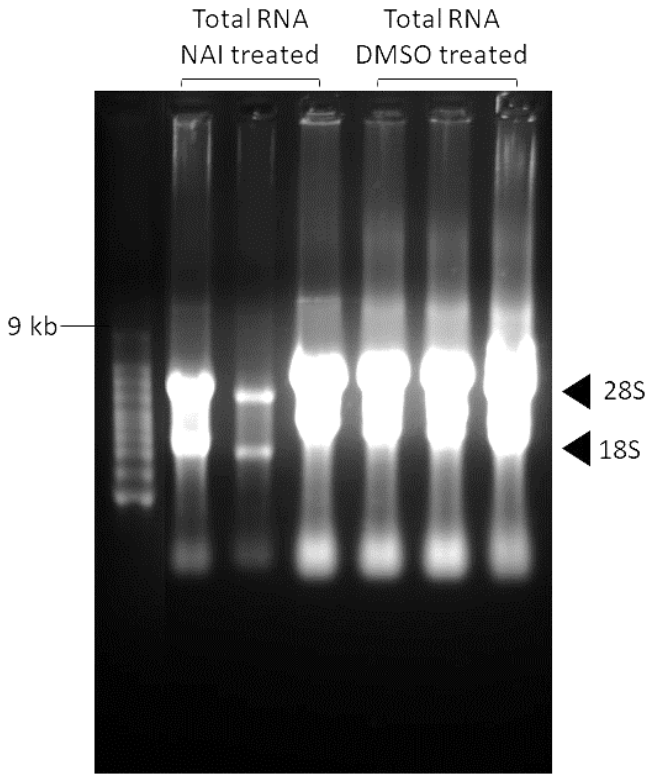


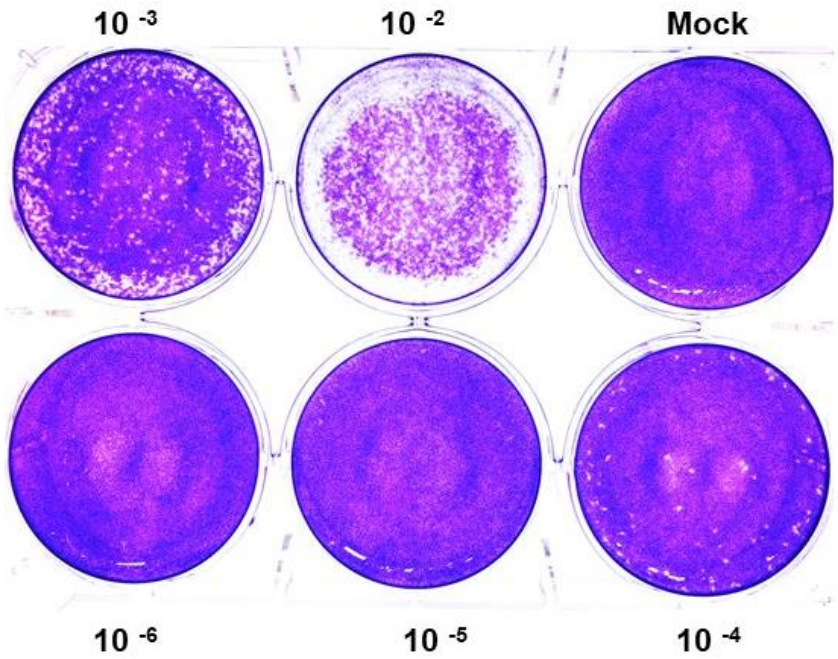

**Figure 72 Quality analysis of total RNA extracted from stably expressing DENV HEK293T cells:**

*Total RNA analysed via denaturing MOPS agarose gel electrophoresis in comparison to a Millennium RNA size ladder (Ambion). Intact 28S and 18S ribosomal subunits are visualised.*

### 5.2.1.3 The culture of infectious DENV-2 and use in Intracellular SHAPE

Infectious DENV-2 was produced following transfection into BHK-21 cells and further expansion in C6/36 cells for 1 week before determination of viral titre via plaque assay (Fig. 73A). For intracellular SHAPE, we infected 70% confluent T175 flasks of BHK-21 or C6/36 cells with an MOI of 0.0002 for 72 hours prior to total RNA extraction (Fig. 73B).

**A)**




**B) Determination of *pfu/ml* and MOI:**

***Pfu*/ml*:***

$$\frac{\mathbf{57}}{\mathbf{0.001\ \times 0.2\ \boldsymbol{ml}}} = \mathbf{2.85\ \times 10^{6}}$$

**MOI:**

$$\frac{\mathbf{0.0015 \times 2850000}}{\mathbf{1840000}} = \mathbf{0.0002}$$

**Figure 73 DENV-2 plaque assay - *pfu* and MOI calculations:**

***A)*** *DENV-2 plaque assay in 6 well format conducted using BHK-21 cells. DENV-2 was titred following the infection of C6/36 cells for dilutions labelled accordingly. Mock refers to an uninfected negative control.* ***B)*** *Pfu/ml and MOI calculations.*

#### 5.2.1.4 Optimisation of DENV infection time for intracellular SHAPE analysis

It was critical to establish a suitable DENV infection time frame, prior to SHAPE reagent treatment and total RNA extraction, allowing time for the establishment of infection but before mammalian host cell death, resulting in the inability to extract intact total RNA. This was more of a consideration in mammalian cells than insect cells, where DENV can establish a persistent infection with no apparent cytopathic effect. We therefore sought to determine the optimal time-frame for DENV infection, prior to induction of cytopathic effect in mammalian cells, whereby viral proteins could be detected. It was determined, following time-course infection experimentation and analysis by western blot of the presence of DENV-2 NS5 in comparison to the host cellular control of β-Actin, that the optimal time-frame for infection prior to RNA extraction was 72 hours (Fig. 74). This time-frame was therefore used throughout intracellular SHAPE experimentation using infectious DENV-2.

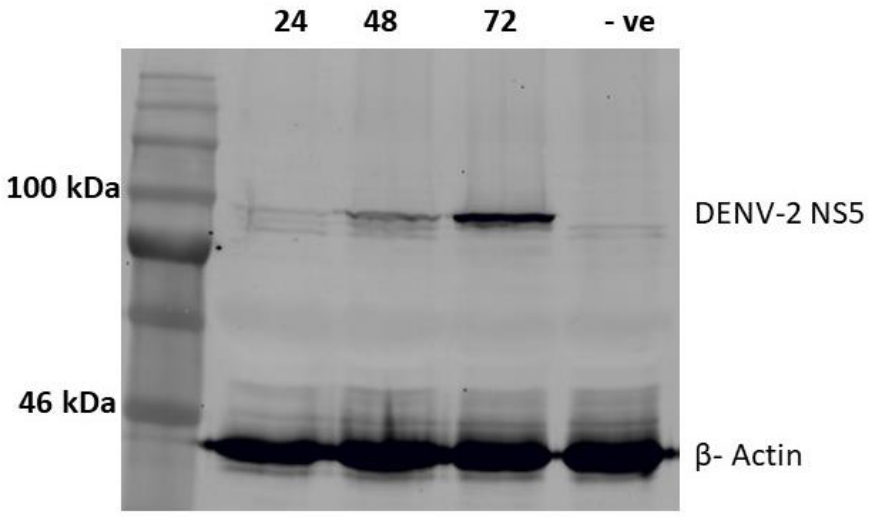



**Figure 74 Western blot analysis of DENV-2 infection time-course for optimisation of Intracellular SHAPE:**

*BHK-21 cells were infected with DENV-2 and lysed at 24, 48 and 72 hours post-infection. Since DENV-2 NS5 could be sufficiently detected at 72 hours, this time frame was used for intracellular SHAPE experimentation. DENV-2 NS5 is 100 kDa and the house keeping gene β-Actin is 42 kDa. The negative control (-ve) refers to un-infected BHK-21 cells.*

#### 5.2.1.5 Extraction of high quality total RNA following viral infection of mammalian and insect cells

As previously discussed, SHAPE experimentation is reliant on high quality RNA. Therefore, it was critical that high quality total RNA was extracted following virus infection of both mammalian and insect cells, BHK-21 and C6/36 cells, respectively. The extracted total RNA was analysed using denaturing MOPS agarose gel electrophoresis to confirm quality (Fig. 75). As seen in Figure 75 A, the mammalian extracted RNA was of high-quality since the 18S and 28S ribosomal subunits were clearly detected without evidence of degradation. Additionally, total RNA extracted from insect cells displayed a typical ribosomal RNA profile whereby heat denaturation causes the 28S subunit to migrate alongside the 18S subunit due to heat scission of the 28S subunit (Winnebeck, Millar and Warman, 2010). Therefore, total RNA extracted from insect cells was intact and viable for experimental use due to the clear presence of the undegraded 18S subunit (Fig. 75B).

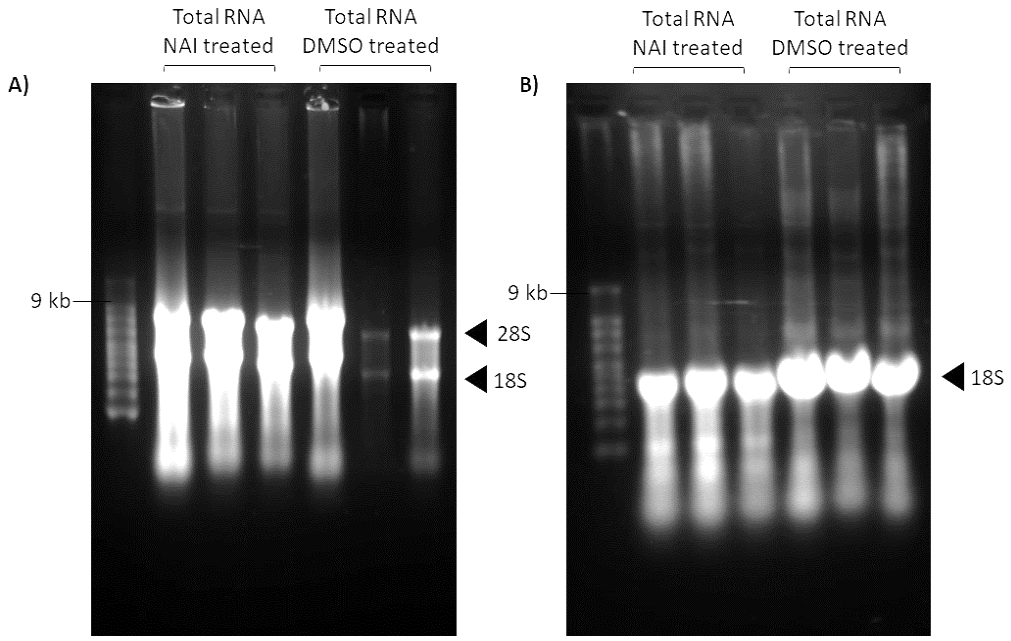



**Figure 75 Quality analysis of total RNA extracted from DENV infected BHK-21 and C6/36 cells:**

*Total RNA analysed via denaturing MOPS agarose gel electrophoresis in comparison to a Millennium RNA size ladder (Ambion).* ***A)*** *Total RNA extracted from DENV-2 infected mammalian BHK-21 cells. Intact 28S and 18S ribosomal subunits are visualised.* ***B)*** *Total RNA extracted from DENV-2 infected mosquito C6/36 cells and intact 18S ribosomal subunits are visualised.*

#### 5.2.1.6 Confirmation of DENV protein expression following total RNA extraction

Following total RNA extraction, DENV infection or stable expression of DENV-2 non-structural proteins, the presence of viral proteins was assessed via RT-PCR amplification of the DENV NS3 genome region. Cellular lysate was analysed for the presence DENV NS5 by western blot. The action further confirmed active DENV replication in mammalian and insect cells. Following confirmation, the total RNA extracted was therefore viable for intracellular SHAPE analysis (Figs. 76 and 77).

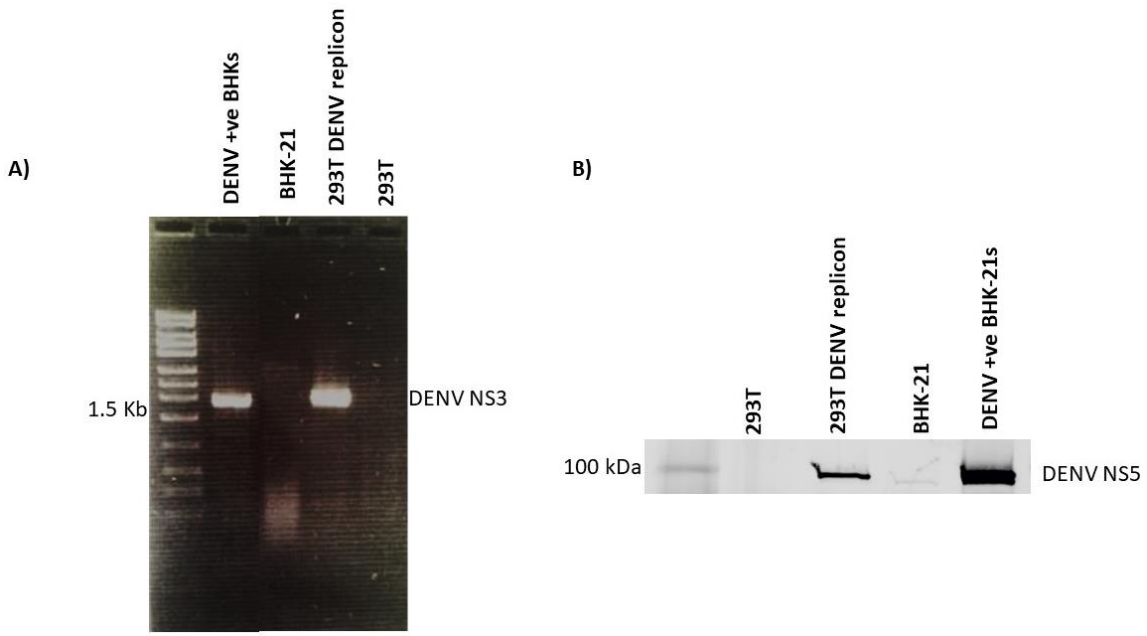



**Figure 76 Confirmation of DENV infection and the stable expression of DENV-2 non-structural proteins in mammalian cells via RT-PCR and western blot:**

***A)*** *RT-PCR amplification of DENV-NS3 following infection of BHK-21 cells with DENV in comparison to un-infected BHK-21s and RT-PCR amplification of DENV-NS3 following the stable expression of DENV non-structural proteins in HEK293T cells in comparison to naïve HEK293T cells. The DENV-2 NS3 amplicon is 1.8 kb. Size confirmed in comparison to a exACTGene 1 Kb DNA ladder (Fisher).* ***B)*** *DENV-2 NS5 Western blot of PLB lysates following the stable expression of DENV non-structural proteins in HEK293T cells in comparison to naïve HEK293T cells and DENV infection of BHK-21 cells in comparison to naïve BHK-21 cells. Equal amounts of protein were loaded. DENV-2 NS5 is 100 kDa. Size confirmed in comparison to a Colour protein standard, broad range ladder (NEB).*

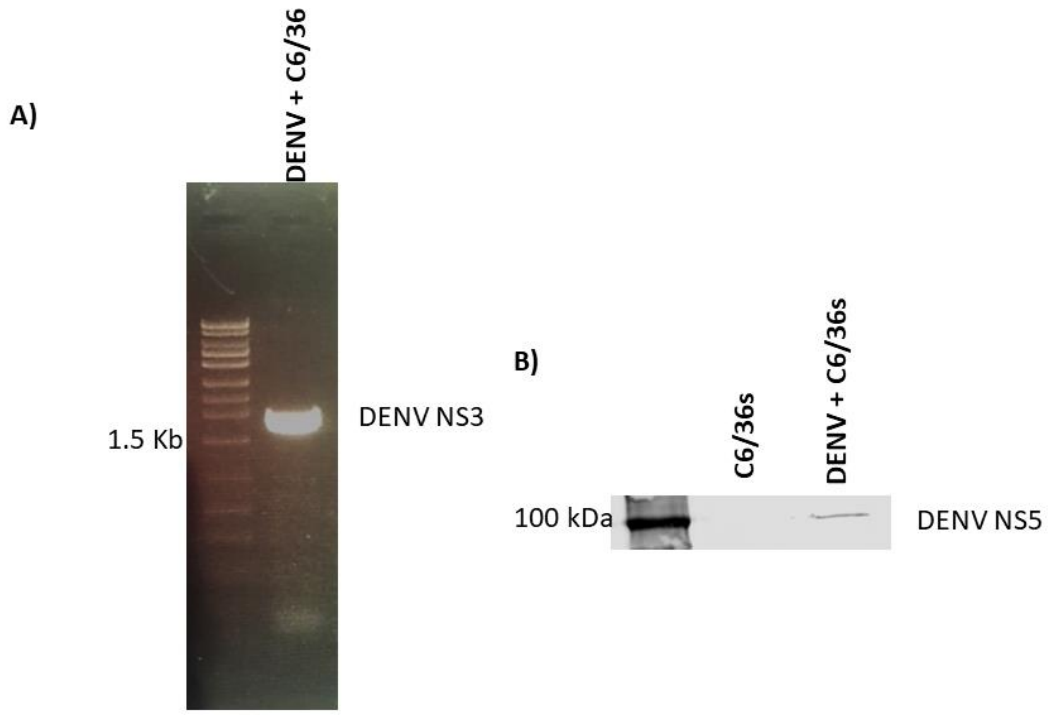



**Figure 77 Confirmation of DENV infection of C6/36 cells via RT-PCR and western blot:**

***A)*** *RT-PCR amplification of DENV-2 NS3 following DENV infection of C6/36 cells. DENV-2 NS3 is 1.8 Kb. Size confirmed in comparison to a exACTgene 1 kb DNA size ladder.* ***B)*** *Western blot analysis of DENV-2 NS5 following DENV infection of C6/36 cells in comparison to un-infected C6/36 cells. Size confirmed in comparison to a Colour protein standard, broad range ladder (NEB). DENV-2 NS5 is 100 kDa.*

#### 5.2.1.7 Cellular fractionation, membrane enrichment and total RNA extraction

Stably expressing DENV HEK293T replicon cells were treated with NAI SHAPE reagent, fractionated and total RNA was extracted from the isolated membrane fraction enriched for active viral replication complexes. The efficacy of the cellular fractionation protocol was analysed using western blot via the identification of DENV NS5 and the nuclear porin protein NUP98 (Fig. 78A). As seen in figure 78A, the cellular fractionation was successful as indicated by the enrichment of DENV-2 NS5 in the membrane fraction and NUP98 in the nuclear fraction. However, due to the fact that cellular fractionation is an enrichment process and not a purification procedure a low level of DENV-2 NS5 was detected in the nuclear and cytoplasmic fractions as well as low level NUP98 detected in the membrane fraction (Fig. 78A).

Additionally, total RNA extracted from the membrane fraction was analysed for quality using denaturing MOPS agarose gel electrophoresis (Fig. 78B). As seen in figure 78B, the extracted RNA was of high-quality since the 18S and 28S ribosomal subunits were clearly detected without evidence of degradation.

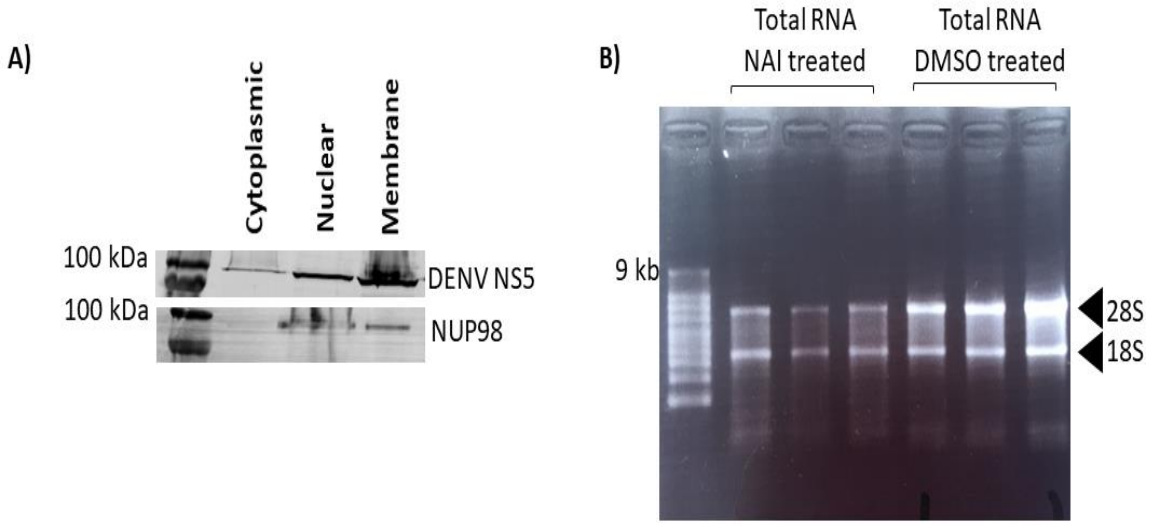



**Figure 78 Confirmation of the cellular fractionation of stably expressing DENV HEK293T replicon cells and analysis of total RNA extracted from the membrane fraction:**

***A)*** *DENV-2 NS5 western blot analysis of the cellular fractionation of stably expressing DENV HEK293T replicon cells. DENV-NS5 is 100 kDa, NUP98 is 98 kDa.* ***B)*** *Quality analysis of total RNA extracted from the membrane fraction of stably expressing DENV HEK293T cells intact 28S and 18S ribosomal subunits are visualised.*

#### 5.2.1.8 The execution of Intracellular SHAPE

The optimised protocol for intracellular SHAPE, including both intracellular SHAPE conducted on total viral RNA and intracellular SHAPE conducted on membrane fractions enriched for active viral replication complexes, is described in figure 79. In brief, cells are treated in culture with NAI, a SHAPE reagent analogous to NMIA. SHAPE fragment analysis was conducted using capillary electrophoresis by DNA sequencing and services, Dundee University. Data analysis was carried out using QuSHAPE (Karabiber *et al.*, 2013). The QuSHAPE pipeline is further described in section 4.2.1. Structural predictions were carried out via reactivity overlay (as described in section 4.2.1) or by utilising RNAstructure (Reuter and Mathews, 2010) whereby SHAPE reactivities were read as a soft pseudo free energy constraint to produce an RNA secondary structure based on the experimental data (example thermodynamic predictions can be viewed in Appendix II). As described in section 4.2.2 the linear or circular conformation of viral RNA examined was deduced using specific structural motifs diagnostic of either form.

