Establishing a model system to study CHIKV nsP4 and its host interactome during genome replication and translation.

Nuala Mairead Sweeney

Submitted in accordance with the requirements for the degree of Master of Science by Research

The University of Leeds

Faculty of Biological Sciences

School of Molecular and Cellular Biology

April 2025

I confirm that the work submitted is my own and that appropriate credit has been given where reference has been made to the work of others.

Acknowledgements

Firstly, I would like to thank my supervisor Dr Andrew Tuplin for all his help and support throughout this project and for giving me the opportunity to work with him. I would also like to thank Professor Mark Harris for his additional guidance during this project, and Professor Andres Merits for the kind gifts of the two antibodies and virus construct that made this research project a reality.

My thanks also go to the members of Tuplin lab, for all their kindness, support, and advice throughout this project. In particular, I would like to thank Zhaoxia, Yuqian and Sam for their work in BSL-3 that helped support this project.

I want to thank my parents for their unending support throughout the past year, Hibah, who has always been by my side commiserating and celebrating everything with me and, lastly, I would like to thank Ben for all his endless encouragement, love, and Tim Robinson quotes to help get me through everything.

Abstract

Chikungunya virus (CHIKV) is an arthropod-borne alphavirus, that causes large outbreaks that have high morbidity, due to arthralgia symptoms associated with chronic pain. CHIKV encodes four non-structural proteins (nsPs), including the RNA dependent RNA polymerase nsP4. Interactions of nsP4 with host cell proteins are poorly categorised, due to its rapid degradation *in vivo* and the insolubility of recombinantly expressed protein.

The aim of this study was to investigate interactions between nsP4 and host cell proteins during CHIKV replication and translation, using a sub-genomic replicon (SGR) system that encodes CHIKV nsPs but has a luciferase reporter gene in place of the structural viral proteins. Attempts to optimise an antibody against nsP4 for use in co-immunoprecipitation and quantitative mass spectrometry were hindered by the unreliable and unspecific nature of available nsP4 antibodies. Consequently, this study went on to develop a novel system to study nsP4 interactions in a CHIKV SGR expressing recombinant nsP4, labelled with N-terminal FLAG tags (termed nsP4-3XF SGR).

nsP4-3XF SGR was successfully engineered by cloning the sequence of a tagged nsP4 from an existing recombinant virus into the SGR, which was then validated for replication efficiency and nsP4-3XF expression. Expression of nsP4-3XF decreased the level of SGR replication in comparison to the wildtype SGR, but replicated to a sufficient level that it could be used for further study of nsP4 interactions. Western blotting demonstrated that nsP4-3XF could be detected in transfected BHK-21 cell lysates but not in RD, Huh7 or C6/36 cells. Preliminary co-immunoprecipitation assays from transfected BHK-21 cells confirmed purification of nsP4-3XF in complex

with host cell proteins HSP90 and TMEM45B which have previously been described as interactors of CHIKV nsP4. As such, the preliminary data demonstrated that nsP4-3XF SGR, engineered and validated by this study will be an invaluable system for studying the CHIKV nsP4 proteome complex during CHIKV replication and translation.

Table of Contents

List of Tables and Figures

Page number

Tables

Figures

List of abbreviations

A – alanine

Arbovirus – arthropod borne virus

AUD – alphavirus unique domain

BHK – baby hamster kidney

bp - base pair

C protein – capsid protein

CIP – calf intestinal phosphate

CHIKV – Chikungunya virus

CHOP - C/EBP homologous protein

Co-IP – co-immunoprecipitation

DEPC - diethyl pyrocarbonate

DENV – dengue virus

DMEM - Dulbecco's Modified Eagle Medium

DNA – deoxyribonucleic acid

E (1-3) protein – envelope protein

ECSA – Eastern, Central, Southern African CHIKV genotype

EEEV – Eastern equine encephalitis virus

eIF2α - eukaryotic initiation factor α

FBS – foetal bovine serum

GFP – green fluorescent protein

GLB – Glasgow lysis buffer

HSP (70/90) – heat shock protein

hpi – hours post infection

hpt – hours post transfection

HVD – hypervariable domain

IFA – immunofluorescent assay

IOL – Indian Ocean lineage

IP (lysis buffer) – immunoprecipitation (lysis buffer)

mRNA – messenger RNA

nsP (1-4) – non-structural protein

nsP (123/1234/23) – non-structural polyprotein

nsP4-3XF – non-structural protein 4 (with Strep II and three FLAG tags) in CHIKV SGR

nsP4-FLAG – non-structural protein 4 (with Strep II and three FLAG tags) in infectious virus

ONNV – O'nyong'nyong virus

ORF (-1/-2) – open reading frame

PBS – phosphate buffered saline

pE2 – precursor envelope protein 2

PERK – PKR-like endoplasmic reticulum kinases

PCR – polymerase chain reaction

PLB – passive lysis buffer

R – arginine

RC – replication complex

RdRp – RNA dependent RNA polymerase

RNA – ribonucleic acid

RRV – Ross River virus

S – serine

Spp. – species

SGR – sub-genomic replicon

SINV – Sindbis virus

SFV – Semliki Forest virus

TATase – terminal adenylyltransferase

TBS(T) – tris buffered saline (tween)