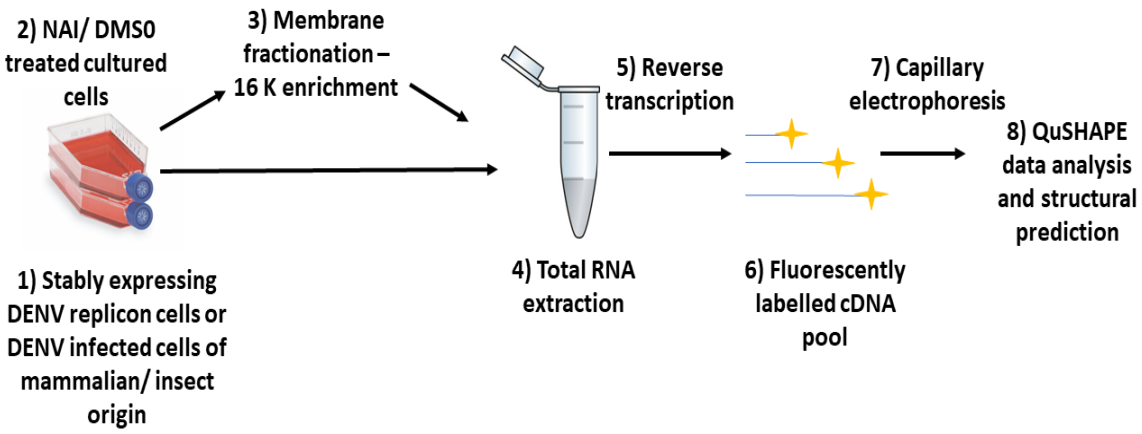



**Figure 79 The optimised Intracellular SHAPE protocol:**

***1)*** *Stably expressing DENV HEK293T replicon cells or DENV infected BHK-21 or C6/36 cells are cultured.* ***2)*** *The cultured cells are NAI or DMSO treated.* ***3)*** *Stably expressing DENV HEK293T cells are fractionated and the membrane fraction isolated.* ***4)*** *Total RNA is extracted from cultured cells or from the isolated membrane fraction of stably expressing DENV HEK293T replicon cells.* ***5 and 6)*** *Primer extension and reverse transcription is conducted on DENV genomic RNA producing a pool of cDNA fragments corresponding to NAI adduct formation.* ***7 and 8)*** *the cDNA pool is analysed by fluorescent capillary electrophoresis and RNA structure determined using QuSHAPE and the visualisation applet VARNA* (Karabiber *et al.*, 2013).

### 5.2.2 Intracellular SHAPE conducted on stably expressing DENV HEK293T replicon cells

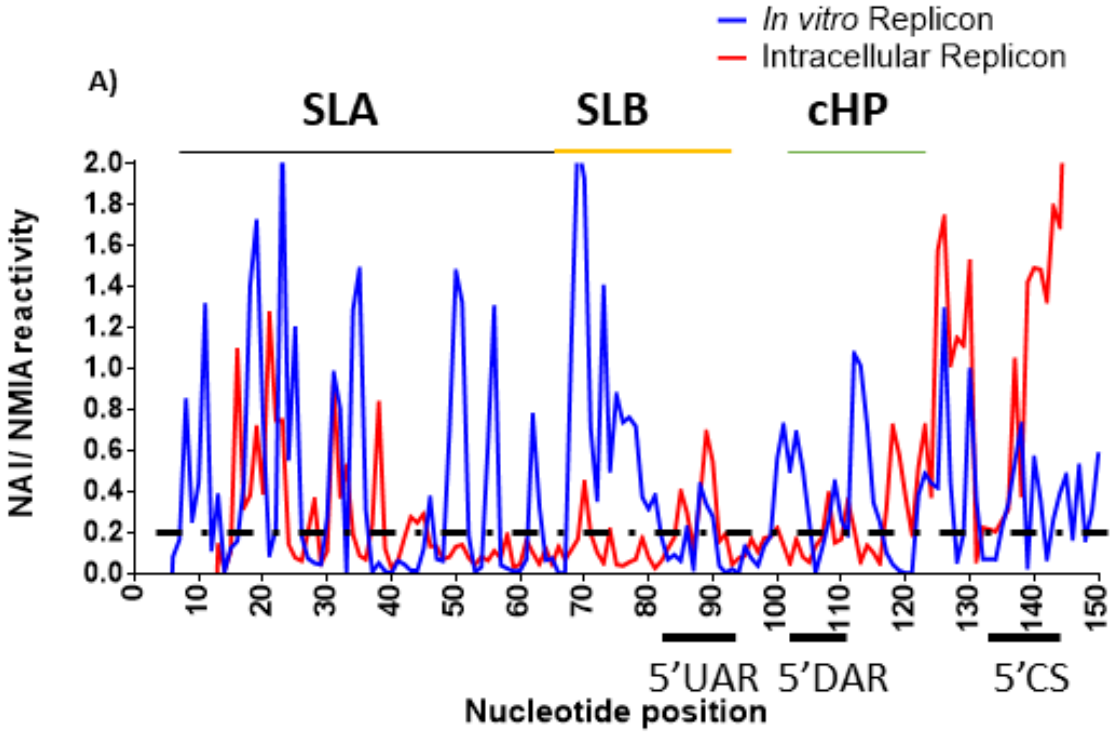
To date and to our knowledge, intracellular SHAPE mapping has not been performed, analysing the effect of active viral replication on DENV RNA structure, focusing on the 5´ extremity of DENV RNA in the context of the DENV replicon. We determined the structure of DENV RNA following NAI treatment and total RNA extraction from HEK293T cells stably expressing the DENV replicon. We completed this experimentation in the aim of further deducing the molecular mechanisms behind genomic conformational change during active viral replication. We compared our findings to the structure of full-length replicon *in vitro* transcribed RNA in the absence of *trans*-activating factors determined by *in vitro* SHAPE mapping thereby interrogating the observed structural differences between *in vitro* and intracellular SHAPE mapping.

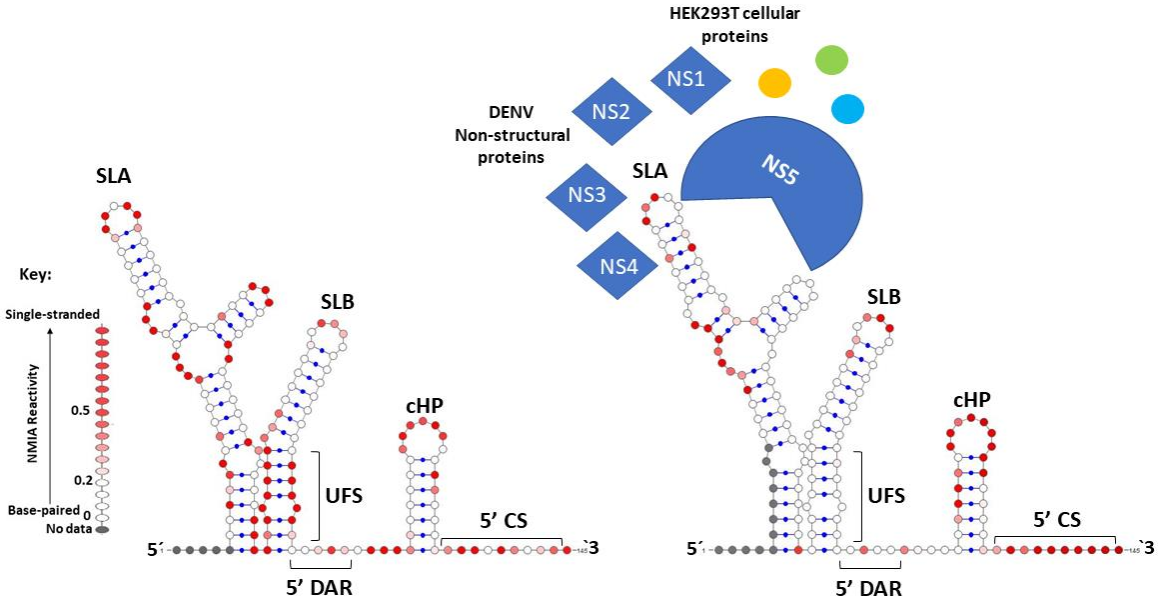
In figure 80, the normalised nucleotide reactivities to NAI/NMIA are compared, in the context of the full-length DENV replicon transcript, before being overlaid onto the predicted structure of the DENV-2 5´extremity RNA at 37 °C

*In vitro* SHAPE mapping of full-length DENV replicon RNA revealed DENV genomic RNA to be in the linear conformation, as determined by a NMIA reactive top-loop of SLB, indicating the presence of the stem-loop structure and a reactive 5´ CS (Fig. 80). The structure of cHP was determined to be relatively stable as indicated by the presence of a reactive top-loop but a predominantly base-paired stem therefore indicating the presence of this structure. The UFS, at the base of SLB was not present as displayed by the base of SLB displaying high-reactivity to NMIA and indicating this region to be predominantly single-stranded.

Similarly, intracellular SHAPE mapping of full-length DENV replicon RNA also indicated the DENV genome to be in the linear conformation, indicated by a NAI reactive top-loop of SLB and a reactive 5´ CS (Fig. 80). During active replication, in the context of the DENV replicon the cHP is more un-stable as indicated by nucleotides in the stem of the structure becoming more accessible to NAI. However, the UFS is observed for the first time to stabilise, as displayed by a reduced reactivity to NAI at the base of SLB indicating this structure to be present and stable in the context of active viral replication.

Critically, during active viral replication, in the context of stably expressing HEK293T replicon cells we observed a stable UFS for the first time, not observed using *in vitro* SHAPE mapping. Therefore, we sought to map the structure of DENV 5´ extremity RNA during DENV infection.




**B) *In vitro* replicon** **C) Intracellular replicon**

**Figure 80 *In vitro* vs intracellular SHAPE mapping normalised NMIA/NAI reactivities corresponding to full-length DENV replicon RNA at mammalian physiological temperature:**

**A)** *NMIA/NAI reactivities corresponding to the full-length replicon transcript as determined by in vitro SHAPE and as determined by intracellular SHAPE are shown in blue and red respectively. The dashed line corresponds to a NMIA reactivity of 0.2. Reactivities above this line are partially reactive and above 0.5 are highly reactive to NMIA and therefore single-stranded. Nucleotide positions of RNA structures and cyclisation sequences are indicated.* ***B)*** *Full-length replicon NMIA reactivities overlaid onto the first 145 nts of DENV-2 genomic RNA.* ***C)*** *Full-length replicon NAI reactivities overlaid onto the first 145 nts of DENV-2 genomic RNA. Overlays of NMIA/ NAI reactivities conducted using VARNA.*

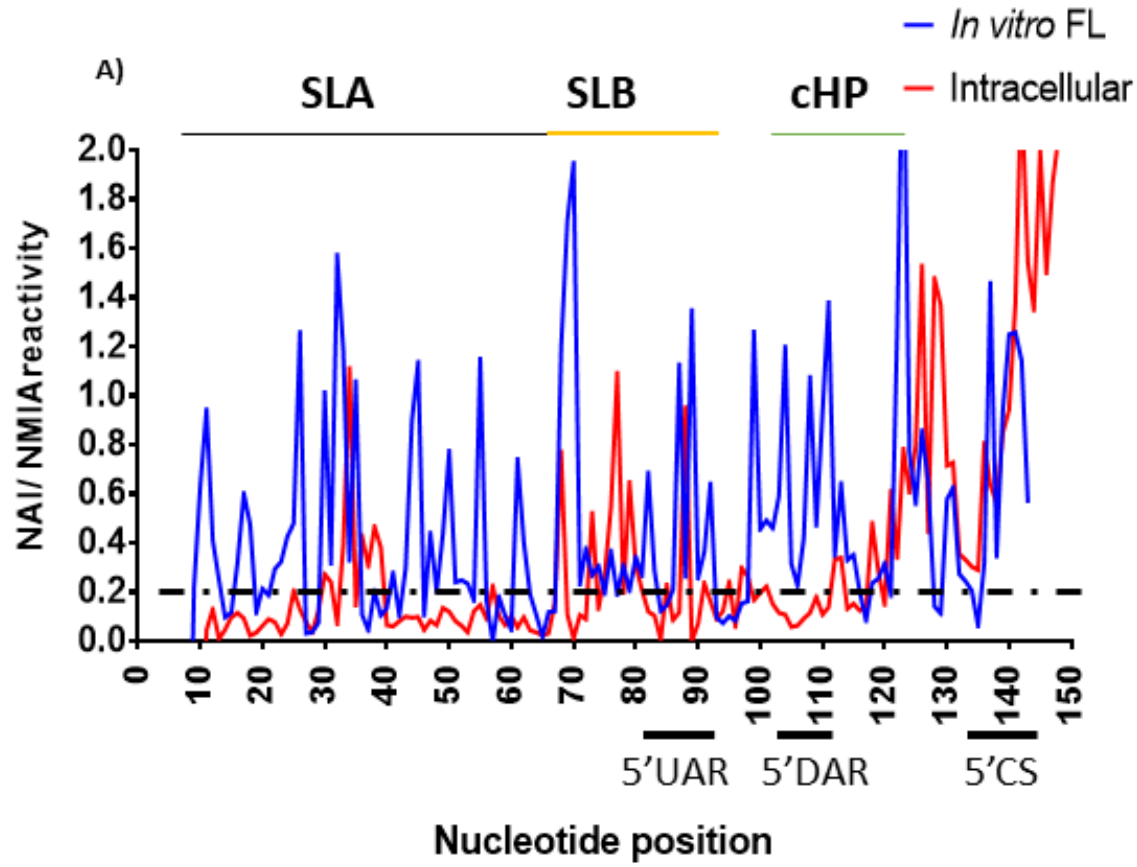
### 5.2.3 Intracellular SHAPE conducted on DENV infected cells of mammalian origin

To date and to our knowledge, the structure of the 5´ extremity of DENV genomic RNA has not been determined utilising intracellular SHAPE mapping following DENV infection, considering the mammalian host environment. We therefore determined the structure of DENV genomic RNA following the infection of BHK-21 cells with DENV-2. The determined viral RNA structure was compared to the RNA structural predictions made utilising *in vitro* SHAPE mapping of full-length DENV genomic RNA in the absence of *trans*-activating factors therefore acting to aid our understanding of DENV RNA structural requirements during active viral replication.

In figure 81, the normalised nucleotide reactivities to NAI/NMIA are compared, in the context of the full-length infectious DENV genomic transcript, before being overlaid onto the predicted structure of the DENV-2 5´extremity RNA at 37 °C.

As previously discussed in section 4.2.3, the addition of *trans*-activating factors *in vitro,* in the context of the full-length infectious DENV genome, does not result in DENV genomic cyclisation but is hypothesised to prepare the viral RNA for genomic cyclisation, and therefore viral replication, via the modulation of *cis*-acting RNA structures such as SLB and cHP. The UFS is unstable *in vitro* in the context of the DENV full-length genome and therefore the addition of *trans*-activating factors does not induce the linear to circular conformational switch.

Intracellular SHAPE mapping of DENV-2 RNA structure following DENV infection of BHK-21 cells has determined the viral genome to be predominantly in the linear conformation (Fig. 81). The linear conformation was determined due to the reactivity profile displayed by both the top-loop of SLB and the 5´CS as reactive to NAI and therefore single-stranded. The cHP structure is inherently unstable as indicated by nucleotides corresponding to the base of the stem predicted to be single-stranded and therefore the structure ceases to form in this context. The UFS stabilises in the context of DENV infection as indicated by reduced accessibility of nucleotides within this region, displaying a low reactivity profile and are therefore inferred to be base-paired indicating the presence of the structure in the context of active viral replication (Fig. 81).


A)
In vitro FL
Intracellular
SLA
SLB
cHP
NAI/ NMIAreactivity
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
5'UAR
5'DAR
5'CS
Nucleotide position

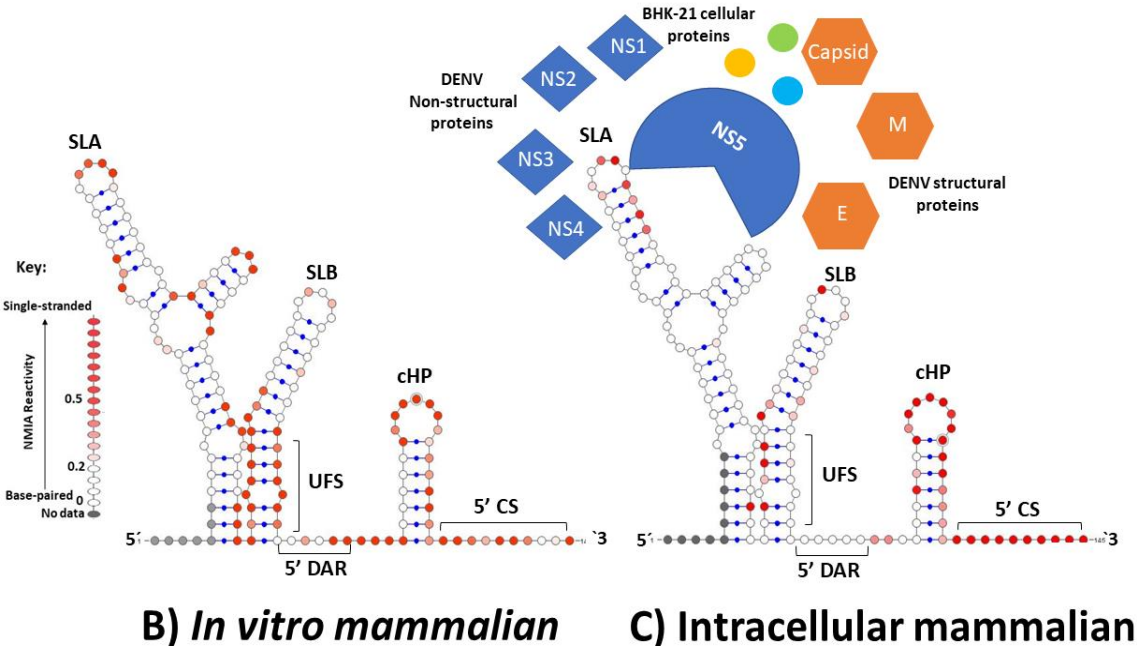

BHK-21 cellular proteins
NS1
NS2
DENV Non-structural proteins
NS3
NS4
NS5
Capsid
M
E
DENV structural proteins
SLA
SLB
cHP
UFS
5' CS
5' DAR
Key:
Single-stranded
NMIA Reactivity
0.5
0.2
Base-paired
No data
B) In vitro mammalian
C) Intracellular mammalian

**Figure 81 *In vitro* vs intracellular SHAPE mapping normalised NMIA/NAI reactivities corresponding to full-length DENV RNA at mammalian physiological temperature:**

**A)** *NMIA/NAI reactivities corresponding to the full-length (FL) infectious transcript as determined by in vitro SHAPE and as determined by intracellular SHAPE are shown in blue and red respectively. The dashed line corresponds to a NMIA reactivity of 0.2. Reactivities above this line are partially reactive and above 0.5 are highly reactive to NMIA and therefore single-stranded. Nucleotide positions of RNA structures and cyclisation sequences are indicated.* ***B)*** *Full-length infectious NMIA reactivities overlaid onto the first 145 nts of DENV-2 genomic RNA.* ***C)*** *Full-length replicon NAI reactivities overlaid onto the first 145 nts of DENV-2 genomic RNA. Overlays of NMIA/ NAI reactivities conducted using VARNA.*

### 5.2.4 Intracellular SHAPE conducted on DENV infected cells of insect origin

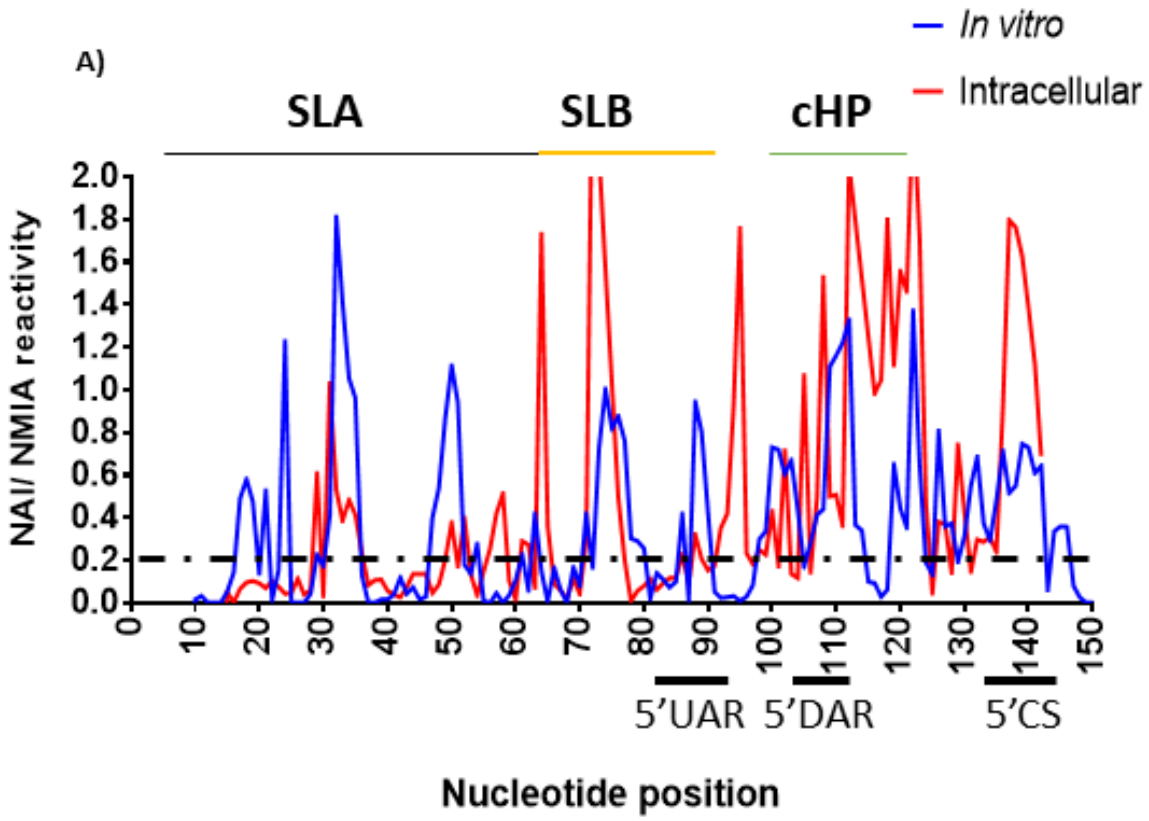
DENV cycles between mammalian and insect hosts and therefore must adapt rapidly to each environment to establish a robust infection. In contrast to humans, DENV infection of mosquito cells leads to a persistent infection without apparent cytopathic effect. Additionally, previous work has identified differential RNA sequence requirements for DENV replication between mosquito and human cells (Villordo and Gamarnik, 2013). However, the mechanistic basis of this host adaptation is not fully understood at the RNA structural level. We therefore sought to analyse the structure of the DENV-2 5´ extremity RNA at 28°C (insect physiological temperature). We determined the structure of DENV genomic RNA following the infection of C6/36 cells with DENV-2. The determined viral RNA structure was compared to the RNA structural predictions made utilising *in vitro* SHAPE mapping of full-length DENV genomic RNA in the absence of *trans*-activating factors to further underpin the mechanisms behind the linear to circular genomic switch during DENV infection of insect cells

In figure 82, the normalised nucleotide reactivities to NAI/NMIA are compared, in the context of the full-length infectious DENV genomic transcript, before being overlaid onto the predicted structure of the DENV-2 5´extremity RNA at 28 °C.

As previously discussed in section 4.2.6, our *in vitro* SHAPE mapping at insect physiological temperature in the context of the full-length genome in the absence of DENV-2 RdRp revealed the genomic structure of DENV-2 RNA to be predominantly in the linear conformation. Following the addition of the *trans*-activating factor DENV-2 RdRp, the genome remains in the linear conformation as was indicated by the same characteristic structural motifs indicative of the linear genomic conformation. The UFS also remained de-stabilized in the presence of DENV-2 RdRp. However, the cHP structure stabilised.

Taking these findings into account we mapped the structure of DENV genomic RNA following DENV infection of C6/36 cells therefore determining the structure of viral RNA during active viral replication. We determined the extracted DENV RNA to be predominantly in the linear conformation, as indicated by a NAI reactivity profile indicating the top-loop of SLB and the 5´CS to be single-stranded (Fig. 82).

During active viral replication the structure of cHP is unmaintained, as indicated by a NAI reactivity profile corresponding to the increased reactivity of nucleotides comprising the base of cHP indicating this region to be predominantly single-stranded and therefore the unfolding of the structure. During active viral replication the top of SLB is modified, indicating the partial melting of this structure. However, the UFS at the base of SLB appears to stabilise indicating the presence of this structure during active viral replication (Fig. 82).


A)
In vitro
Intracellular
SLA
SLB
cHP
NAI/ NMIA reactivity
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
5'UAR
5'DAR
5'CS
Nucleotide position

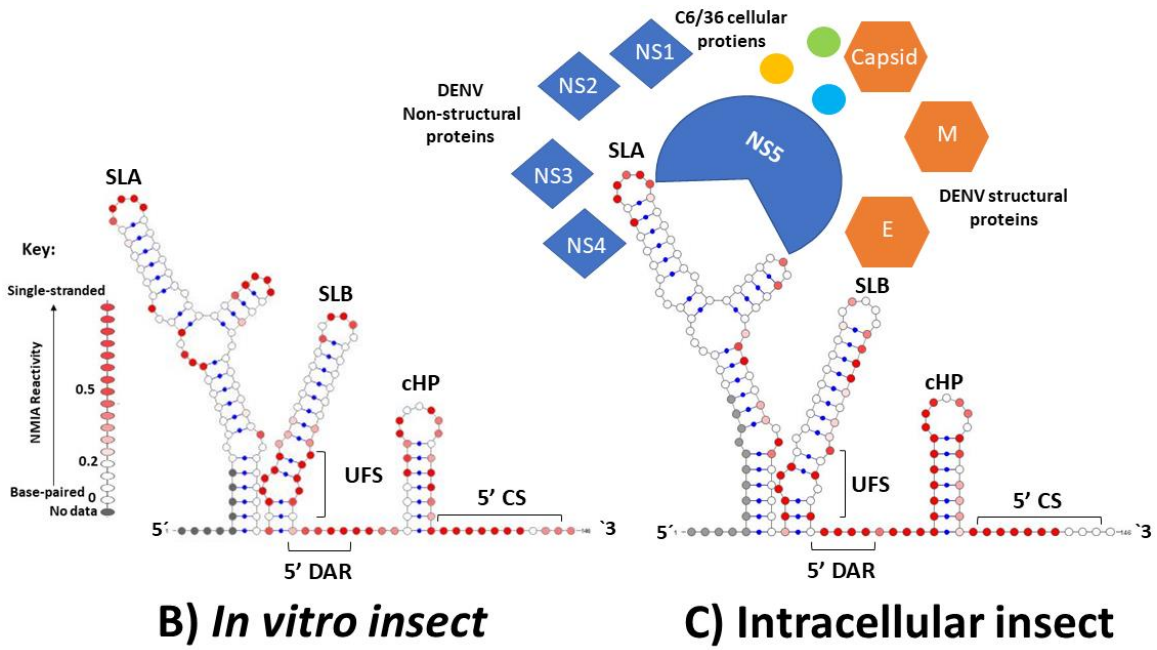

C6/36 cellular protiens
NS1
NS2
NS3
NS4
NS5
DENV Non-structural proteins
Capsid
M
E
DENV structural proteins
SLA
SLB
cHP
UFS
5' CS
5' DAR
Key:
Single-stranded
NMIA Reactivity
0.5
0.2
Base-paired 0
No data
B) In vitro insect
C) Intracellular insect

**Figure 82 *In vitro vs* intracellular SHAPE mapping normalised NMIA/NAI reactivities corresponding to full-length DENV RNA at insect physiological temperature:**

**A)** *NMIA/NAI reactivities corresponding to the full-length infectious transcript as determined by in vitro SHAPE and as determined by intracellular SHAPE are shown in blue and red respectively. The dashed line corresponds to a NMIA/NAI reactivity of 0.2. Reactivities above this line are partially reactive and above 0.5 are highly reactive and therefore single-stranded. Nucleotide positions of RNA structures and cyclisation sequences are indicated.* ***B)*** *Full-length infectious NMIA reactivities overlaid onto the first 145 nts of DENV-2 genomic RNA.* ***C)*** *Full-length replicon NAI reactivities overlaid onto the first 145 nts of DENV-2 genomic RNA. Experimentation conducted at 28°C – insect physiological temperature. Overlays of NMIA/ NAI reactivities conducted using VARNA.*

### 5.2.5 Intracellular SHAPE mapping of fractionated stably expressing DENV HEK293T cells - enriched for active viral replication complexes

It is well documented that DENV compartmentalises during its life-cycle, acting to modify host cell membrane structures, inducing the formation of double membrane vesicles to house the active viral replication complex (Gillespie *et al.*, 2010; Inoue and Tsai, 2013). Therefore, we extended our studies to analyse DENV genomic RNA structure in an environment conducive towards viral replication. Stably expressing DENV HEK293T replicon cells were fractionated, enabling the isolation of the membrane fraction. We conducted intracellular SHAPE on viral RNA extracted from the membrane fraction, enriched for active viral replication complexes in the aim of further deducing RNA structural requirements during viral replication and the biological mechanisms behind the linear to circular genomic switch.

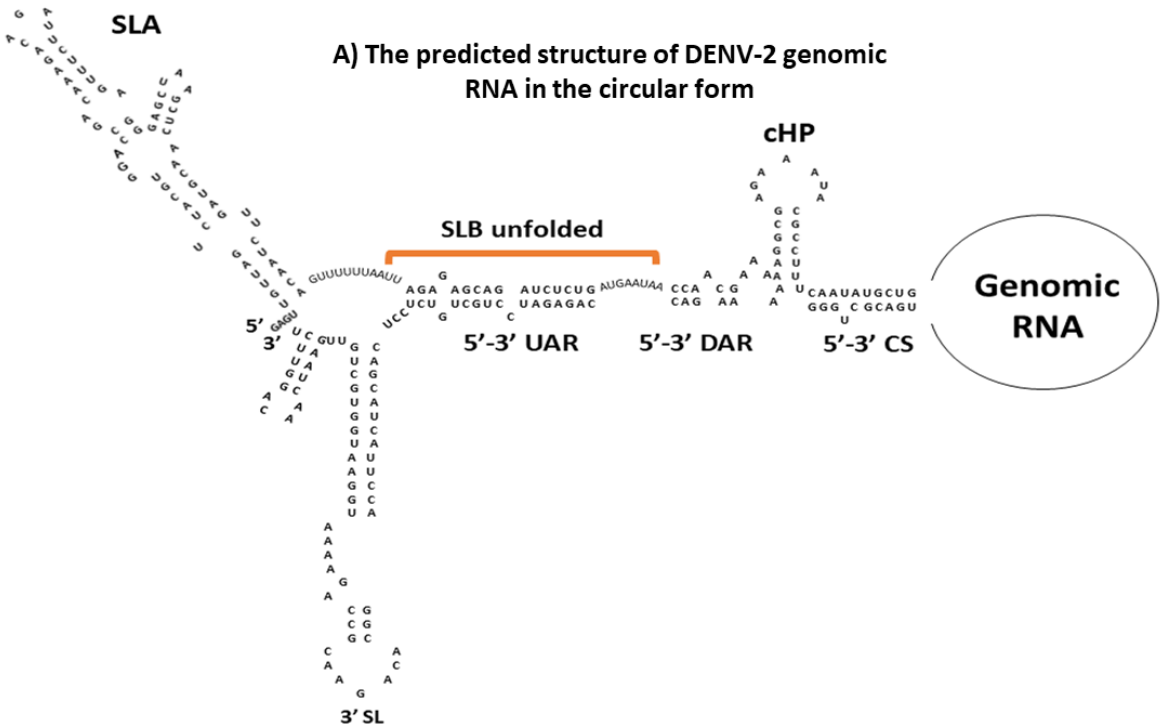
In figure 83, the intracellular SHAPE mapping of viral RNA extracted from the membrane fraction of stably expressing DENV HEK293T replicon cells is described and compared and contrasted with the intracellular SHAPE mapping of total RNA extracted from stably expressing DENV HEK293T cells. Firstly, the normalized nucleotide reactivities to NAI are displayed and compared to the reactivities determined from total DENV RNA structural analysis from DENV stably expressing HEK293T replicon cells before being overlaid onto the predicted structure of the DENV-2 5´extremity RNA at 37 °C.

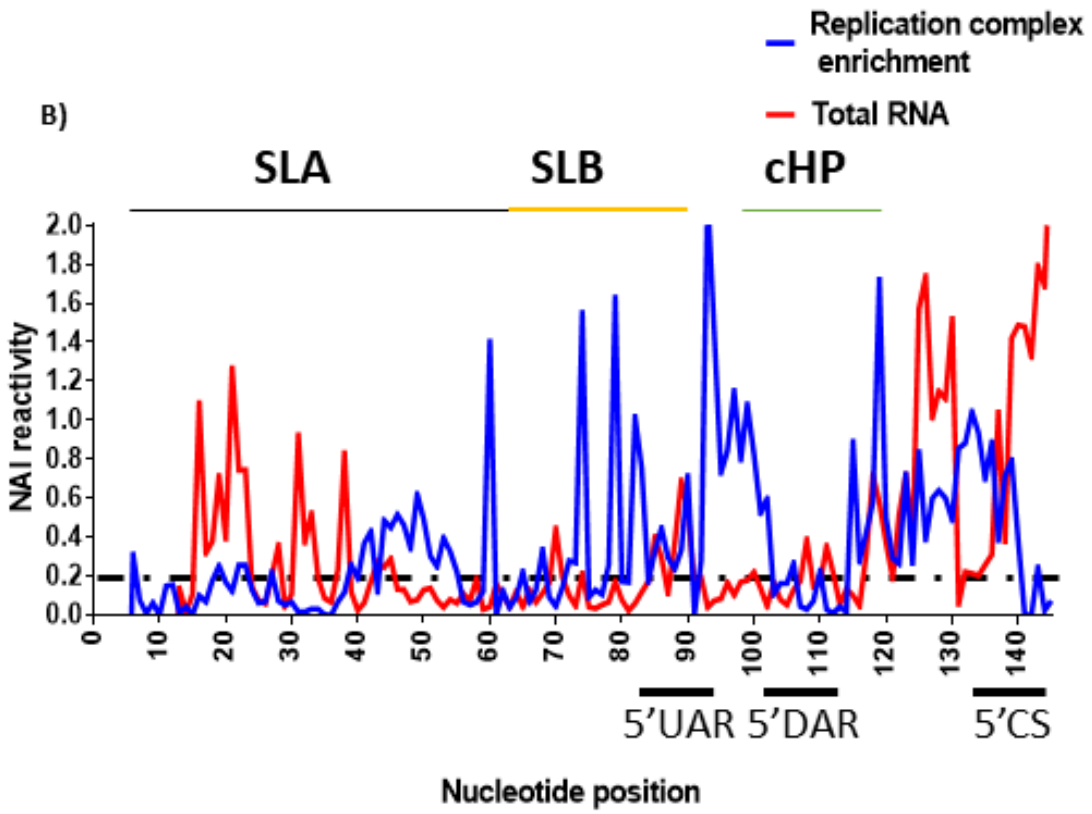
The intracellular SHAPE mapping of full-length DENV replicon RNA following total RNA extraction is further discussed in section 5.2.2. In brief, the linear conformation was determined. Moreover, the cHP structure is inherently unstable. Additionally, the UFS stabilizes indicating the presence of the structure in the context of active viral replication.

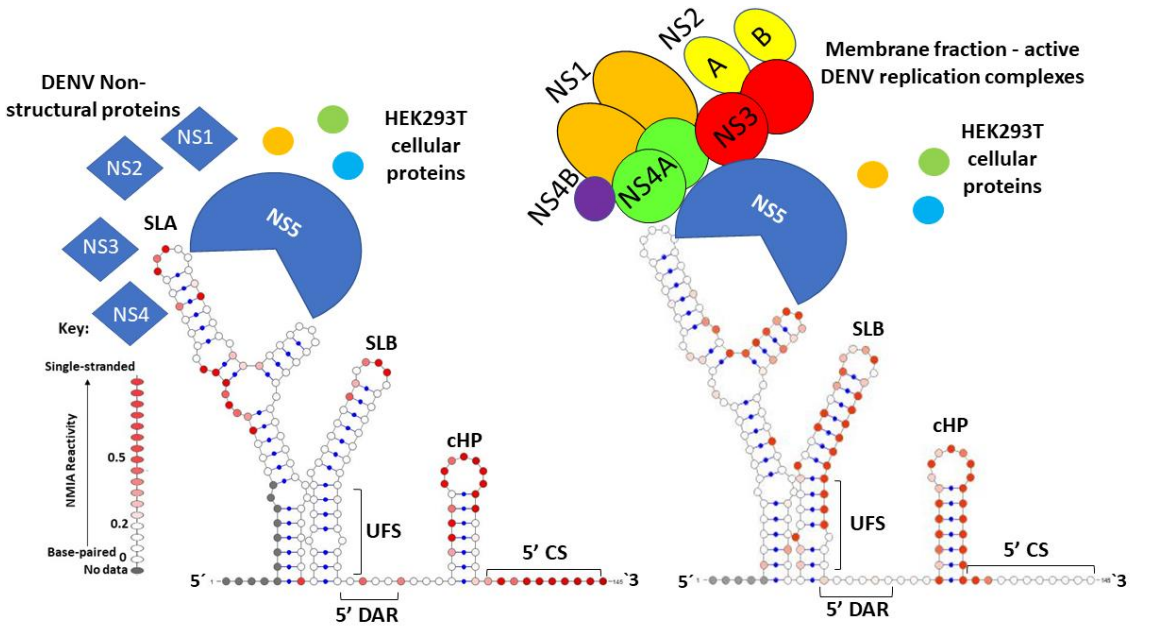
In contrast, intracellular SHAPE conducted on the isolated membrane fraction has predicted substantial genomic structural rearrangements in the 5´ extremity DENV genomic RNA region of interest (Fig. 83). The structure of SLB is unfolded, as indicated by an increased accessibility to NAI displayed by nucleotides corresponding to the stem of the structure indicating this region to be

predominantly single- stranded. Additionally, the UFS does not form as the structure of SLB itself ceases to exist. Moreover, the cHP does not form. The entire stem-loop is predicted to be single-stranded as indicated by the increased accessibility to NAI displayed by nucleotides corresponding to the stem-loop structure.

We predict, for the first time, to have mapped the DENV-2 RNA genome in the circular genomic conformation, the genomic conformation whereby negative sense RNA synthesis can occur. The structure of SLB is unfolded and the NAI reactivity profile corresponding to nucleotides comprising the 5´ CS and 5´DAR are reduced, inferring the cyclisation sequences to be base-paired. Therefore, the DENV-2 RNA structurally interrogated following membrane fraction isolation of stably expressing DENV HEK293T replicon cells, is predicted to form the circular DENV genomic conformation in an environment conducive for DENV-2 replication.






C) *Intracellular total RNA replicon* D) *Intracellular membrane fraction*

**Figure 83 Intracellular SHAPE mapping of total RNA extracted from stably expressing DENV HEK293T replicon cells vs Intracellular SHAPE mapping of total RNA extracted from the membrane fraction of stably expressing DENV HEK293T cells:**

***A)*** *The DENV-2 circular genomic form.* ***B)*** *Full-length DENV-2 total RNA extracted from stably expressing DENV HEK293T cells. Intracellular SHAPE NAI reactivities overlaid onto the first 145 nts of DENV-2 genomic RNA.* ***C)*** *Full-length DENV-2 RNA extracted from the membrane fraction of stably expressing DENV HEK293T replicon cells. NAI reactivities overlaid onto the first 145 nts of the DENV-2 genomic RNA.*

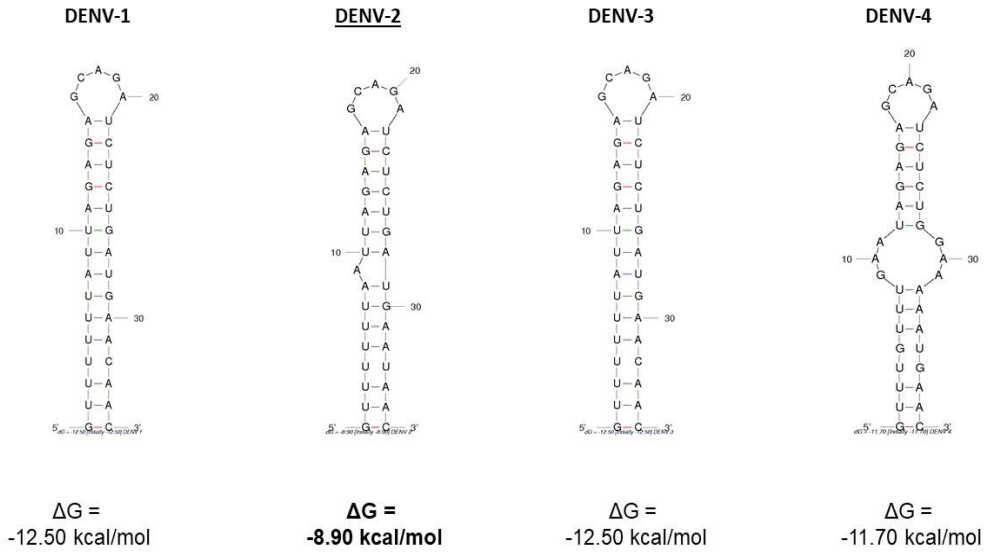
### 5.2.6 Pan-serotype thermodynamic analysis of the DENV-UFS

Our *in vitro* and intracellular structural studies have revealed that active replication is a fundamental requirement of the structural integrity of the DENV-2 UFS. In line with published work, the DENV-2 UFS is inherently unstable *in vitro* (Liu *et al.*, 2016). Consequently, we sought to interrogate these observations further, analysing the thermodynamics of the DENV-UFS across all four serotypes in the aim of analysing why the DENV-2 UFS is so structurally unstable.

Upon further *in silico* investigation, utilising Mfold, considering the thermodynamics of the DENV-2 UFS in comparison to UFS structures from other DENV serotypes, it is evidently the most thermodynamically unstable at both mammalian and insect temperatures in the absence of *trans*-activating factors (Zuker, 2003) (Figs. 84 and 85). We suggest this is due to the limited number of hydrogen bonds in the UFS stem region as evidenced by the comparative thermodynamic details described in figures 84 and 85.

Gibbs free energy (**δ**G) is used to describe the spontaneity of a process and thus predict the feasibility of chemical reactions. Reactions with a negative **δ**G occur spontaneously without energy input, reactions with a **δ**G of zero are at equilibrium and reactions with a positive **δ**G require energy to occur (GM, 2000). The DENV-2 UFS at both mammalian and insect physiological temperature has a **δ**G closest to zero, indicating more energy is required for the structure to form in comparison to the same structure utilised by the other serotypes. We predict this information is reflected in our *in vitro* experimental results and that the DENV-2 UFS ceases to form in this context due to its fundamental instability, which cannot be rescued by the addition of *trans*-activating factors during *in vitro* SHAPE experimentation. However, in the intracellular environment, *trans*-activating factors act to stabilise the DENV-2 UFS and it is only in this environment, during active replication, that the UFS stabilises.

**SLB 37 °C Mfold structural and thermodynamic predictions:**




**37 °C DENV-2**

| Structural element | δG | Information |
|---|---|---|
| External loop | 0.00 | 0 ss bases & 1 closing helices. |
| Stack | -2.20 | External closing pair is G $^{1}$-C $^{36}$ |
| Stack | -0.90 | External closing pair is U $^{2}$-A $^{35}$ |
| **Helix** | -3.10 | 3 base pairs. |
| Interior loop | 1.50 | External closing pair is U $^{3}$-A $^{34}$ |
| Stack | -0.90 | External closing pair is U $^{5}$-A $^{32}$ |
| Stack | -1.30 | External closing pair is U $^{6}$-A $^{31}$ |
| Stack | -1.00 | External closing pair is U $^{7}$-G $^{30}$ |
| **Helix** | -3.20 | 4 base pairs. |
| Bulge loop | 2.70 | External closing pair is A $^{8}$-U $^{29}$ |
| Stack | -1.30 | External closing pair is U $^{10}$-A $^{28}$ |
| Stack | -1.00 | External closing pair is U $^{11}$-G $^{27}$ |
| Stack | -2.10 | External closing pair is A $^{12}$-U $^{26}$ |
| Stack | -2.40 | External closing pair is G $^{13}$-C $^{25}$ |
| Stack | -2.10 | External closing pair is A $^{14}$-U $^{24}$ |
| Stack | -2.40 | External closing pair is G $^{15}$-C $^{23}$ |
| **Helix** | -11.30 | 7 base pairs. |
| Hairpin loop | 4.50 | Closing pair is A $^{16}$-U $^{22}$ |

**37 °C DENV-3**

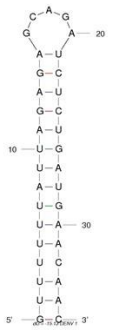
| Structural element | δG | Information |
|---|---|---|
| External loop | 0.00 | 0 ss bases & 1 closing helices. |
| Stack | -2.20 | External closing pair is G $^{1}$-C $^{35}$ |
| Stack | -0.90 | External closing pair is U $^{2}$-A $^{34}$ |
| **Helix** | -3.10 | 3 base pairs. |
| Interior loop | 1.70 | External closing pair is U $^{3}$-A $^{33}$ |
| Stack | -0.90 | External closing pair is U $^{5}$-A $^{31}$ |
| Stack | -1.30 | External closing pair is U $^{6}$-A $^{30}$ |
| Stack | -1.00 | External closing pair is U $^{7}$-G $^{29}$ |
| Stack | -1.10 | External closing pair is A $^{8}$-U $^{28}$ |
| Stack | -1.30 | External closing pair is U $^{9}$-A $^{27}$ |
| Stack | -1.00 | External closing pair is U $^{10}$-G $^{26}$ |
| Stack | -2.10 | External closing pair is A $^{11}$-U $^{25}$ |
| Stack | -2.40 | External closing pair is G $^{12}$-C $^{24}$ |
| Stack | -2.10 | External closing pair is A $^{13}$-U $^{23}$ |
| Stack | -2.40 | External closing pair is G $^{14}$-C $^{22}$ |
| **Helix** | -15.60 | 11 base pairs. |
| Hairpin loop | 4.50 | Closing pair is A $^{15}$-U $^{21}$ |

**Figure 84 Thermodynamic structural predictions of serotype specific DENV SLB at mammlian physiological temperature:**

*Thermodynamic analysis conducted using MFold* (Zuker, 2003).

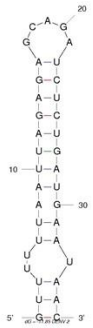
**SLB 28 °C Mfold structural and thermodynamic predictions:**

**DENV-1**



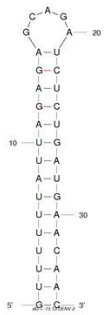
ΔG =
-15.13 kcal/mol

**DENV-2**



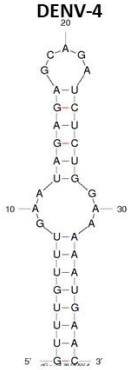
**ΔG =**
**-11.85 kcal/mol**

**DENV-3**



ΔG =
-15.13 kcal/mol

**DENV-4**



ΔG =
-14.76 kcal/mol

**28 °C DENV-2**

| Structural element | δG | Information |
|---|---|---|
| External loop | 0.00 | 0 ss bases & 1 closing helices. |
| Stack | -2.34 | External closing pair is G $^{1}$-C $^{36}$ |
| Stack | -1.07 | External closing pair is U $^{2}$-A $^{35}$ |
| **Helix** | -3.41 | 3 base pairs. |
| Interior loop | 1.22 | External closing pair is U $^{3}$-A $^{34}$ |
| Stack | -1.07 | External closing pair is U $^{6}$-A $^{32}$ |
| **Helix** | -1.07 | 2 base pairs. |
| Interior loop | 0.78 | External closing pair is U $^{7}$-A $^{31}$ |
| Stack | -1.04 | External closing pair is A $^{9}$-U $^{29}$ |
| Stack | -1.17 | External closing pair is U $^{10}$-A $^{28}$ |
| Stack | -0.89 | External closing pair is U $^{11}$-G $^{27}$ |
| Stack | -1.87 | External closing pair is A $^{12}$-U $^{26}$ |
| Stack | -2.62 | External closing pair is G $^{13}$-C $^{25}$ |
| Stack | -1.87 | External closing pair is A $^{14}$-U $^{24}$ |
| Stack | -2.62 | External closing pair is G $^{15}$-C $^{23}$ |
| **Helix** | -12.08 | 8 base pairs. |
| Hairpin loop | 2.71 | Closing pair is A $^{16}$-U $^{22}$ |

**28 °C DENV-3**

| Structural element | δG | Information |
|---|---|---|
| External loop | 0.00 | 0 ss bases & 1 closing helices. |
| Stack | -2.34 | External closing pair is G $^{1}$-C $^{35}$ |
| Stack | -1.07 | External closing pair is U $^{2}$-A $^{34}$ |
| **Helix** | -3.41 | 3 base pairs. |
| Interior loop | 0.78 | External closing pair is U $^{3}$-A $^{33}$ |
| Stack | -1.07 | External closing pair is U $^{5}$-A $^{31}$ |
| Stack | -1.17 | External closing pair is U $^{6}$-A $^{30}$ |
| Stack | -0.89 | External closing pair is U $^{7}$-G $^{29}$ |
| Stack | -1.04 | External closing pair is A $^{8}$-U $^{28}$ |
| Stack | -1.17 | External closing pair is U $^{9}$-A $^{27}$ |
| Stack | -0.89 | External closing pair is U $^{10}$-G $^{26}$ |
| Stack | -1.87 | External closing pair is A $^{11}$-U $^{25}$ |
| Stack | -2.62 | External closing pair is G $^{12}$-C $^{24}$ |
| Stack | -1.87 | External closing pair is A $^{13}$-U $^{23}$ |
| Stack | -2.62 | External closing pair is G $^{14}$-C $^{22}$ |
| **Helix** | -15.21 | 11 base pairs. |
| Hairpin loop | 2.71 | Closing pair is A $^{15}$-U $^{21}$ |

**Figure 85 Thermodynamic structural predictions of serotype specific DENV SLB at insect physiological temperature:**

*Thermodynamic analysis conducted using MFold* (Zuker, 2003).

## 5.3 Discussion

### 5.3.1 The linear genomic conformation and the DENV-2 UFS

It has been proposed that a highly conserved, conformation tuneable, *cis*-acting element termed the 5´UAR flanking stem (UFS), at the base of SLB, controls viral replication via regulating the binding of the RdRp to SLA. Stabilisation of the UFS is proposed to impair genomic cyclisation and therefore inhibit the transfer of DENV-2 RdRp to the 3´SL, promoting the linear genomic state and preventing viral replication. Genomic cyclisation unwinds SLB and consequently the UFS resulting in DENV-2 RdRp losing its affinity for SLA and promoting the transfer to the 3´SL, initiating viral replication. Nevertheless, it is suggested that the DENV-2 UFS is inherently unstable, as determined by *in vitro* SHAPE mapping and thermodynamically (section 5.2.6) (Liu *et al.*, 2016).

Our *in vitro* studies, discussed in Chapter 4, analysing the 5´ 500 nt transcript alongside the full-length DENV-2 transcript in the absence of *trans*-activating factors, agreed with published experiments that the DENV-2 UFS was unstable in this context. However, upon the addition of *trans*-activating factors, i.e stably expressing DENV HEK293T replicon cellular extract or DENV-2 RdRp, it's level of instability remained largely the same, thus indicating that neither the cellular extract or the RdRp alone could act to stabilise the UFS *in vitro*.

In contrast, intracellular SHAPE analysis conducted following DENV infection on total extracted RNA, in the context of active viral replication, at both insect and mammalian physiological temperature revealed a stable DENV-2 UFS for the first time. The intracellular environment is therefore predicted to be critical to the stabilisation of the DENV-2 UFS. We hypothesise that the linear genomic RNA form is the predominant structural form of the DENV genome in the host cell cytosol following total RNA extraction. It may be the case that the linear genomic conformation dominates within the host cell cytosol as a result of viral and host protein factor genomic RNA binding. For example, the interaction of DENV genomic RNA with host cell translational machinery may favour the linear conformation for efficient translation. Additionally, virus encapsidation is one of the most understudied steps of the Flavivirus replication cycle. For example, it is very much unclear how the capsid protein recruits the viral genome for assembly and

release (Samsa *et al.*, 2009). It may well be the case that a certain genomic conformation is required for efficient viral encapsidation and could consequently be the reason why the linear genomic form predominates during our intracellular studies. A specific genomic conformation has previously been shown to be required for retroviral encapsidation (Huthoff and Berkhout, 2001; D'Souza and Summers, 2004).To study genomic structural requirements for DENV replication, techniques such as Isothermal Titration Calorimetry or Nuclear Magnetic Resonance could be utilised to analyse the binding of DENV proteins critical to encapsidation, such as the Capsid protein, to either the linear or circular genomic form.

We further hypothesise that the fundamental instability of DENV-2 RNA is due to its increased replication kinetics and the possible interaction with host and viral protein factors is why the linear genomic conformation of the DENV genome is the predominant conformational form in the host cell cytosol. This may be why we have sampled the linear conformation both *in vitro* and intracellularly following total RNA extraction procedures (Sasmono *et al.*, 2015). The linear DENV genomic form may be required for viral encapsidation or translation procedures throughout the viral life-cycle. The UFS is only able to stabilise in the intracellular context due to the presence of *trans*-activating factors we were not able to replicate in an *in vitro* context. Furthermore, the presence of a stable UFS aids the binding of the DENV-2 RdRp to SLA, the first stage in viral replication (Liu *et al.*, 2016).

### 5.3.2 Intracellular structural interrogation of DENV cHP

The cHP sequence is conserved among all Flaviviruses and its presence and position are suggested to be critical to efficient translation ribosomal start-codon selection. It is suggested that the DENV-2 cHP selects the first AUG proportional to its stability and therefore, structural integrity, in a sequence independent manner and is required for efficient translation (Clyde and Harris, 2006). However, the conservation of cHP and evidence that mutation of the stem loop results in replicative disruption implicates this structure to have a role during viral replication, far more extensive than efficient translation of the Capsid protein. It is suggested that the overall topology of the stem-loop may contribute to the structural integrity of the 5´ end of DENV genomic RNA, acting to aid the binding of the viral replicase

complex or by stabilising the cyclisation interaction between the DENV cyclisation sequences and therefore encouraging the formation of the circular DENV genomic form, aiding viral replication (Clyde, Barrera and Harris, 2008).

Both *in vitro* and intracellularly, we have found the structure of cHP to vary considerably, with structural integrity determined as largely unmaintained throughout intracellular experimentation and in the context of membrane protein enrichment. We hypothesise this is due to the fact that in the intracellular environment cHP is required for a multitude of roles and interactions and therefore our SHAPE data is sampling these interactions all at one time. Consequently, we sought to analyse genomic RNA structure within an enriched membrane fraction therefore interrogating genomic RNA within active viral replication complexes. True structural integrity may not be required during active viral replication within the replication complex, in an environment where translation and therefore ribosomal start codon selection is not taking place. Therefore, using DENV-2 we predict there is no fundamental requirement for the maintenance of cHP structural integrity in the context of active viral replication.

### 5.3.3 Interrogation of DENV genomic RNA structure within the membrane fraction of stably expressing DENV HEK293T replicon cells

As previously discussed, viruses compartmentalise within the host cell as a fundamental requirement of their life-cycle, utilising the host cell to increase replication capability and manipulating the immune response. DENV is known to induce the formation of double-membrane vesicles from the ER, resulting in the formation of the viral replication complex. The site of viral replication is therefore restrained to the membrane of the ER (Salonen, Ahola and Kaariainen, 2005; Gillespie *et al.*, 2010; Inoue and Tsai, 2013). We chose to exploit this characteristic and, by the process of crude cellular fractionation, investigated the structure of viral RNA in the membrane protein fraction of stably expressing DENV HEK293T replicon cells, enriched for active viral replication complexes.

Our structural analysis of viral RNA within the membrane fraction has revealed, for the first time, the full-length replicon DENV-2 genomic RNA to be predominantly in the circular genomic conformation, as indicated by the complete loss of the

structural integrity of SLB and the reduced reactivity of the 5´ DAR and 5´CS – structural motifs for the determination of genomic conformation. Therefore, within the membrane fraction, enriched for active viral replication complexes, the predominant form of DENV-2 RNA is circular, the genomic conformation required for negative-sense RNA synthesis and DENV replication. This observation has led us to further disprove the theory of concatermisation or homoduplex formation of the DENV genome as neither of the forms could be present. This is evidenced as in the concatemeric conformation the SLB structure is maintained. Additionally, the homoduplex displays a degree of flexibility seen in the 5´ end of the genome that is not seen in the formation of the circular complex and which we have not observed.

Although we have successfully investigated the structure of DENV viral RNA within the membrane fraction of stably expressing DENV HEK293T cells and sampled the cyclic genomic form, further experimental repeats are required to strengthen our findings. To further broaden our experimental capabilities, experiments have yet to be optimised to interrogate viral RNA structure within the membrane fraction of both mammalian and insect cells following DENV-2 infection.

### 5.3.4 Limitations of Intracellular SHAPE mapping

Secondary and tertiary RNA structures are likely to be different within the intracellular environment, influenced by an infinite amount of *trans*-activating factors and the host cell environment, for example virus induced compartmentalisation, difficult to reproduce *in vitro*. Therefore, the ability to analyse RNA structure intracellularly is of critical importance and it is predicted that this field of chemical biology is poised to make a huge impact in the area of RNA structural biology (Kubota, Tran and Spitale, 2015). However, despite recent advances intracellular SHAPE mapping still carries inherent limitations.

A major limitation of intracellular SHAPE mapping following total RNA extraction is that the SHAPE reagent probing pattern obtained is from an average of structures present within the intracellular environment at one time and data analysis therefore infers the presence of only one structure within the cell, the average most dominant structural form of the RNA of interest. This is an inherent problem that cannot be over-come without specific targeting of sub-cellular locations and compartments.

Structured RNAs are present in various sub-cellular compartments – viral replication complexes being one example. To determine RNA structure intracellularly within these compartments, whilst maintaining their activity, probed cells must be lysed gently and fractionated. Cellular fractionation is crude, resulting in only the enrichment of certain proteins within each fraction, not complete purification. To overcome this, for example, an intracellular SHAPE reagent could be designed which localised solely to the viral replication complex acting to exclusively modify RNA within this location. Localising reagents are prevalent, for example Triphenylphosphonium salts have been shown to localise to the mitochondrial matrix and therefore could be utilised in the understanding of the mitochondrial transcriptome (Panja, Schu and Woodson, 2013). This characteristic could be translated to the intracellular RNA structural probing of viral replication complexes, replacing cellular fractionation, in the examination of viral RNA structural requirements in a truly selective context.

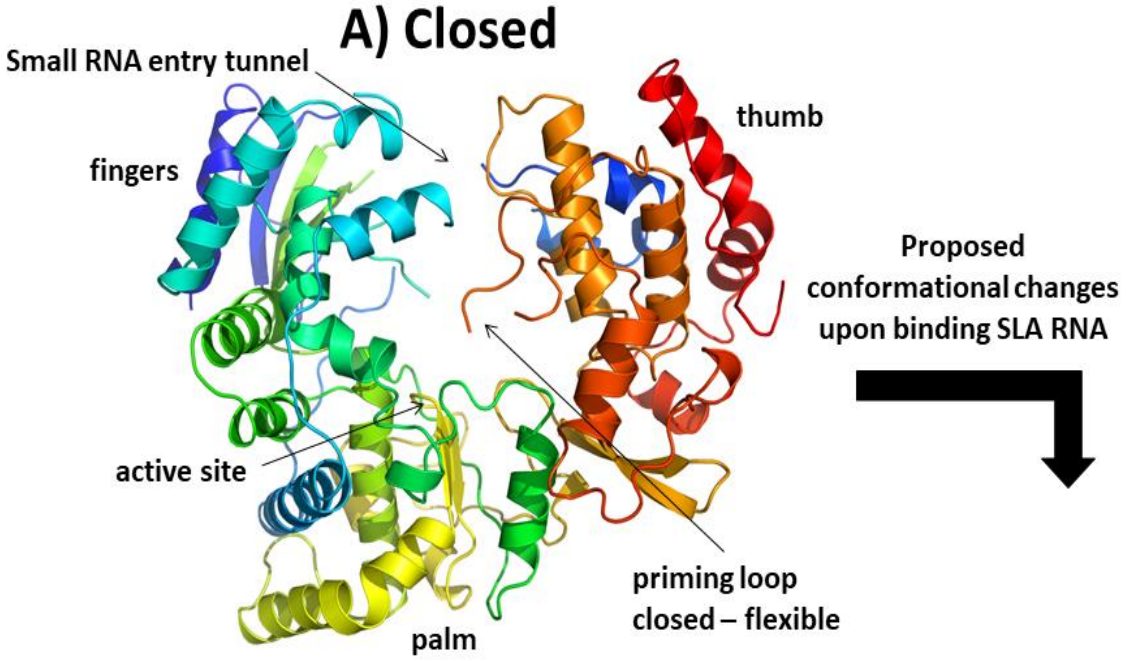
SHAPE probing reagents have been primarily designed to measure RNA secondary structure and therefore do not truly take into account the tertiary structure of RNA, for example, long-range base-pairing and the formation of pseudoknots and dumbbell structures, known to be important for regulating RNA function, exemplified by DENV replication and the genomic switch between the circular and linear form during viral replication. This limitation could be overcome by focusing on the development of RNA structure prediction programs that account for long-distance or three-dimensional RNA-RNA contacts. Recently, one such method that has been devised is crosslinking, ligation and sequencing of hybrids or CLASH (Kudla *et al.*, 2011). Using this technique RNA is irradiated causing the formation of RNA-RNA cross-links which, following sequencing, have identified RNA-RNA crosslinks that can be separated by great lengths of primary sequence. This technique could be utilised in intracellular mapping. Cultured cells could be irradiated and probed, fractionated if required, and RNA structure interrogated. This would act to bring intracellular structural probing into three dimensions, vital to truly understanding the influence of RNA structure.

# Chapter 6 General discussion

Flavivirus replication strategy demands further understanding at the molecular level. Identifying and interrogating the role of significant protein/RNA interactions during the DENV life-cycle could unlock potential strategies for anti-viral therapies, desperately needed in the treatment of DENV infection, translatable throughout the conserved Flavivirus family, and therefore the focus of our studies.

Chapter 3 presents the successful expression, purification, activity determination and first molecular structure of DENV-2 RdRp (unbound to RNA) – using X-ray crystallography solved by molecular replacement to a resolution of ~ 2.2 Å. The molecular structure revealed a typical right-hand viral polymerase in the RNA unbound “closed” pre-initiation state conformation with the priming loop extending towards the active site of the protein. This is a novel DENV serotype structure, increasing our knowledge of serotype specific DENV RdRps which bare approximately 70% amino-acid sequence homology, and improving our structural knowledge of the protein – critical for structure directed direct acting anti-viral development.

As discussed in detail in Chapter 1, DENV-2 RdRp binding to SLA RNA is a fundamental protein/RNA interaction within the DENV life-cycle. SLA is the 5´promoter of negative-sense RNA synthesis, acting to transfer the RdRp, via genomic cyclisation to the 3´SL to initiate negative-sense RNA synthesis. It is suggested that the “closed conformation” of the RdRp, prior to RNA binding, forms a pre-initiation conformation with an RNA tunnel within the protein that is too small to accommodate the growing double-stranded RNA intermediate during negative-sense RNA synthesis. In the “closed conformation” the priming loop of the polymerase extends towards the active site of the RdRp but must change conformation during extension to move away from the active site to stabilise and accommodate the nascent RNA chain (Fig. 86). Nevertheless, it is predicted that these structural movements occur, yet, for DENV, these changes have yet to be interrogated at the structural level and was therefore a major experimental focus.


A) Closed
Small RNA entry tunnel
fingers
thumb
Proposed
conformational changes
upon binding SLA RNA
active site
priming loop
closed – flexible
palm

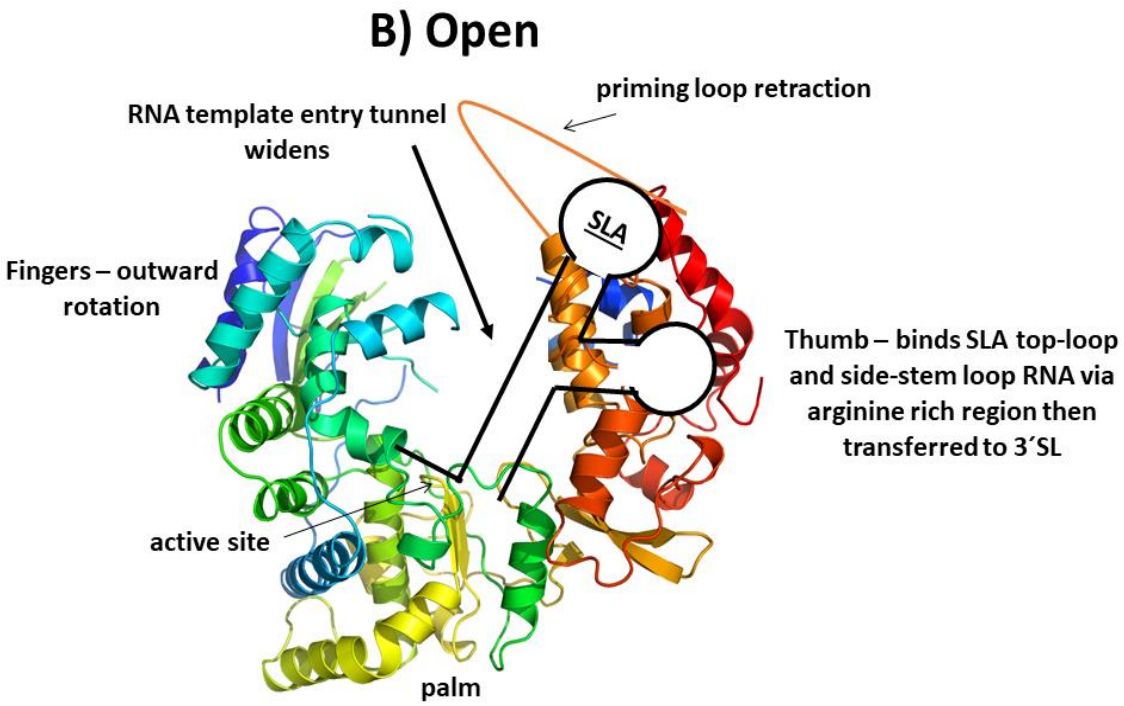

B) Open
priming loop retraction
RNA template entry tunnel
widens
SLA
Fingers – outward
rotation
Thumb – binds SLA top-loop
and side-stem loop RNA via
arginine rich region then
transferred to 3´SL
active site
palm

**Figure 86 Proposed structural model of the transition of the DENV RdRp from the "closed" to the "open" conformation during the initiation of DENV replication:**

*Schematic representation of the proposed DENV-RdRp conformational changes during replication initiation. The DENV-2 RdRp was solved by molecular replacement to a resolution of 2.2 Å.* ***A)*** *DENV-2 RdRp in the 'closed' pre- initiation conformation with the priming loop extended towards the active site of the protein and an RNA entry tunnel too small to accommodate the double-stranded replicative intermediate.* ***B)*** *By homology to another member of the Flaviviridae family HCV NS5B RdRp in the open elongation conformation we propose a model of the DENV-2 RdRp in the 'open' elongation conformation following binding to SLA promoter RNA via the thumb subdomain. The priming loop retracts away from the active site of the protein. The Fingers domain is suggested to outwardly rotate to widen the RNA entry tunnel which can then accommodate the double-stranded replicative intermediate.* (Appleby *et al.*, 2015).

Following determination of the crystal structure of RNA unbound DENV-2 RdRp SLA RNA co-crystallisation studies were attempted to determine the crystal structure of DENV-2 RdRp bound to SLA RNA utilising a range of co-crystallisation approaches. Despite multiple attempts, the resultant crystals either did not diffract or diffracted poorly with electron density corresponding to the "closed conformation" of the protein and therefore unbound to RNA.

The HCV RdRp NS5B is the only member of the *Flaviviridae* family to be successfully structurally interrogated in this manner bound to RNA, following years of work yielding initiation state 'open conformation' RdRp crystal structures which are being used to more accurately target HCV replication in the development of direct acting antivirals against HCV infection (Appleby *et al.*, 2015). Our considered co-crystallisation and soaking experiments were applicable, relevant approaches, however the protocols and strategies used in the determination of HCV RdRp replicative initiation structures could be translated to the study of DENV-2 RdRp with the aim of interrogating this vital stage in the DENV replicative cycle. Future work could include:

1. Select a DENV-2 genotype or mutant that displays increased RNA processivity and RNA binding kinetics.
2. Utilise a conformation stabilisation strategy whereby DENV NS5 resistance is produced in culture to a nucleoside analogue. Following sequencing, the resistant mutant could be cloned and expressed from a suitable expression construct.
3. Utilise nucleotide di-phosphates to stall DENV RdRp.
4. Optimise ionic concentrations, such as $Mn^{2+}$, to lower the Michaelis constant ($k_m$) of the initiating nucleotide - acting to stabilise the RdRp/RNA complex.
5. Perform X-ray crystallography and structurally interrogate a DENV RdRp domain which has stabilised a specific conformational state along the initiation pathway.

Despite the fact that this study was not able to obtain a high-resolution crystal structure of the DENV-2 RdRp in an initiation state 'open conformation', a high-resolution crystal structure of a DENV-2 RdRp domain was solved which, to our

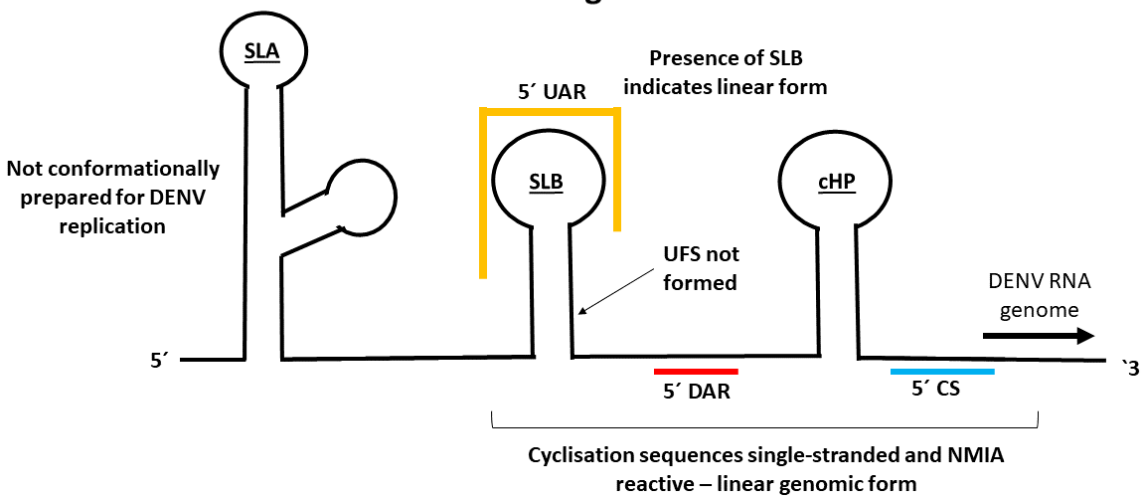
knowledge, has not been previously published and further increases our structural knowledge of the DENV replicative process.

Aside from the structural interrogation of the DENV-2 RdRp, much of our experimental investigations focussed on the mechanistic and structural function of the linear to circular genomic switch utilising full-length DENV genomic RNA templates in the context of DENV replication. We utilised novel *in vitro* SHAPE mapping approaches and optimised intracellular SHAPE mapping during active virus replication – a novel and previously unpublished approach for the analysis of DENV RNA structure. Unlike the technique of intracellular SHAPE utilised to interrogate the structure of non-replicating foot-and-mouth disease virus RNA transcripts or in virion intracellular SHAPE utilised, for example, to interrogate the structure of full-length poliovirus, we have developed a method that functions to interrogate viral RNA structure within the host cell during active viral replication, a true methodology development in the virology field.

Firstly, we utilised the technique of *in vitro* SHAPE mapping to examine any structural changes to the first 145 nts of DENV genomic RNA in the absence and presence of *trans*-activating factors at both mosquito and mammalian physiological temperature. Our findings, summarised in figure 87, subsequently led us to hypothesise that *in vitro:*

1. Full-length DENV genomic RNA does not form the circular genomic conformation *in vitro* and the presence of *trans*-activating factors is not sufficient to induce cyclisation.
2. The DENV-2 UFS does not form in the absence of intracellular conditions and *trans*-activating factors alone are not sufficient to stabilise its formation.
3. Viral and/or cellular *trans*-activating factors prepare DENV genomic RNA for cyclisation (a critical first stage in the initiation of DENV genome replication) via the destabilisation of downstream *cis*-acting RNA stem-loops.

**A) DENV RNA structure in the absence of *trans*-activating factors *in vitro*:**




**B) DENV RNA structure in the presence of *trans*-activating factors *in vitro*:**

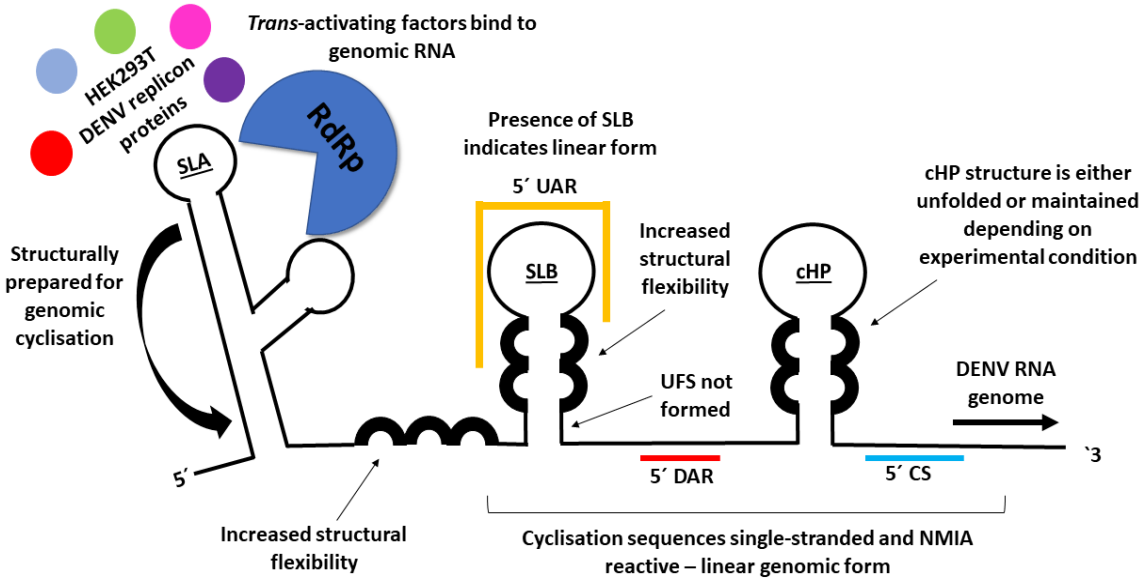


**Figure 87 A schematic of *in vitro* SHAPE findings at both mammalian and insect physiological temperature and proposed RNA structural hypothesis:**

***A)*** *A schematic representation of the structure of the first 145 nts of DENV genomic RNA in the absence of trans-activating factors. Without trans-activating factors the DENV genome is in the linear form, indicated by the presence of SLB and cyclisation sequences reactive to NMIA. In the absence of trans-activating factors the DENV genome is not conformationally prepared for replication and the UFS does not form.* ***B)*** *A schematic representation of the structure of the first 145 nts of DENV genomic RNA in the presence of trans-activating factors (DENV RdRp and stably expressing DENV HEK293T replicon cell extract). Interactions with trans-activating factors destabilise local RNA structures, a prerequisite for genome cyclisation. Consequently, we hypothesise that the trans-activating factor destabilisation of local RNA structures observed is an early mechanism involved in the initiation of negative-sense replication. Despite structurally preparing the genome, the addition of trans-activating factors is not sufficient to induce the linear to circular genomic switch or vice-versa. UAR = Upstream of AUG region, DAR = Downstream of AUG region and CS = cyclisation sequence.*

Following the interrogation of DENV RNA structure *in vitro* in the absence and presence of *trans*-activating factors, we optimised a novel technique utilising intracellular SHAPE mapping to investigate the structure of DENV RNA during active viral replication. We utilised intracellular SHAPE mapping to analyse the structure of total RNA following extraction from DENV-2 sub-genomic replicon stably expressing HEK293T cells or following DENV-2 infection of mammalian or mosquito cells. Our findings, summarised in figure 88, revealed:

1. Intracellular SHAPE for the determination of viral RNA structure during active viral replication is technically possible.
2. The linear genomic conformation is the dominant genomic conformation in the host cell.
3. The DENV-2 UFS is stabilised within the cell during active DENV replication.

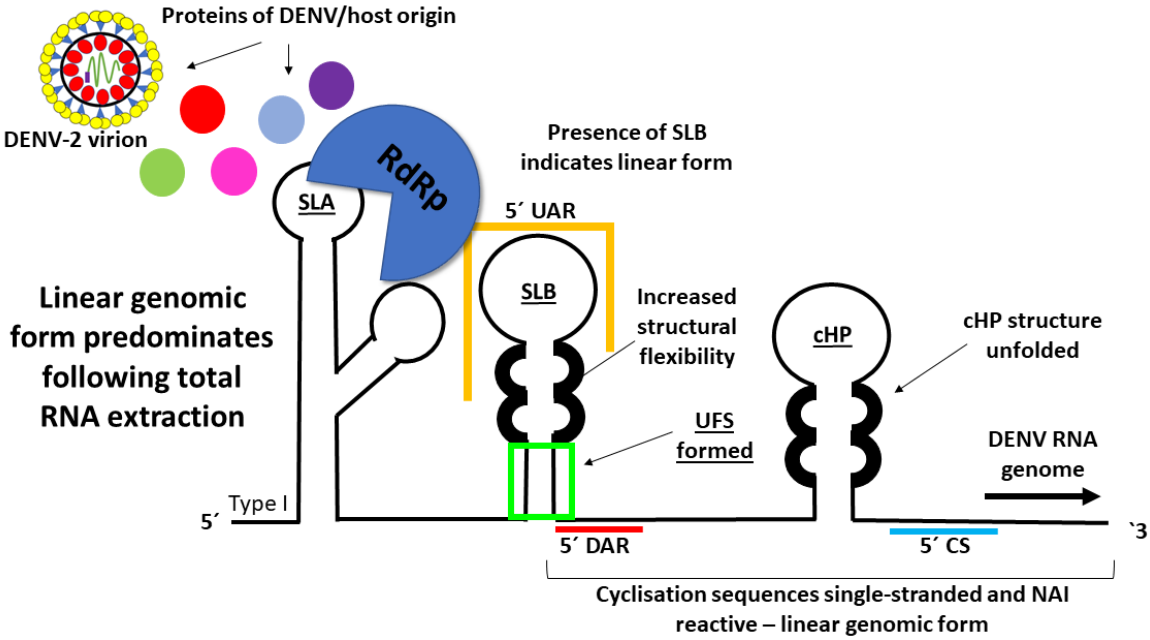



**Figure 88 Schematic representation of intracellular DENV RNA structure following total RNA extraction:**

*Intracellular SHAPE was utilised to interrogate the structure of DENV total RNA extracted from actively replicating stably expressing DENV HEK293T replicon cells or DENV-2 infected mammalian or insect cells. DENV genomic RNA remains in the linear conformation, which predominates in the host cell cytosol, indicated by the reactivity of the cyclisation sequences to NAI and the presence of the SLB structure. Importantly, the UFS structure is formed during active replication. UAR = Upstream of AUG region, DAR = Downstream of AUG region and CS = cyclisation sequence.*

Taking our optimised, validated technique of intracellular SHAPE one step further we sought to interrogate the structure of DENV genomic RNA within the replication complex during active virus replication, therefore analysing the structure of DENV genomic RNA in the true cellular location and environment of DENV genome replication. Utilising stably expressing DENV HEK293T replicon cells we interrogated DENV RNA structure within the isolated membrane fraction enriched for intact, active replication complexes. Our findings, summarised in figure 89, revealed:

1. DENV genomic RNA forms the circular conformation in an environment enriched for active replication complexes.

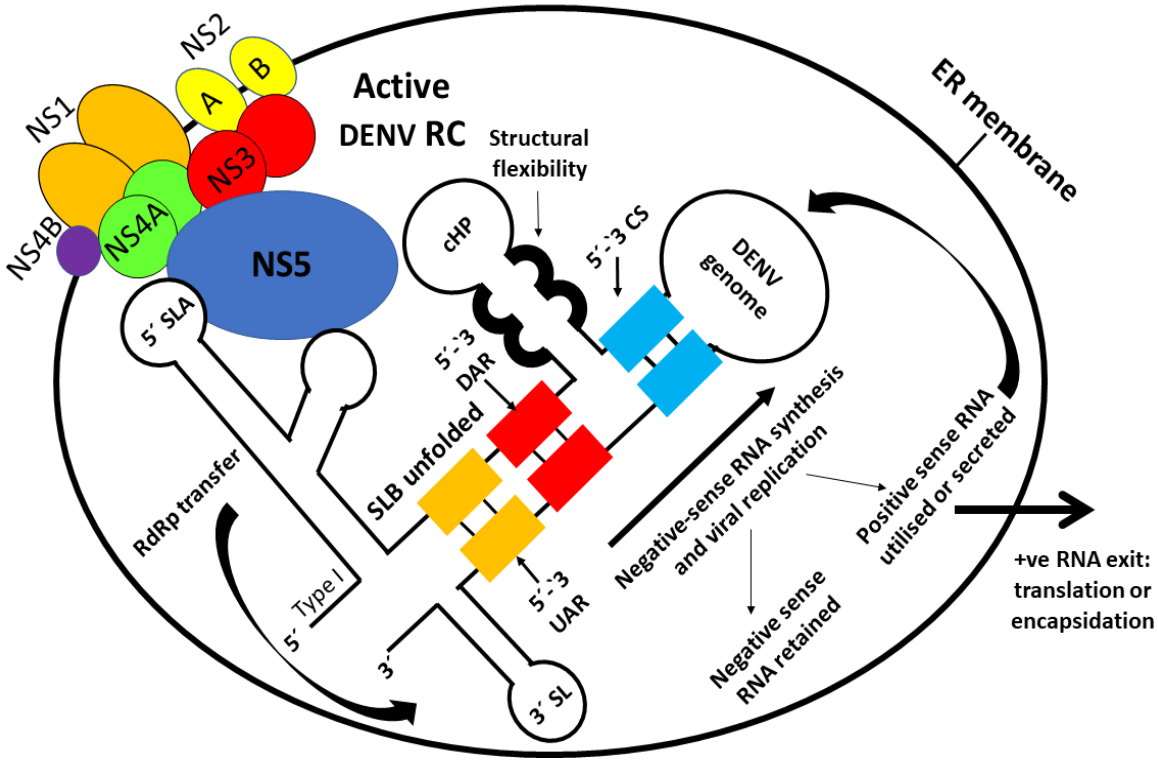



**Figure 89 Schematic representation of DENV genomic RNA structure within the replication complex:**

*A schematic representation of DENV RNA structure within the replication complex (RC) of actively replicating, stably expressing DENV HEK293T replicon cells. Within the replication complex DENV genomic RNA predominantly forms the circular conformation. The 5´ and 3´ cyclisation sequences hybridise, reducing their reactivity to NAI. The SLB structure is unfolded due to the hybridisation of the 5´-`3 UAR. The circular genomic conformation allows for negative-sense RNA synthesis and therefore viral replication within the replication complex. UAR = Upstream of AUG region, DAR = Downstream of AUG region and CS = cyclisation sequence. ER = Endoplasmic Reticulum, Type I= Type I cap.*

Consequently, taking all *in vitro* and intracellular experimental evidence into account, we therefore hypothesise that the linear form of the genome predominates in the host cell cytosol. The linear genomic conformation may therefore be required, for example, for genomic translation or encapsidation purposes and this conformation may be mediated via genomic RNA interaction with host and viral proteins within the cytosol. However, within the replication complex the circular form of the genome forms, and it is only in this cellular compartment, the true physiological environment of DENV genome replication, that full-length DENV genome cyclisation and hence initiation of viral replication occurs. This hypothesis is depicted in figure 90.

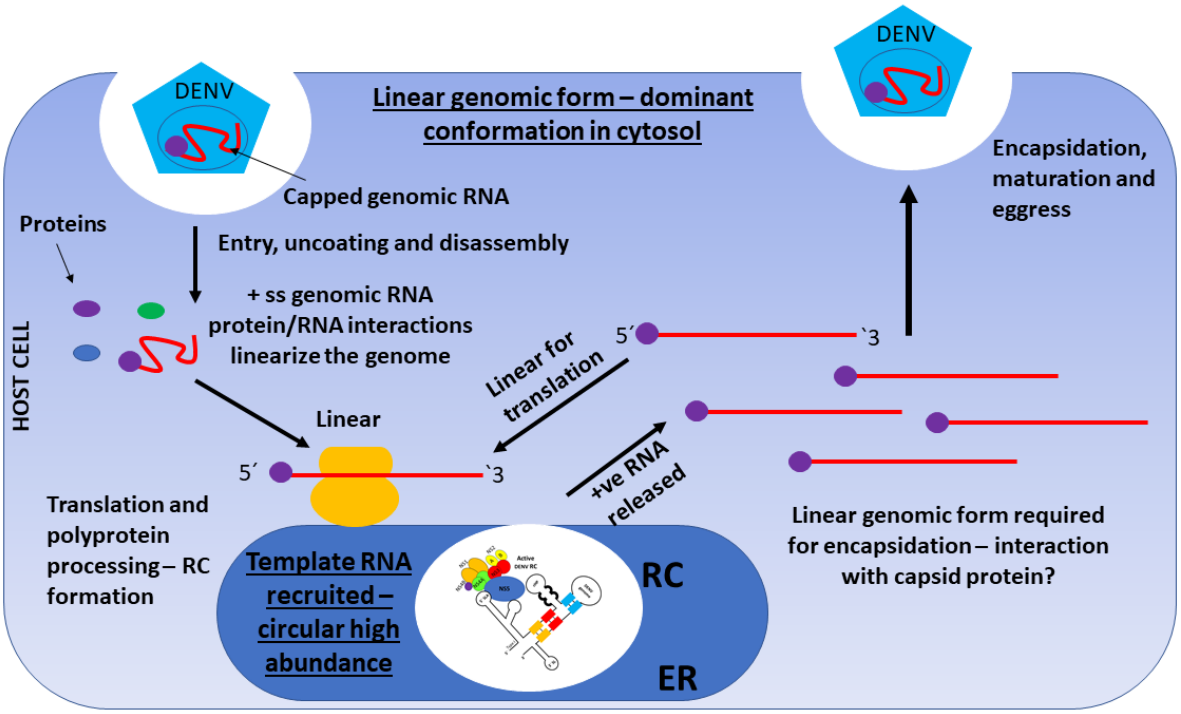



**Figure 90 Proposed DENV genomic conformation life-cycle model:**

*A hypothesis of the genomic conformations required throughout the DENV life-cycle, based on extensive in vitro and intracellular SHAPE mapping data. RC= Replication complex, ER = Endoplasmic reticulum, + SS = positive-sense single-stranded, + ve = positive-sense.*

The future direction of this study requires the analysis of DENV RNA structure within the membrane fraction of DENV infected mammalian and mosquito cells, rather than stably expressing DENV replicon HEK293T cells, to further investigate the true physiological environment of a DENV infection and genome replication. Following the novel methodologies optimised in the current study, this investigation would further elucidate DENV RNA structural requirements during virus infection which would likely be conserved across the Flavivirus family, increasing our knowledge of DENV infection and our understanding of Flavivirus molecular biology as a whole.

Finally, with the increasingly recognised regulatory role of dynamic RNA structure interactions across a range of viral and cellular systems, it is likely that the intracellular SHAPE methodologies optimised in the current study will be amenable and useful across a wide range of virological and cellular systems. Critically, this action would significantly increase our knowledge at the molecular level of a range of viruses that pose a significant threat to public health, desperately required in the development of direct acting anti-viral therapies.

## Bibliography

Achee, N. L. *et al.* (2015) ‘A Critical Assessment of Vector Control for Dengue Prevention’, *PLOS Neglected Tropical Diseases*. Public Library of Science, 9(5), p. e0003655. ava: https://doi.org/10.1371/journal.pntd.0003655.

Aguiar, M., Stollenwerk, N. and Halstead, S. B. (2018) ‘The risks behind Dengvaxia recommendation’, *The Lancet Infectious Diseases*. Elsevier, 16(8), pp. 882–883. doi: 10.1016/S1473-3099(16)30168-2.

Alla, P. and Alexander, S. (2014) ‘Structure and function of pseudoknots involved in gene expression control’, *Wiley Interdisciplinary Reviews: RNA*. Wiley-Blackwell, 5(6), pp. 803–822. doi: 10.1002/wrna.1247.

Allison, S. L. *et al.* (1995) ‘Oligomeric rearrangement of tick-borne encephalitis virus envelope proteins induced by an acidic pH.’, *Journal of Virology* , 69(2), pp. 695–700. http://jvi.asm.org/content/69/2/695.abstract.

Alvarez, Lodeiro, M. F., *et al.* (2005) ‘Long-Range RNA-RNA Interactions Circularize the Dengue Virus Genome’, *Journal of Virology*. American Society for Microbiology, 79(11), pp. 6631–6643. doi: 10.1128/JVI.79.11.6631-6643.2005.

Alvarez, Laura De Lella Ezcurra, A., *et al.* (2005) *Role of RNA structures present at the 3'UTR of dengue virus on translation, RNA synthesis, and viral replication*, *Virology*. doi: 10.1016/j.virol.2005.06.009.

Alvarez, M. *et al.* (2006) ‘Dengue hemorrhagic Fever caused by sequential dengue 1-3 virus infections over a long time interval: Havana epidemic, 2001-2002.’, *The American journal of tropical medicine and hygiene*. United States, 75(6), pp. 1113–1117.

Andino, R., Rieckhof, G. E. and Baltimore, D. (1990) ‘A functional ribonucleoprotein complex forms around the 5’ end of poliovirus RNA.’, *Cell*. United States, 63(2), pp. 369–380.

Angulo, J. *et al.* (2016) ‘LOOP IIId of the HCV IRES is essential for the structural rearrangement of the 40S-HCV IRES complex.’, *Nucleic acids research*. England, 44(3), pp. 1309–1325. doi: 10.1093/nar/gkv1325.

Antal, M. *et al.* (2002) ‘Analysis of the structure of human telomerase RNA in

vivo', *Nucleic Acids Research*. Oxford, UK: Oxford University Press, 30(4), pp. 912–920. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC100349/.

Appleby, T. C. *et al.* (2015) 'Structural basis for RNA replication by the hepatitis C virus polymerase', *Science*, 347(6223), p. 771 LP-775. http://science.sciencemag.org/content/347/6223/771.abstract.

Avirutnan, P. *et al.* (2007) 'Secreted NS1 of Dengue Virus Attaches to the Surface of Cells via Interactions with Heparan Sulfate and Chondroitin Sulfate E', *PLOS Pathogens*. Public Library of Science, 3(11), p. e183. https://doi.org/10.1371/journal.ppat.0030183.

Balsitis, S. J. *et al.* (2010) 'Lethal Antibody Enhancement of Dengue Disease in Mice Is Prevented by Fc Modification', *PLOS Pathogens*. Public Library of Science, 6(2), p. e1000790. https://doi.org/10.1371/journal.ppat.1000790.

Beltramello, M. *et al.* (2010) *The Human Immune Response to Dengue Virus Is Dominated by Highly Cross-Reactive Antibodies Endowed with Neutralizing and Enhancing Activity*, *Cell host & microbe*. doi: 10.1016/j.chom.2010.08.007.

Bente, D. A. and Rico-Hesse, R. (2006) 'Models of dengue virus infection', *Drug Discovery Today: Disease Models*, 3(1), pp. 97–103. doi: https://doi.org/10.1016/j.ddmod.2006.03.014.

Bergfors, T. (2003) 'Seeds to crystals.', *Journal of structural biology*. United States, 142(1), pp. 66–76.

Bhatt, S. *et al.* (2013) 'The global distribution and burden of dengue', *Nature*, 496(7446), pp. 504–507. doi: 10.1038/nature12060.

Bian, G. *et al.* (2010) 'The Endosymbiotic Bacterium Wolbachia Induces Resistance to Dengue Virus in Aedes aegypti', *PLOS Pathogens*. Public Library of Science, 6(4), p. e1000833. https://doi.org/10.1371/journal.ppat.1000833.

Blitvich, B. J. and Firth, A. E. (2015) 'Insect-specific flaviviruses: a systematic review of their discovery, host range, mode of transmission, superinfection exclusion potential and genomic organization', *Viruses*, 7(4), p. 1927—1959. doi: 10.3390/v7041927.

de Borba, L. *et al.* (2015) 'Overlapping Local and Long-Range RNA-RNA

Interactions Modulate Dengue Virus Genome Cyclization and Replication', *Journal of Virology* , 89(6), pp. 3430–3437. doi: 10.1128/JVI.02677-14.

Bradrick, S. S., Walters, R. W. and Gromeier, M. (2006) 'The hepatitis C virus 3'-untranslated region or a poly(A) tract promote efficient translation subsequent to the initiation phase', *Nucleic Acids Research*, 34(4), p. 1293—1303. doi: 10.1093/nar/gkl019.

Brathwaite Dick, O. *et al.* (2012) 'The History of Dengue Outbreaks in the Americas', *The American Journal of Tropical Medicine and Hygiene*. The American Society of Tropical Medicine and Hygiene, 87(4), pp. 584–593. doi: 10.4269/ajtmh.2012.11-0770.

Burrill, C. P. *et al.* (2013) 'Global RNA Structure Analysis of Poliovirus Identifies a Conserved RNA Structure Involved in Viral Replication and Infectivity', *Journal of Virology* , 87(21), pp. 11670–11683. doi: 10.1128/JVI.01560-13.

Callister, W. and Rethwisch, D. (2011) 'Materials Science and Engineering'.

Capeding, M. R. *et al.* (2018) 'Clinical efficacy and safety of a novel tetravalent dengue vaccine in healthy children in Asia: a phase 3, randomised, observer-masked, placebo-controlled trial', *The Lancet*. Elsevier, 384(9951), pp. 1358–1365. doi: 10.1016/S0140-6736(14)61060-6.

Caragata, E. P. *et al.* (2013) 'Dietary Cholesterol Modulates Pathogen Blocking by Wolbachia', *PLOS Pathogens*. Public Library of Science, 9(6), p. e1003459. https://doi.org/10.1371/journal.ppat.1003459.

Carvalho, D. O. *et al.* (2015) 'Suppression of a Field Population of Aedes aegypti in Brazil by Sustained Release of Transgenic Male Mosquitoes.', *PLoS neglected tropical diseases*. United States, 9(7), p. e0003864. doi: 10.1371/journal.pntd.0003864.

Castro, C. *et al.* (2009) 'Nucleic acid polymerases use a general acid for nucleotidyl transfer.', *Nature structural & molecular biology*. United States, 16(2), pp. 212–218. doi: 10.1038/nsmb.1540.

CDC (2018) *Dengue Epidemiology*. https://www.cdc.gov/dengue/epidemiology/index.html (Accessed: 7 March 2018).

Chambers, T. J. *et al.* (1990) ‘Flavivirus Genome Organization, Expression, and Replication’, *Annual Review of Microbiology*. Annual Reviews, 44(1), pp. 649–688. doi: 10.1146/annurev.mi.44.100190.003245.

Chan, M. and Johansson, M. A. (2012) ‘The Incubation Periods of Dengue Viruses’, *PLoS ONE*. Edited by N. Vasilakis. San Francisco, USA: Public Library of Science, 7(11), p. e50972. doi: 10.1371/journal.pone.0050972.

Chung, K. M. *et al.* (2006) ‘West Nile virus nonstructural protein NS1 inhibits complement activation by binding the regulatory protein factor H’, *Proceedings of the National Academy of Sciences*, 103(50), p. 19111 LP-19116. http://www.pnas.org/content/103/50/19111.abstract.

Clyde, K., Barrera, J. and Harris, E. (2008) ‘The capsid-coding region hairpin element (cHP) is a critical determinant of dengue virus and West Nile virus RNA synthesis’, *Virology*, 379(2), pp. 314–323. doi: 10.1016/j.virol.2008.06.034.

Clyde, K. and Harris, E. (2006) ‘RNA Secondary Structure in the Coding Region of Dengue Virus Type 2 Directs Translation Start Codon Selection and Is Required for Viral Replication’, *Journal of Virology*. American Society for Microbiology, 80(5), pp. 2170–2182. doi: 10.1128/JVI.80.5.2170-2182.2006.

Conway, M. J. *et al.* (2014) ‘Mosquito Saliva Serine Protease Enhances Dissemination of Dengue Virus into the Mammalian Host’, *Journal of Virology* , 88(1), pp. 164–175. doi: 10.1128/JVI.02235-13.

Creative Biostructure (2018) *Custom Size-exclusion Chromatography*. https://www.creative-biostructure.com/custom-size-exclusion-chromatography-service-259.htm (Accessed: 22 February 2018).

D’Souza, V. and Summers, M. F. (2004) ‘Structural basis for packaging the dimeric genome of Moloney murine leukaemia virus.’, *Nature*. England, 431(7008), pp. 586–590. doi: 10.1038/nature02944.

Deigan, K. E. *et al.* (2009) ‘Accurate SHAPE-directed RNA structure determination.’, *Proceedings of the National Academy of Sciences of the United States of America*. United States, 106(1), pp. 97–102. doi: 10.1073/pnas.0806929106.

Dejnirattisai, W. *et al.* (2015) ‘A new class of highly potent, broadly neutralizing antibodies isolated from viremic patients infected with dengue virus’, *Nature immunology*, 16(2), p. 170—177. doi: 10.1038/ni.3058.

Diaz-Toledano, R., Lozano, G. and Martinez-Salas, E. (2017) ‘In-cell SHAPE uncovers dynamic interactions between the untranslated regions of the foot-and-mouth disease virus RNA’, *Nucleic Acids Research*. Oxford University Press, 45(3), pp. 1416–1432. doi: 10.1093/nar/gkw795.

Dong, H. *et al.* (2012) ‘2’-O methylation of internal adenosine by flavivirus NS5 methyltransferase’, *PLoS pathogens*, 8(4), p. e1002642. doi: 10.1371/journal.ppat.1002642.

Dreher, T. (1999) *Functions of the 3’-untranslated regions of positive strand RNA viral genomes*, *Annual review of phytopathology*. doi: 10.1146/annurev.phyto.37.1.151.

Du, Y. (2015) ‘Fluorescence Polarization Assay to Quantify Protein-Protein Interactions in an HTS Format BT - Protein-Protein Interactions: Methods and Applications’, in Meyerkord, C. L. and Fu, H. (eds). New York, NY: Springer New York, pp. 529–544. doi: 10.1007/978-1-4939-2425-7_35.

DuBridge, R. B. *et al.* (1987) ‘Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system.’, *Molecular and Cellular Biology*, 7(1), pp. 379–387. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC365079/.

Dyer, O. (2017) ‘Philippines halts dengue immunisation campaign owing to safety risk’, *BMJ*. BMJ Publishing Group Ltd, 359. doi: 10.1136/bmj.j5759.

Emsley, P. and Cowtan, K. (2004) ‘Coot: model-building tools for molecular graphics.’, *Acta crystallographica. Section D, Biological crystallography*. United States, 60(Pt 12 Pt 1), pp. 2126–2132. doi: 10.1107/S0907444904019158.

European Centre for Disease Prevention and Control (2018) *Infectious diseases and public health*. https://ecdc.europa.eu/en/climate-change/climate-change-europe/vector-borne-diseases (Accessed: 21 February 2018).

Fairbrother, A. and Yuill, T. M. (1987) ‘Experimental and horizontal transmission of Modoc virus in Deer mice (Peromyscus Maniculatus)’, *Journal of Wildlife*

*Diseases*. Wildlife Disease Association, 23(2), pp. 179–185. doi: 10.7589/0090-3558-23.2.179.

Feng, C., Chan, D. and Spitale, R. C. (2017) 'Assaying RNA Structure Inside Living Cells with SHAPE BT - mRNA Processing: Methods and Protocols', in Shi, Y. (ed.). New York, NY: Springer New York, pp. 247–256. doi: 10.1007/978-1-4939-7204-3_18.

Filomatori, C. V *et al.* (2006) 'A 5′ RNA element promotes dengue virus RNA synthesis on a circular genome', *Genes & Development* , 20(16), pp. 2238–2249. doi: 10.1101/gad.1444206.

Filomatori, C. V *et al.* (2011) 'RNA Sequences and Structures Required for the Recruitment and Activity of the Dengue Virus Polymerase', *Journal of Biological Chemistry* , 286(9), pp. 6929–6939. doi: 10.1074/jbc.M110.162289.

GE Healthcare Life Sciences (2018) *Protein Purification Methods*. https://www.gelifesciences.com/solutions/protein-research/knowledge-center/protein-purification-methods/size-exclusion-chromatography (Accessed: 22 February 2018).

Gibbons, R. V *et al.* (2012) 'Dengue and US Military Operations from Spanish–American War through Today', *Emerging Infectious Disease journal*, 18(4), p. 623. doi: 10.3201/eid1804.110134.

Gilles, H. M. (1978) 'A World Geography of Human Diseases', *Journal of the Royal Society of Medicine*, 71(12), p. 939—939. http://europepmc.org/articles/PMC1436329.

Gillespie, L. K. *et al.* (2010) 'The endoplasmic reticulum provides the membrane platform for biogenesis of the flavivirus replication complex.', *Journal of virology*. United States, 84(20), pp. 10438–10447. doi: 10.1128/JVI.00986-10.

Gjenero-Margan, I. *et al.* (2011) *Autochthonous Dengue fever in Croatia, August-September 2010*, *Euro surveillance : bulletin européen sur les maladies transmissibles = European communicable disease bulletin*.

GM, C. (2000) *The Cell: A Molecular Approach*. 2 nd.

Goubau, D. *et al.* (2014) 'Antiviral immunity via RIG-I-mediated recognition of

RNA bearing 5'-diphosphates.', *Nature*. England, 514(7522), pp. 372–375. doi: 10.1038/nature13590.

Greenpeace (2018) *Annotated map of Southeast Asia*. http://www.greenpeace.org/sweden/Global/sweden/p2/skog/image/2006/map-of-paradise-forests.jpg (Accessed: 30 April 2018).

Gubler, D. J. (1998) 'Dengue and Dengue Hemorrhagic Fever', *Clinical Microbiology Reviews* , 11(3), pp. 480–496. http://cmr.asm.org/content/11/3/480.abstract.

Gubler, D. J. and Clark, G. G. (1995) 'Dengue/Dengue Hemorrhagic Fever: The Emergence of a Global Health Problem', *Emerging Infectious Disease journal*, 1(2), p. 55. doi: 10.3201/eid0102.950204.

Guirakhoo, F., Bolin, R. A. and Roehrig, J. T. (1992) 'The Murray Valley encephalitis virus prM protein confers acid resistance to virus particles and alters the expression of epitopes within the R2 domain of E glycoprotein', *Virology*, 191(2), pp. 921–931. doi: https://doi.org/10.1016/0042-6822(92)90267-S.

Guy, B. *et al.* (2017) 'A recombinant live attenuated tetravalent vaccine for the prevention of dengue', *Expert Review of Vaccines*. Taylor & Francis, 16(7), pp. 671–683. doi: 10.1080/14760584.2017.1335201.

Guzman, M. G. and Alvarez, M. et al. (2013) *Secondary infection as a risk factor for dengue hemorrhagic fever/dengue shock syndrome: An historical perspective and role of antibody-dependent enhancement of infection*, *Archives of virology*. doi: 10.1007/s00705-013-1645-3.

Hajdin, C. E. *et al.* (2013) 'Accurate SHAPE-directed RNA secondary structure modeling, including pseudoknots.', *Proceedings of the National Academy of Sciences of the United States of America*. United States, 110(14), pp. 5498–5503. doi: 10.1073/pnas.1219988110.

Hall-Mendelin, S. *et al.* (2016) 'The insect-specific Palm Creek virus modulates West Nile virus infection in and transmission by Australian mosquitoes', *Parasites & Vectors*, 9(1), p. 414. doi: 10.1186/s13071-016-1683-2.

Halstead, S. (2003) *Neutralization and Antibody-Dependent Enhancement of*

*Dengue Viruses*, *Advances in virus research*. doi: 10.1016/S0065-3527(03)60011-4.

Halstead, S. (2008) 'Dengue: Overview and History', *Tropical Medicine: Science and Practice*. doi: 10.1142/9781848162297_0001.

Halstead, S. B. (2014) 'Dengue Antibody-Dependent Enhancement: Knowns and Unknowns', *Microbiology Spectrum*, 2(6). http://www.asmscience.org/content/journal/microbiolspec/10.1128/microbiolspec.AID-0022-2014.

Harris, E. *et al.* (2000) 'Clinical, epidemiologic, and virologic features of dengue in the 1998 epidemic in Nicaragua.', *The American journal of tropical medicine and hygiene*. United States, 63(1–2), pp. 5–11.

Heinz, F. X. *et al.* (1994) 'Structural changes and functional control of the tick-borne encephalitis virus glycoprotein E by the heterodimeric association with protein prM.', *Virology*. United States, 198(1), pp. 109–117. doi: 10.1006/viro.1994.1013.

Hodge, K. *et al.* (2016) 'Identification of a Conserved RNA-dependent RNA Polymerase (RdRp)-RNA Interface Required for Flaviviral Replication', *Journal of Biological Chemistry* , 291(33), pp. 17437–17449. doi: 10.1074/jbc.M116.724013.

Huthoff, H. and Berkhout, B. (2001) 'Two alternating structures of the HIV-1 leader RNA.', *RNA (New York, N.Y.)*. United States, 7(1), pp. 143–157.

IDT (2018) *RNA modification by SHAPE reagents*. http://www.idtdna.com/pages/education/decoded/article/single-nucleotide-resolution-of-rna-structure (Accessed: 27 February 2018).

Igarashi, A. (1978) 'Isolation of a Singh's Aedes albopictus cell clone sensitive to Dengue and Chikungunya viruses.', *The Journal of general virology*. England, 40(3), pp. 531–544. doi: 10.1099/0022-1317-40-3-531.

IGEM (2018) *The Biochemistry Behind Protein Purification*. http://2015.igem.org/Team:Freiburg/Project/Protein_Purification (Accessed: 22 February 2018).

Iglesias, N. G., Filomatori, C. V and Gamarnik, A. V (2011) 'The F1 Motif of

Dengue Virus Polymerase NS5 Is Involved in Promoter-Dependent RNA Synthesis', *Journal of Virology* , 85(12), pp. 5745–5756. doi: 10.1128/JVI.02343-10.

Inoue, T. and Tsai, B. (2013) 'How Viruses Use the Endoplasmic Reticulum for Entry, Replication, and Assembly', *Cold Spring Harbor Perspectives in Biology*. Cold Spring Harbor Laboratory Press, 5(1), p. a013250. doi: 10.1101/cshperspect.a013250.

Ito, T., Tahara, S. M. and Lai, M. M. C. (1998) 'The 3′-Untranslated Region of Hepatitis C Virus RNA Enhances Translation from an Internal Ribosomal Entry Site', *Journal of Virology*. American Society for Microbiology, 72(11), pp. 8789–8796. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC110295/.

J. Johnson, A. and Roehrig, J. (1999) *New Mouse Model for Dengue Virus Vaccine Testing*, *Journal of virology*. doi: https://doi.org/10.1016/j.vaccine.2015.09.112.

Jin, Z. *et al.* (2011) 'Characterization of the Elongation Complex of Dengue Virus RNA Polymerase: assembly, kinetics of nucelotide incorporation, and fidelity', *Journal of Biological Chemistry*, 286(3), pp. 2067–2077. doi: 10.1074/jbc.M110.162685.

Junjhon, J. *et al.* (2014) 'Ultrastructural Characterization and Three-Dimensional Architecture of Replication Sites in Dengue Virus-Infected Mosquito Cells', *Journal of Virology*. Edited by R. W. Doms. 1752 N St., N.W., Washington, DC: American Society for Microbiology, 88(9), pp. 4687–4697. doi: 10.1128/JVI.00118-14.

Kao, C. C., Singh, P. and Ecker, D. J. (2001) 'De Novo Initiation of Viral RNA-Dependent RNA Synthesis', *Virology*, 287(2), pp. 251–260. doi: https://doi.org/10.1006/viro.2001.1039.

Karabiber, F. *et al.* (2013) 'QuShape: rapid, accurate, and best-practices quantification of nucleic acid probing information, resolved by capillary electrophoresis.', *RNA (New York, N.Y.)*. United States, 19(1), pp. 63–73. doi: 10.1261/rna.036327.112.

Kenyon, J. C. *et al.* (2013) 'In-gel probing of individual RNA conformers within a

mixed population reveals a dimerization structural switch in the HIV-1 leader.', *Nucleic acids research*. England, 41(18), p. e174. doi: 10.1093/nar/gkt690.

Kenyon, J., Prestwood, L. and Lever, A. (2014) 'Current perspectives on RNA secondary structure probing.', *Biochemical Society transactions*. England, 42(4), pp. 1251–1255. doi: 10.1042/BST20140084.

Khromykh, A. A. *et al.* (2001) 'Essential Role of Cyclization Sequences in Flavivirus RNA Replication', *Journal of Virology* , 75(14), pp. 6719–6728. doi: 10.1128/JVI.75.14.6719-6728.2001.

Klungthong, C. *et al.* (2004) 'The molecular epidemiology of dengue virus serotype 4 in Bangkok, Thailand', *Virology*, 329(1), pp. 168–179. doi: https://doi.org/10.1016/j.virol.2004.08.003.

Kovermann, M., Rogne, P. and Wolf-Watz, M. (2016) 'Protein dynamics and function from solution state NMR spectroscopy.', *Quarterly reviews of biophysics*. England, 49, p. e6. doi: 10.1017/S0033583516000019.

Kubota, M., Tran, C. and Spitale, R. C. (2015) 'Progress and challenges for chemical probing of RNA structure inside living cells', *Nature chemical biology*, 11(12), pp. 933–941. doi: 10.1038/nchembio.1958.

Kudla, G. *et al.* (2011) 'Cross-linking, ligation, and sequencing of hybrids reveals RNA-RNA interactions in yeast.', *Proceedings of the National Academy of Sciences of the United States of America*. United States, 108(24), pp. 10010–10015. doi: 10.1073/pnas.1017386108.

Kuhn, R. J. *et al.* (2002) 'Structure of Dengue Virus: Implications for Flavivirus Organization, Maturation, and Fusion', *Cell*, 108(5), pp. 717–725. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4152842/.

Kupitz, C. *et al.* (2014) 'Microcrystallization techniques for serial femtosecond crystallography using photosystem II from Thermosynechococcus elongatus as a model system', *Philosophical Transactions of the Royal Society of London B: Biological Sciences*. The Royal Society, 369(1647). doi: 10.1098/rstb.2013.0316.

Kuruvilla, J. G. *et al.* (2007) 'Dengue virus infection and immune response in humanized RAG2−/−γc−/− (RAG-hu) mice', *Virology*, 369(1), pp. 143–152. doi:

https://doi.org/10.1016/j.virol.2007.06.005.

Lee, B. *et al.* (2017) ‘Comparison of SHAPE reagents for mapping RNA structures inside living cells’, *RNA*. Cold Spring Harbor Laboratory Press, 23(2), pp. 169–174. doi: 10.1261/rna.058784.116.

Leontis, N. B., Stombaugh, J. and Westhof, E. (2002) ‘The non-Watson–Crick base pairs and their associated isostericity matrices’, *Nucleic Acids Research*, 30(16), pp. 3497–3531. http://dx.doi.org/10.1093/nar/gkf481.

Leung, J. Y. *et al.* (2008) ‘Role of Nonstructural Protein NS2A in Flavivirus Assembly ’, *Journal of Virology*. American Society for Microbiology (ASM), 82(10), pp. 4731–4741. doi: 10.1128/JVI.00002-08.

Li, Y. *et al.* (2015) ‘Membrane topology of NS2B of dengue virus revealed by NMR spectroscopy’, *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1848(10, Part A), pp. 2244–2252. doi: https://doi.org/10.1016/j.bbamem.2015.06.010.

Lim, S. P. *et al.* (2013) ‘A Crystal Structure of the Dengue Virus Non-structural Protein 5 (NS5) Polymerase Delineates Interdomain Amino Acid Residues That Enhance Its Thermostability and de Novo Initiation Activities ’, *Journal of Biological Chemistry* , 288(43), pp. 31105–31114. doi: 10.1074/jbc.M113.508606.

Lim, S. P. *et al.* (2016) ‘Potent Allosteric Dengue Virus NS5 Polymerase Inhibitors: Mechanism of Action and Resistance Profiling.’, *PLoS pathogens*. United States, 12(8), p. e1005737. doi: 10.1371/journal.ppat.1005737.

Lin, C. *et al.* (1993) ‘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.’, *Journal of Virology* , 67(4), pp. 2327–2335. http://jvi.asm.org/content/67/4/2327.abstract.

Lindenbach, B. D. and Rice, C. M. (1999) ‘Genetic Interaction of Flavivirus Nonstructural Proteins NS1 and NS4A as a Determinant of Replicase Function’, *Journal of Virology* , 73(6), pp. 4611–4621. http://jvi.asm.org/content/73/6/4611.abstract.

Liu, Z.-Y. *et al.* (2016) ‘Viral RNA switch mediates the dynamic control of

flavivirus replicase recruitment by genome cyclization', *eLife*. Edited by J. S. Kieft. eLife Sciences Publications, Ltd, 5, p. e17636. doi: 10.7554/eLife.17636.

Lo, M. K. *et al.* (2003) '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', *Journal of virology*, 77(18), p. 10004—10014. doi: 10.1128/jvi.77.18.10004-10014.2003.

Lodeiro, M. F. and Filomatori, C. (2009) 'Structural and Functional Studies of the Promoter Element for Dengue Virus RNA Replication', *Journal of Virology*, 83(2), pp. 993–1008. doi: 10.1128/JVI.01647-08.

Lott, W. B. and Doran, M. R. (2013) 'Do RNA viruses require genome cyclisation for replication?', *Trends in Biochemical Sciences*, 38(7), pp. 350–355. doi: https://doi.org/10.1016/j.tibs.2013.04.005.

Lu, G. and Gong, P. (2013) 'Crystal Structure of the Full-Length Japanese Encephalitis Virus NS5 Reveals a Conserved Methyltransferase-Polymerase Interface', *PLOS Pathogens*. Public Library of Science, 9(8), p. e1003549. https://doi.org/10.1371/journal.ppat.1003549.

Lucks, J. B. *et al.* (2011) 'Multiplexed RNA structure characterization with selective 2′-hydroxyl acylation analyzed by primer extension sequencing (SHAPE-Seq)', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 108(27), pp. 11063–11068. doi: 10.1073/pnas.1106501108.

Luo, D. *et al.* (2008) 'Crystal Structure of the NS3 Protease-Helicase from Dengue Virus', *Journal of Virology* , 82(1), pp. 173–183. doi: 10.1128/JVI.01788-07.

Ma, L. *et al.* (2004) 'Solution structure of dengue virus capsid protein reveals another fold', *Proceedings of the National Academy of Sciences of the United States of America*, 101(10), p. 3414—3419. doi: 10.1073/pnas.0305892101.

Mackenzie, J. (2005) 'Wrapping things up about virus RNA replication.', *Traffic (Copenhagen, Denmark)*. England, 6(11), pp. 967–977. doi: 10.1111/j.1600-0854.2005.00339.x.

Mackenzie, J. M., Jones, M. K. and Young, P. R. (1996) ‘Immunolocalization of the Dengue Virus Nonstructural Glycoprotein NS1 Suggests a Role in Viral RNA Replication’, *Virology*, 220(1), pp. 232–240. doi: https://doi.org/10.1006/viro.1996.0307.

Macpherson, I. and Stoker, M. (1962) ‘Polyoma transformation of hamster cell clones--an investigation of genetic factors affecting cell competence.’, *Virology*. United States, 16, pp. 147–151.

Malet, H. *et al.* (2007) ‘Crystal Structure of the RNA Polymerase Domain of the West Nile Virus Non-structural Protein 5’, *Journal of Biological Chemistry* , 282(14), pp. 10678–10689. doi: 10.1074/jbc.M607273200.

Malet, H. *et al.* (2008) *The flavivirus polymerase as a target for drug discovery*, *Antiviral research*. doi: 10.1016/j.antiviral.2008.06.007.

Manas, M. *et al.* (2015) ‘Stacking geometry for non-canonical G:U wobble base pair containing dinucleotide sequences in RNA: dispersion-corrected DFT-D study’, *Biopolymers*. Wiley-Blackwell, 103(6), pp. 328–338. doi: 10.1002/bip.22616.

Martina, B. E. E., Koraka, P. and Osterhaus, A. D. M. E. (2009) ‘Dengue Virus Pathogenesis: an Integrated View’, *Clinical Microbiology Reviews*. American Society for Microbiology (ASM), 22(4), pp. 564–581. doi: 10.1128/CMR.00035-09.

Marx, V. (2017) ‘Structural biology: doors open at the European XFEL’, *Nature Methods*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 14, p. 843. http://dx.doi.org/10.1038/nmeth.4394.

McCoy, A. J. *et al.* (2007) ‘Phaser crystallographic software.’, *Journal of applied crystallography*. United States, 40(Pt 4), pp. 658–674. doi: 10.1107/S0021889807021206.

McGraw, E. and O’Neill, S. (2013) *Beyond insecticides: New thinking on an ancient problem*, *Nature reviews. Microbiology*. doi: 10.1038/nrmicro2968.

McKnight, K. L. and Lemon, S. M. (1998) ‘The rhinovirus type 14 genome contains an internally located RNA structure that is required for viral replication’,

*RNA (New York, N.Y.)*, 4(12), p. 1569—1584. doi: 10.1017/s1355838298981006.

McLean, J. E. *et al.* (2011) 'Flavivirus NS4A-induced Autophagy Protects Cells against Death and Enhances Virus Replication', *Journal of Biological Chemistry* , 286(25), pp. 22147–22159. doi: 10.1074/jbc.M110.192500.

McSherry, J. A. (1982) 'Some Medical Aspects of the Darien Scheme: Was it Dengue?', *Scottish Medical Journal*. SAGE Publications, 27(2), pp. 183–184. doi: 10.1177/003693308202700215.

Megret, F. *et al.* (1992) 'Use of recombinant fusion proteins and monoclonal antibodies to define linear and discontinuous antigenic sites on the dengue virus envelope glycoprotein', *Virology*, 187(2), pp. 480–491. doi: https://doi.org/10.1016/0042-6822(92)90450-4.

Messina, J. P. *et al.* (2014) 'Global spread of dengue virus types: mapping the 70 year history', *Trends in Microbiology*, 22(3), pp. 138–146. doi: https://doi.org/10.1016/j.tim.2013.12.011.

Miller, J. L. *et al.* (2008) 'The Mannose Receptor Mediates Dengue Virus Infection of Macrophages', *PLOS Pathogens*. Public Library of Science, 4(2), p. e17. https://doi.org/10.1371/journal.ppat.0040017.

Miller, S. and Krijnse-Locker, J. (2008) 'Modification of intracellular membrane structures for virus replication.', *Nature reviews. Microbiology*. England, 6(5), pp. 363–374. doi: 10.1038/nrmicro1890.

Miller, S., Sparacio, S. and Bartenschlager, R. (2006) 'Subcellular Localization and Membrane Topology of the Dengue Virus Type 2 Non-structural Protein 4B', *Journal of Biological Chemistry* , 281(13), pp. 8854–8863. doi: 10.1074/jbc.M512697200.

Modis, Y. *et al.* (2004) *Structure of the dengue virus envelope protein after membrane fusion*, *Nature*. doi: 10.1038/nature02165.

Moi, M. L., Takasaki, T. and Kurane, I. (2016) 'Human antibody response to dengue virus: implications for dengue vaccine design', *Tropical Medicine and Health*, 44(1), p. 1. doi: 10.1186/s41182-016-0004-y.

Mores, C. N., Christofferson, R. C. and Davidson, S. A. (2014) 'The Role of the

Mosquito in a Dengue Human Infection Model', *The Journal of Infectious Diseases*, 209(suppl_2), pp. S71–S78. http://dx.doi.org/10.1093/infdis/jiu110.

Mota, J. and Rico-Hesse, R. (2009) 'Humanized Mice Show Clinical Signs of Dengue Fever according to Infecting Virus Genotype', *Journal of Virology* , 83(17), pp. 8638–8645. doi: 10.1128/JVI.00581-09.

Murata, K. and Wolf, M. (2018) 'Cryo-electron microscopy for structural analysis of dynamic biological macromolecules', *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1862(2), pp. 324–334. doi: https://doi.org/10.1016/j.bbagen.2017.07.020.

Murshudov, G. N. *et al.* (2011) ' REFMAC5 for the refinement of macromolecular crystal structures', *Acta Crystallographica Section D: Biological Crystallography*. International Union of Crystallography, 67(Pt 4), pp. 355–367. doi: 10.1107/S0907444911001314.

Mustafa, M. S. *et al.* (2015) 'Discovery of fifth serotype of dengue virus (DENV-5): A new public health dilemma in dengue control', *Medical Journal, Armed Forces India*. Elsevier, 71(1), pp. 67–70. doi: 10.1016/j.mjafi.2014.09.011.

Nagy, P. and Pogany, J. (2011) *The dependence of viral RNA replication on co-opted host factors*, *Nature reviews. Microbiology*. doi: 10.1038/nrmicro2692.

Nature Education (2018) 'What is Dengue Fever?' https://www.nature.com/scitable/topicpage/what-is-dengue-fever-22399100.

Nayak, V. *et al.* (2009) 'Crystal Structure of Dengue Virus Type 1 Envelope Protein in the Postfusion Conformation and Its Implications for Membrane Fusion', *Journal of Virology* , 83(9), pp. 4338–4344. doi: 10.1128/JVI.02574-08.

Nealon, J. *et al.* (2016) *Symptomatic Dengue Disease in Five Southeast Asian Countries: Epidemiological Evidence from a Dengue Vaccine Trial*, *PLoS neglected tropical diseases*. doi: 10.1371/journal.pntd.0004918.

Nicholson, B. L. and White, K. A. (2014) 'Functional long-range RNA–RNA interactions in positive-strand RNA viruses', *Nature Reviews Microbiology*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 12, p. 493. http://dx.doi.org/10.1038/nrmicro3288.

Nikolcheva, T. and Woodson, S. A. (1999) ‘Facilitation of Group I Splicing in Vivo: Misfolding of the Tetrahymena IVS and the Role of Ribosomal RNA Exons’, *Journal of Molecular Biology*, 292(3), pp. 557–567. doi: https://doi.org/10.1006/jmbi.1999.3083.

Nilsen, T. W. (2013) ‘RNA Structure Determination Using Nuclease Digestion’, *Cold Spring Harbor Protocols* , 2013(4), p. pdb.prot072330. doi: 10.1101/pdb.prot072330.

Noble, C. G. *et al.* (2013) ‘Conformational Flexibility of the Dengue Virus RNA-Dependent RNA Polymerase Revealed by a Complex with an Inhibitor’, *Journal of Virology* , 87(9), pp. 5291–5295. doi: 10.1128/JVI.00045-13.

Noble, C. G. and Shi, P.-Y. (2012) ‘Structural biology of dengue virus enzymes: Towards rational design of therapeutics’, *Antiviral Research*, 96(2), pp. 115–126. doi: https://doi.org/10.1016/j.antiviral.2012.09.007.

Nomaguchi, M. *et al.* (2003) ‘De Novo Synthesis of Negative-Strand RNA by Dengue Virus RNA-Dependent RNA Polymerase In Vitro: Nucleotide, Primer, and Template Parameters’, *Journal of Virology* , 77(16), pp. 8831–8842. doi: 10.1128/JVI.77.16.8831-8842.2003.

Offerdahl, D. K. *et al.* (2012) ‘A three-dimensional comparison of tick-borne flavivirus infection in mammalian and tick cell lines.’, *PloS one*. United States, 7(10), p. e47912. doi: 10.1371/journal.pone.0047912.

OhAinle, M. *et al.* (2011) ‘Dynamics of Dengue Disease Severity Determined by the Interplay Between Viral Genetics and Serotype-Specific Immunity’, *Science Translational Medicine*, 3(114), p. 114ra128 LP-114ra128. http://stm.sciencemag.org/content/3/114/114ra128.abstract.

Oliphant, T. *et al.* (2007) *Antibody Recognition and Neutralization Determinants on Domains I and II of West Nile Virus Envelope Protein*, *Journal of virology*. doi: 10.1128/JVI.01732-06.

Oxitec (2018a) *How it works*. http://www.oxitec.com/our-solution/technology/how-it-works/ (Accessed: 8 March 2018).

Oxitec (2018b) *The Science*. http://www.oxitec.com/our-solution/technology/the-

science/ (Accessed: 8 March 2018).

Panavas, T., Sanders, C. and Butt, T. R. (2009) ‘SUMO fusion technology for enhanced protein production in prokaryotic and eukaryotic expression systems.’, *Methods in molecular biology (Clifton, N.J.)*. United States, 497, pp. 303–317. doi: 10.1007/978-1-59745-566-4_20.

Panja, S., Schu, D. J. and Woodson, S. A. (2013) ‘Conserved arginines on the rim of Hfq catalyze base pair formation and exchange.’, *Nucleic acids research*. England, 41(15), pp. 7536–7546. doi: 10.1093/nar/gkt521.

Paul, L. M. *et al.* (2016) ‘Dengue Virus Antibodies Enhance Zika Virus Infection’, *bioRxiv*. http://biorxiv.org/content/early/2016/04/25/050112.abstract.

PDB (2018) *Introduction to Biological Assemblies and the PDB Archive*. https://pdb101.rcsb.org/learn/guide-to-understanding-pdb-data/biological-assemblies (Accessed: 3 April 2018).

Perera, R. and Kuhn, R. J. (2008) ‘Structural Proteomics of Dengue Virus’, *Current opinion in microbiology*, 11(4), pp. 369–377. doi: 10.1016/j.mib.2008.06.004.

Pichlmair, A. *et al.* (2006) ‘RIG-I-Mediated Antiviral Responses to Single-Stranded RNA Bearing 5'-Phosphates’, *Science*, 314(5801), p. 997 LP-1001. http://science.sciencemag.org/content/314/5801/997.abstract.

Potisopon, S. *et al.* (2014) ‘The methyltransferase domain of dengue virus protein NS5 ensures efficient RNA synthesis initiation and elongation by the polymerase domain’, *Nucleic acids research*, 42(18), p. 11642—11656. doi: 10.1093/nar/gku666.

Potterton, E. *et al.* (2003) ‘A graphical user interface to the CCP4 program suite.’, *Acta crystallographica. Section D, Biological crystallography*. United States, 59(Pt 7), pp. 1131–1137.

Promega (2018) *Protein Purification and Analysis*. https://www.promega.co.uk/resources/product-guides-and-selectors/protocols-and-applications-guide/protein-purification-and-analysis/ (Accessed: 22 February 2018).

Raut, C. *et al.* (1996) *Susceptibility of laboratory-bred rodent to the experimental infection with dengue virus type 2*, *Acta virologica*.

RCSB PDB (2018) *Nature of 3D Structural Data*. http://www.rcsb.org/pdb/static.do?p=general_information%5Cabout_pdb%5Cnature_of_3d_structural_data.html (Accessed: 22 February 2018).

Reddy, T. and Sansom, M. S. P. (2018) 'The Role of the Membrane in the Structure and Biophysical Robustness of the Dengue Virion Envelope', *Structure*. Elsevier, 24(3), pp. 375–382. doi: 10.1016/j.str.2015.12.011.

Reuter, J. S. and Mathews, D. H. (2010) 'RNAstructure: software for RNA secondary structure prediction and analysis.', *BMC bioinformatics*. England, 11, p. 129. doi: 10.1186/1471-2105-11-129.

Rey, F. A. (2003) 'Dengue virus envelope glycoprotein structure: New insight into its interactions during viral entry', *Proceedings of the National Academy of Sciences*, 100(12), p. 6899 LP-6901. http://www.pnas.org/content/100/12/6899.abstract.

Rice, G. M., Leonard, C. W. and Weeks, K. M. (2014) 'RNA secondary structure modeling at consistent high accuracy using differential SHAPE.', *RNA (New York, N.Y.)*. United States, 20(6), pp. 846–854. doi: 10.1261/rna.043323.113.

Richter, M. K. S. *et al.* (2014) 'Immature dengue virus is infectious in human immature dendritic cells via interaction with the receptor molecule DC-SIGN.', *PloS one*. United States, 9(6), p. e98785. doi: 10.1371/journal.pone.0098785.

Rivinoja, A. *et al.* (2011) *Golgi pH, its regulation and roles in human disease*, *Annals of medicine*. doi: 10.3109/07853890.2011.579150.

Rodenhuis-Zybert, I. A. *et al.* (2010) 'Immature Dengue Virus: A Veiled Pathogen?', *PLOS Pathogens*. Public Library of Science, 6(1), p. e1000718. https://doi.org/10.1371/journal.ppat.1000718.

Rodenhuis-Zybert, I. A. *et al.* (2011) 'A fusion-loop antibody enhances the infectious properties of immature flavivirus particles', *Journal of virology*, 85(22), p. 11800—11808. doi: 10.1128/jvi.05237-11.

Romero-Brey, I. and Bartenschlager, R. (2014) *Membranous Replication*

*Factories Induced by Plus-Strand RNA Viruses*, *Viruses*. doi: 10.3390/v6072826.

Rose, W. *et al.* (2017) 'Dengue illness in children', *Current Medical Issues*, 15(2), pp. 95–105. doi: 10.4103/cmi.cmi_25_17.

Rouvinski, A. *et al.* (2017) *Covalently linked dengue virus envelope glycoprotein dimers reduce exposure of the immunodominant fusion loop epitope*, *Nature Communications*. doi: 10.1038/ncomms15411.

La Ruche, G. *et al.* (2010) *First two autochthonous dengue virus infection in metropolitan France, September 2010*, *Euro surveillance : bulletin européen sur les maladies transmissibles = European communicable disease bulletin*.

Ryan, M., Monaghan, S. and Flint, M. (1998) *Virus-encoded proteinases of the Flaviviridae*, *The Journal of general virology*. doi: 10.1099/0022-1317-79-5-947.

Sabchareon, A. *et al.* (2018) 'Protective efficacy of the recombinant, live-attenuated, CYD tetravalent dengue vaccine in Thai schoolchildren: a randomised, controlled phase 2b trial', *The Lancet*. Elsevier, 380(9853), pp. 1559–1567. doi: 10.1016/S0140-6736(12)61428-7.

El Sahili, A. and Lescar, J. (2017) 'Dengue Virus Non-Structural Protein 5', *Viruses* . doi: 10.3390/v9040091.

Salonen, A., Ahola, T. and Kaariainen, L. (2005) 'Viral RNA replication in association with cellular membranes.', *Current topics in microbiology and immunology*. Germany, 285, pp. 139–173.

Samsa, M. M. *et al.* (2009) 'Dengue virus capsid protein usurps lipid droplets for viral particle formation.', *PLoS pathogens*. United States, 5(10), p. e1000632. doi: 10.1371/journal.ppat.1000632.

Sasmono, R. T. *et al.* (2015) 'Genomic analysis and growth characteristic of dengue viruses from Makassar, Indonesia', *Infection, Genetics and Evolution*, 32, pp. 165–177. doi: https://doi.org/10.1016/j.meegid.2015.03.006.

Schlichting, I. (2005) 'X-Ray Crystallography of Protein-Ligand Interactions BT - Protein-Ligand Interactions: Methods and Applications', in Ulrich Nienhaus, G. (ed.). Totowa, NJ: Humana Press, pp. 155–165. doi: 10.1385/1-59259-912-5:155.

Schroeder, R. *et al.* (2002) ‘RNA folding in vivo.’, *Current opinion in structural biology*. England, 12(3), pp. 296–300.

Selisko, B. *et al.* (2006) ‘Comparative mechanistic studies of de novo RNA synthesis by flavivirus RNA-dependent RNA polymerases’, *Virology*, 351(1), pp. 145–158. doi: https://doi.org/10.1016/j.virol.2006.03.026.

Selisko, B. *et al.* (2012) ‘Molecular basis for nucleotide conservation at the ends of the dengue virus genome.’, *PLoS pathogens*. United States, 8(9), p. e1002912. doi: 10.1371/journal.ppat.1002912.

Selisko, B. *et al.* (2014) ‘Regulation of Flavivirus RNA synthesis and replication’, *Current Opinion in Virology*, 9, pp. 74–83. doi: https://doi.org/10.1016/j.coviro.2014.09.011.

Semenza, J. C. *et al.* (2014) ‘International Dispersal of Dengue through Air Travel: Importation Risk for Europe’, *PLOS Neglected Tropical Diseases*. Public Library of Science, 8(12), p. e3278. https://doi.org/10.1371/journal.pntd.0003278.

Shresta, S. *et al.* (2006) ‘Murine Model for Dengue Virus-Induced Lethal Disease with Increased Vascular Permeability’, *Journal of Virology* , 80(20), pp. 10208–10217. doi: 10.1128/JVI.00062-06.

Shrivastava, G. *et al.* (2017) *NS2A comprises a putative viroporin of Dengue virus 2*, *Virulence*. doi: 10.1080/21505594.2017.1356540.

Shu, B. and Gong, P. (2017) ‘The uncoupling of catalysis and translocation in the viral RNA-dependent RNA polymerase.’, *RNA biology*. United States, 14(10), pp. 1314–1319. doi: 10.1080/15476286.2017.1300221.

da Silva-Voorham, J. M. *et al.* (2009) ‘Dengue: a growing risk to travellers to tropical and sub-tropical regions.’, *Nederlands tijdschrift voor geneeskunde*. Netherlands, 153, p. A778.

Sim, S. and Hibberd, M. L. (2016) ‘Genomic approaches for understanding dengue: insights from the virus, vector, and host’, *Genome biology*, 17, p. 38. doi: 10.1186/s13059-016-0907-2.

Smith, S. A. *et al.* (2016) ‘Dengue Virus prM-Specific Human Monoclonal

Antibodies with Virus Replication-Enhancing Properties Recognize a Single Immunodominant Antigenic Site', *Journal of Virology* , 90(2), pp. 780–789. doi: 10.1128/JVI.01805-15.

Song, C. and Simon, A. (1995) *Requirement of a 3'-Terminal Stem-loop inin VitroTranscription by an RNA-dependent RNA Polymerase*, *Journal of molecular biology*. doi: 10.1006/jmbi.1995.0594.

Soni, A. *et al.* (2001) *Management of dengue fever in ICU*, *Indian journal of pediatrics*. doi: 10.1007/BF02722356.

Sousa, C. *et al.* (2012) *Rapid communications ongoing outbreak of dengue type 1 in the autonomous region of Madeira, Portugal: Preliminary report*, *Euro surveillance: bulletin europeen sur les maladies transmissibles = European communicable disease bulletin*. doi: 10.2807/ese.17.49.20333-en.

Spasic, A. *et al.* (2018) 'Modeling RNA secondary structure folding ensembles using SHAPE mapping data.', *Nucleic acids research*. England, 46(1), pp. 314–323. doi: 10.1093/nar/gkx1057.

Spitale, R. C. *et al.* (2013) 'RNA SHAPE analysis in living cells.', *Nature chemical biology*. United States, 9(1), pp. 18–20. doi: 10.1038/nchembio.1131.

Stadler, K. *et al.* (1997) *Proteolytic activation of tick-borne encephalitis virus by furin*, *Journal of virology*.

Staple, D. W. and Butcher, S. E. (2005) 'Pseudoknots: RNA Structures with Diverse Functions', *PLOS Biology*. Public Library of Science, 3(6), p. e213. https://doi.org/10.1371/journal.pbio.0030213.

Stern, O. *et al.* (2013) 'An N-Terminal Amphipathic Helix in Dengue Virus Nonstructural Protein 4A Mediates Oligomerization and Is Essential for Replication', *Journal of Virology* , 87(7), pp. 4080–4085. doi: 10.1128/JVI.01900-12.

Stewart, L., Clark, R. and Behnke, C. (2002) 'High-throughput crystallization and structure determination in drug discovery', *Drug Discovery Today*, 7(3), pp. 187–196. doi: https://doi.org/10.1016/S1359-6446(01)02121-3.

Sun, C. *et al.* (2013) 'Two RNA-binding motifs in eIF3 direct HCV IRES-

dependent translation', *Nucleic Acids Research*. Oxford University Press, 41(15), pp. 7512–7521. doi: 10.1093/nar/gkt510.

Sztuba-Solinska, J. *et al.* (2013) 'Structural complexity of Dengue virus untranslated regions: cis -acting RNA motifs and pseudoknot interactions modulating functionality of the viral genome ', *Nucleic Acids Research*, 41(9), pp. 5075–5089. http://dx.doi.org/10.1093/nar/gkt203.

Sztuba-Solinska, J. and Le Grice, S. F. J. (2014) 'Insights into Secondary and Tertiary Interactions of Dengue Virus RNA by SHAPE BT - Dengue: Methods and Protocols', in Padmanabhan, R. and Vasudevan, S. G. (eds). New York, NY: Springer New York, pp. 225–239. doi: 10.1007/978-1-4939-0348-1_14.

Technobis Crystallisation systems (2018) *Co-crystallization studies*. https://www.crystallizationsystems.com/media/files/applications/Technobis Crystallization Systems AN5 Co-crystallization studies.pdf.

Teo, C. S. H. and Chu, J. J. H. (2014) 'Cellular Vimentin Regulates Construction of Dengue Virus Replication Complexes through Interaction with NS4A Protein', *Journal of Virology* , 88(4), pp. 1897–1913. doi: 10.1128/JVI.01249-13.

Tuplin, A. (2015) 'Diverse roles and interactions of RNA structures during the replication of positive-stranded RNA viruses of humans and animals.', *The Journal of general virology*. England, 96(Pt 7), pp. 1497–1503. doi: 10.1099/vir.0.000066.

TutorCircle (2018) *Waves interference*. http://physics.tutorcircle.com/waves/wave-interference.html (Accessed: 20 April 2018).

Uchil, P. D., Kumar, A. V. A. and Satchidanandam, V. (2006) 'Nuclear Localization of Flavivirus RNA Synthesis in Infected Cells', *Journal of Virology* , 80(11), pp. 5451–5464. doi: 10.1128/JVI.01982-05.

Vicente, C. R. *et al.* (2016) 'Serotype influences on dengue severity: a cross-sectional study on 485 confirmed dengue cases in Vitória, Brazil', *BMC Infectious Diseases*, 16(1), p. 320. doi: 10.1186/s12879-016-1668-y.

Villordo, S. M. *et al.* (2015) 'Dengue Virus RNA Structure Specialization

Facilitates Host Adaptation', *PLOS Pathogens*. Public Library of Science, 11(1), p. e1004604. https://doi.org/10.1371/journal.ppat.1004604.

Villordo, S. M. *et al.* (2016) 'RNA Structure Duplications and Flavivirus Host Adaptation', *Trends in Microbiology*, 24(4), pp. 270–283. doi: https://doi.org/10.1016/j.tim.2016.01.002.

Villordo, S. M., Alvarez, D. E. and Gamarnik, A. V (2010) 'A balance between circular and linear forms of the dengue virus genome is crucial for viral replication', *RNA*. Cold Spring Harbor Laboratory Press, 16(12), pp. 2325–2335. doi: 10.1261/rna.2120410.

Villordo, S. M. and Gamarnik, A. V (2009) 'Genome cyclization as strategy for flavivirus RNA replication.', *Virus research*. Netherlands, 139(2), pp. 230–239. doi: 10.1016/j.virusres.2008.07.016.

Villordo, S. M. and Gamarnik, A. V (2013) 'Differential RNA Sequence Requirement for Dengue Virus Replication in Mosquito and Mammalian Cells', *Journal of Virology*. 1752 N St., N.W., Washington, DC: American Society for Microbiology, 87(16), pp. 9365–9372. doi: 10.1128/JVI.00567-13.

Vogels, C. B. F. *et al.* (2017) 'Vector competence of European mosquitoes for West Nile virus', *Emerging Microbes &Amp; Infections*. The Author(s), 6, p. e96. http://dx.doi.org/10.1038/emi.2017.82.

Wahala, W. M. P. B. and de Silva, A. M. (2011) 'The Human Antibody Response to Dengue Virus Infection', *Viruses* . doi: 10.3390/v3122374.

Walker, T. *et al.* (2011) *The wMel Wolbachia strain blocks dengue and invades caged Aedes aegypti populations*, *Nature*. doi: 10.1038/nature10355.

Wang, C.-C. *et al.* (2017) 'Insights into the coordinated interplay of the sHP hairpin and its co-existing and mutually-exclusive dengue virus terminal RNA elements for viral replication', *Virology*, 505, pp. 56–70. doi: https://doi.org/10.1016/j.virol.2017.02.007.

Wang, S., He, R. and Anderson, R. (1999) 'PrM- and Cell-Binding Domains of the Dengue Virus E Protein', *Journal of Virology*. American Society for Microbiology, 73(3), pp. 2547–2551. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC104503/.

Wang, Y. *et al.* (2010) ‘A Theoretical Study of the Separation Principle in Size Exclusion Chromatography’, *Macromolecules*. American Chemical Society, 43(3), pp. 1651–1659. doi: 10.1021/ma902377g.

Watters, K. E. *et al.* (2016) ‘Characterizing RNA structures in vitro and in vivo with selective 2′-hydroxyl acylation analyzed by primer extension sequencing (SHAPE-Seq)’, *Methods (San Diego, Calif.)*, 103, pp. 34–48. doi: 10.1016/j.ymeth.2016.04.002.

Wells, S. E. *et al.* (1998) ‘Circularization of mRNA by Eukaryotic Translation Initiation Factors’, *Molecular Cell*. Elsevier, 2(1), pp. 135–140. doi: 10.1016/S1097-2765(00)80122-7.

Welsch, S. *et al.* (2009) ‘Composition and three-dimensional architecture of the dengue virus replication and assembly sites.’, *Cell host & microbe*. United States, 5(4), pp. 365–375. doi: 10.1016/j.chom.2009.03.007.

WHO (2016) ‘Dengue vaccine: WHO position paper - July 2016.’, *Releve epidemiologique hebdomadaire*. Switzerland, 91(30), pp. 349–364.

WHO (2017) *Dengue and Severe Dengue*. http://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue.

WHO (2018a) *Epidemiology*. http://www.who.int/denguecontrol/epidemiology/en/ (Accessed: 7 March 2018).

WHO (2018b) ‘Immunization, Vaccines and Biologicals’. http://www.who.int/immunization/diseases/dengue/en/.

WHO (2018c) *Treatment*. http://www.who.int/csr/resources/publications/dengue/024-33.pdf?ua=1 (Accessed: 8 March 2018).

Winn, M. D. *et al.* (2011) ‘Overview of the CCP4 suite and current developments.’, *Acta crystallographica. Section D, Biological crystallography*. United States, 67(Pt 4), pp. 235–242. doi: 10.1107/S0907444910045749.

Winnebeck, E. C., Millar, C. D. and Warman, G. R. (2010) ‘Why Does Insect RNA Look Degraded?’, *Journal of Insect Science*. University of Wisconsin Library, 10, p. 159. doi: 10.1673/031.010.14119.

Wu, H. *et al.* (2015) ‘Novel Dengue Virus NS2B/NS3 Protease Inhibitors’, *Antimicrobial Agents and Chemotherapy* , 59(2), pp. 1100–1109. doi: 10.1128/AAC.03543-14.

Wu, J., Liu, W. and Gong, P. (2015) *A Structural Overview of RNA-Dependent RNA Polymerases from the Flaviviridae Family*, *International Journal of Molecular Sciences*. doi: 10.3390/ijms160612943.

Xi, Z., Khoo, C. and Dobson, S. (2005) *Ecology: Wolbachia establishment and invasion in an Aedes Aegypti laboratory population*, *Science (New York, N.Y.)*. doi: 10.1126/science.1117607.

Xie, X. *et al.* (2013) ‘Membrane Topology and Function of Dengue Virus NS2A Protein’, *Journal of Virology* , 87(8), pp. 4609–4622. doi: 10.1128/JVI.02424-12.

Yang, X. *et al.* (2017) ‘Triphosphate Reorientation of the Incoming Nucleotide as a Fidelity Checkpoint in Viral RNA-dependent RNA Polymerases.’, *The Journal of biological chemistry*. United States, 292(9), pp. 3810–3826. doi: 10.1074/jbc.M116.750638.

Yap, T. L. *et al.* (2007) ‘Crystal Structure of the Dengue Virus RNA-Dependent RNA Polymerase Catalytic Domain at 1.85-Angstrom Resolution’, *Journal of Virology* , 81(9), pp. 4753–4765. doi: 10.1128/JVI.02283-06.

Ye, Y. H. *et al.* (2015) ‘Wolbachia Reduces the Transmission Potential of Dengue-Infected Aedes aegypti’, *PLOS Neglected Tropical Diseases*. Public Library of Science, 9(6), p. e0003894. https://doi.org/10.1371/journal.pntd.0003894.

Yin, Z. *et al.* (2009) ‘An adenosine nucleoside inhibitor of dengue virus’, *Proceedings of the National Academy of Sciences*, 106(48), p. 20435 LP-20439. http://www.pnas.org/content/106/48/20435.abstract.

You, S. and Padmanabhan, R. (1999) ‘A Novel in Vitro Replication System for Dengue Virus: Initiation of RNA synthesis at the 3’-end of exogenous viral RNA templates requires 5’- and 3’-terminal complementary sequence motifs of the viral RNA’, *Journal of Biological Chemistry*, 274(47), pp. 33714–33722. doi: 10.1074/jbc.274.47.33714.

Yu, I.-M. *et al.* (2008) ‘Structure of the Immature Dengue Virus at Low pH Primes Proteolytic Maturation’, *Science*, 319(5871), p. 1834 LP-1837. http://science.sciencemag.org/content/319/5871/1834.abstract.

Da Yu Protein Sciences (2018) *Biophysical Characterisation*. http://www.dayuenterprises.com/Circular_Dichroism.html (Accessed: 24 February 2018).

Zavala-Castro, J. *et al.* (2014) *Decreased Mortality Rate Among Patients with Dengue Hemorrhagic Fever Treated with Doxycycline Associated With Reduced Inflammatory Cytokine Levels*. doi: 10.13140/2.1.4279.6801.

Zellweger, R. M., Prestwood, T. R. and Shresta, S. (2010) ‘Enhanced Infection of Liver Sinusoidal Endothelial Cells in a Mouse Model of Antibody-Induced Severe Dengue Disease’, *Cell Host & Microbe*, 7(2), pp. 128–139. doi: https://doi.org/10.1016/j.chom.2010.01.004.

Zhang, G. *et al.* (2013) ‘Wolbachia uses a host microRNA to regulate transcripts of a methyltransferase, contributing to dengue virus inhibition in Aedes aegypti’, *Proceedings of the National Academy of Sciences of the United States of America*, 110(25), p. 10276—10281. doi: 10.1073/pnas.1303603110.

Zhang, X. *et al.* (2013) ‘Cryo-EM structure of the mature dengue virus at 3.5-A resolution.’, *Nature structural & molecular biology*. United States, 20(1), pp. 105–110. doi: 10.1038/nsmb.2463.

Zhao, Y., Soh, T. S., Zheng, J., *et al.* (2015) ‘A Crystal Structure of the Dengue Virus NS5 Protein Reveals a Novel Inter-domain Interface Essential for Protein Flexibility and Virus Replication’, *PLOS Pathogens*. Public Library of Science, 11(3), p. e1004682. https://doi.org/10.1371/journal.ppat.1004682.

Zhao, Y., Soh, T. S., Lim, S. P., *et al.* (2015) ‘Molecular basis for specific viral RNA recognition and 2′-O-ribose methylation by the dengue virus nonstructural protein 5 (NS5)’, *Proceedings of the National Academy of Sciences*, 112(48), p. 14834 LP-14839. http://www.pnas.org/content/112/48/14834.abstract.

Zhou, X. *et al.* (2000) ‘Nitric Oxide Induces Thymocyte Apoptosis Via a Caspase-1-Dependent Mechanism’, *The Journal of Immunology*, 165(3), p. 1252 LP-1258. http://www.jimmunol.org/content/165/3/1252.abstract.

Ziehler, W. A. and Engelke, D. R. (2001) ‘Probing RNA Structure with Chemical Reagents and Enzymes’, *Current protocols in nucleic acid chemistry / edited by Serge L. Beaucage ... [et al.]*, 0 6, p. Unit-6.1. doi: 10.1002/0471142700.nc0601s00.

Zompi, S. and Harris, E. (2012) ‘Animal Models of Dengue Virus Infection’, *Viruses*. MDPI, 4(1), pp. 62–82. doi: 10.3390/v4010062.

Zou, J. *et al.* (2015) ‘Characterization of Dengue Virus NS4A and NS4B Protein Interaction’, *Journal of Virology* , 89(7), pp. 3455–3470. doi: 10.1128/JVI.03453-14.

Zug, R., Koehncke, A. and Hammerstein, P. (2012) *Epidemiology in evolutionary time: the case of Wolbachia horizontal transmission between arthropod host species*, *Journal of Evolutionary Biology*. doi: 10.1111/j.1420-9101.2012.02601.x.

Zuker, M. (2003) ‘Mfold web server for nucleic acid folding and hybridization prediction.’, *Nucleic acids research*. England, 31(13), pp. 3406–3415.

# Appendices

## Appendix I

### Buffers

**1 X TAE buffer**

40 mM Tris-Acetate and 1 mM ethylene-diamine-tetraacetic acid (EDTA)

**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

**SDS running buffer**

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

**Towbin buffer**

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

**Lysis buffer**

300 mM NaCl; 20 mM Tris- HCl pH 7.5; 0.5% v/v Triton X-100. 1 unit DNase (NEB) and a cOmplete Ultra protease inhibitor tablet (Roche).

**Binding buffer**

20 mM Tris- pH 7.5, 300 mM NaCl, 20 mM Imidazole and 1 x tablet protease inhibitors (cOmplete, Mini EDTA-free protease inhibitor cocktail (Roche).

**Gel filtration buffer**

300 mM NaCl and 20 mM Tris pH 7.5

**1 X SHAPE folding buffer**

100 mM HEPES pH 8.0, 100 mM NaCl and 10 mM $MgCl_2$

**0.5 X T.E buffer**

5 mM Tris and 0.5 mM EDTA, pH 7.5

**Buffer D**

(20 mM HEPES pH 7.4, 10 mM KCl, 2mM $MgCl_2$, 1 mM EDTA, 1 mM EGTA, 1 x tablet protease inhibitor (cOmplete, Mini EDTA-free protease inhibitor cocktail (Roche))

**TNMg**

(10 mM Tris-HCl (pH 8.0), 10 mM NaCl, 1.5 mM $MgCl_2$, 1 x complete, mini, EDTA-free protease inhibitor (Roche) and 2 µl (80 U) RNasin Rnase inhibitor (Promega))

## Stains

**Coomassie stain**

0.25% (w/v) Coomassie brilliant blue R-250, 50% (v/v) methanol and 10% (v/v) glacial acetic acid

**De-stain:**

40% (v/v) methanol, 10% (v/v) glacial acetic acid

## Media

**SOC**

autoclaved 2% w/v Tryptone, 0.5% w/v yeast extract, 10 mM NaCl, 2.5 mM KCl and sterilized 10 mM $MgCl_2$ and 20 mM glucose

**LB broth**

LB Broth; 10 g tryptone, 10 g NaCl, 5 g yeast extract, autoclaved in 1L $ddH_20$

## Appendix II

### Example *in vitrol* intracellular thermodynamic RNA structural predictions

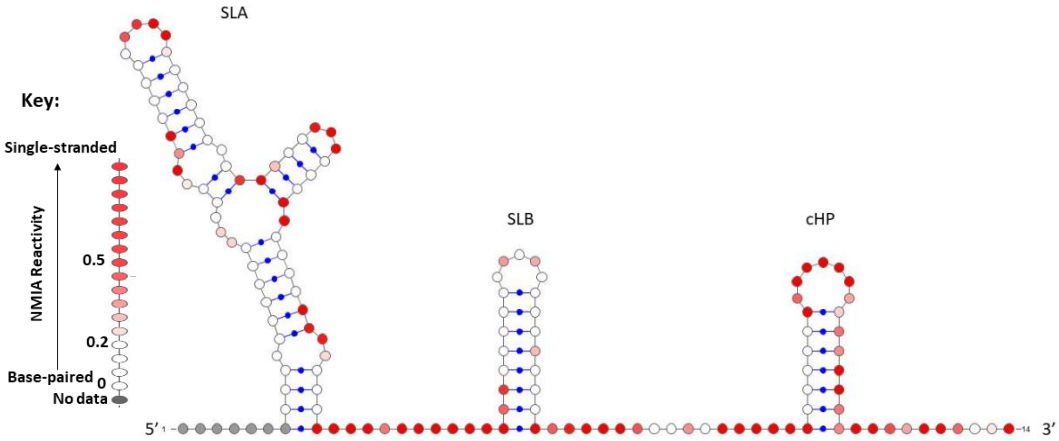



**Figure 91 *In vitro* SHAPE thermodynamically predicted RNA structure of full-length DENV-2 genomic RNA at mammalian physiological temperature in the absence of *trans*-activating factors**

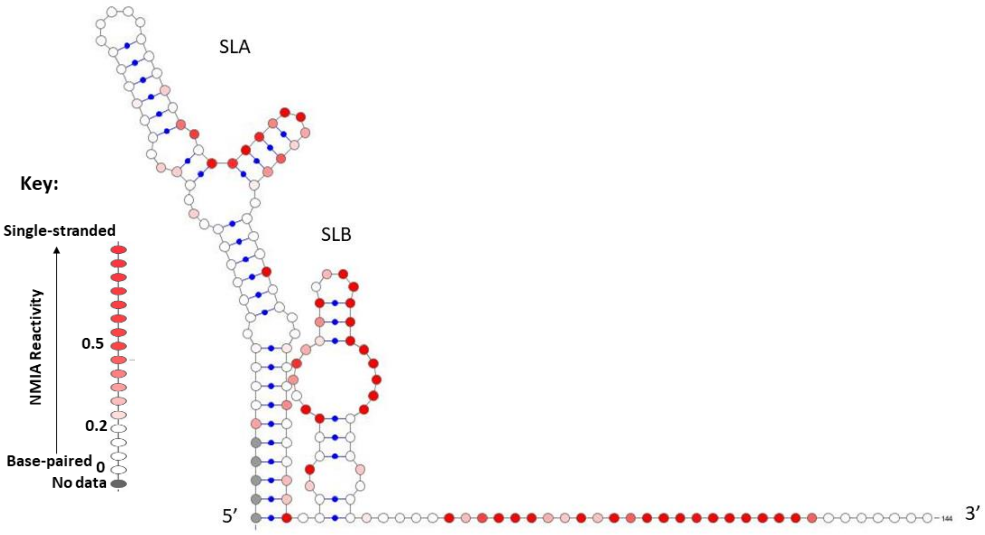



**Figure 92 Intracellular SHAPE thermodynamically predicted RNA structure of full-length DENV-2 genomic RNA extracted from the membrane fraction of stably expressing DENV-2 HEK293T cells at mammalian physiological temperature**