TF – trans frame protein

TMT-MS – tandem mass tag mass spectometry

UTR – untranslated region

UPR – unfolded protein response

V – valine

VEEV – Venezuelan equine encephalitis virus

WEEV – Western equine encephalitis virus

WT – wild type

ZIKV – Zika virus

Chapter 1: Introduction

1.1 Introduction to alphaviruses

Alphaviruses are a genus of positive RNA viruses and the only member of the *Togaviridae* family **(Jain et al., 2016).** There are thirty-two recognised species in the alphavirus genus, the majority of which are arthropod borne viruses (arboviruses) **(Chen et al., 2018).** Many of these arboviruses are transmitted by mosquitoes, and of concern to human health. Non-human alphaviruses include the mosquito specific Eliat virus, and several salmonid alphaviruses for which no vectors have been identified, although they are capable of replicating in mosquito cell lines *in vitro* **(Lello et al., 2021).** Additionally, the alphavirus Southern elephant seal virus is transmitted via sea lice, demonstrating the wide range of arthropod vectors capable of transmitting alphaviruses **(Hyde et al., 2015)**

Alphaviruses can be categorised in one of three ways; by geographic origin and symptoms (Old World and New World alphaviruses), into complexes antigenically, or into clades by phylogenetics **(Van Huizen and McInerney, 2020).** Old World alphaviruses originated in Africa and Asia and are commonly associated with rash, febrile illness and arthralgia, while New World alphaviruses originated in the Americas, and are more commonly associated with neurological symptoms such as encephalitis **(Bakovic et al., 2021).** Old World alphaviruses with relevance to human health include Semliki Forest virus (SFV), Sindbis virus (SINV), Ross River virus (RRV) O'nyong'nyong virus (ONNV) and Chikungunya virus (CHIKV), while New World alphavirus members include Venezuelan, Western and Eastern equine encephalitis viruses (VEEV, WEEV and EEEV). Alphaviruses can also be separated into complexes antigenically, including the SFV complex (SFV, CHIKV, Mayaro virus,

ONNV, RRV), VEEV complex (VEEV, Everglades virus) and WEEV complex (WEEV and SINV) **(Lello et al., 2021).** Despite similarities between members of the same groups, significant differences in the biology of alphavirus species exist, even in closely related members of the genus.

Mosquito-borne alphaviruses are largely transmitted by mosquitoes of the *Aedes* genus. *Aedes albopictus* and *Aedes aegypti* are the main vectors of concern in terms of pathogenic alphaviruses **(Bustos Carrillo et al., 2019).** As alphaviruses are spread by mosquitoes, they are mainly found in tropical and subtropical regions where these vectors are endemic, though cases of local alphavirus transmissions have been reported in the more temperate regions of the United States and in European countries such as Italy and France **(Foy et al., 2013)**. Encroachment of human populations into areas where reservoirs for alphaviruses are present has increased the incidence of disease, due to the increased risk of spillover events. Increased globalisation and travel can also be linked to increased incidence of alphavirus disease, as viruses can spread to naïve areas through infected travellers. Additionally, climate change has the potential to expand the range of the mosquitoes implicated in alphavirus transmission, thereby introducing the viruses to new geographical regions **(Pohjala et al., 2011).**

Despite their relevance to human health, no specific antivirals are available to treat alphavirus infections **(Metibemu et al., 2024).** Consequently, there is a need to study alphaviruses to elucidate the mechanisms behind their ability to cause disease and to develop new specific treatments. SFV and SINV have been widely used as models for Old World alphaviruses, but as different alphaviruses interact with their hosts in a diverse range of ways, biological processes can vary widely throughout the genus **(Van Huizen and McInerney, 2020).** Differences between the biology of

alphavirus species include entry receptors, methods of modulating cellular pathways during infection and the effects of protein interactions in different species **(Carey et al., 2019, Fros and Pijlman, 2016, Rathore et al., 2013, Zimmerman et al., 2023)**. Therefore, there is a need to study a range of alphavirus species that are relevant to human diseases.

1.2.1 Chikungunya virus

Chikungunya fever, the resulting disease from CHIKV infection, was first reported in 1952 during an outbreak in Tanzania **(Müller et al., 2019).** The virus was identified and categorised the following year from blood samples of individuals affected by the outbreak **(Joubert et al., 2012).** Previously, the virus was not seen as a risk for causing major outbreaks, but the more recent re-emergence of CHIKV has resulted in epidemics on a greater magnitude than previously seen and underscored the potential risk of CHIKV to human health **(Li et al., 2013)**. Since its re-emergence in 2004, CHIKV is now endemic in several areas that formerly only experienced sporadic outbreaks, such as India which has been endemic for the virus since a 2006 outbreak that infected over 1.3 million people **(Pareek et al., 2022)**. In addition to endemic circulation in certain regions, CHIKV is the only alphavirus that causes regular epidemics across the globe, with outbreaks often infecting large proportions of the population in the affected area. Alongside the high morbidity associated with CHIKV, impacting the communities it affects socially and financially, the increasing risk of CHIKV globally has resulted in increased concern.

CHIKV is transmitted to mammalian hosts through the bite of an infected female *Aedes spp.* mosquito **(Matkovic et al., 2019)**. The virus is zoonotic and is maintained in both sylvatic and urban cycles, with the sylvatic cycle acting as a

reservoir for the virus with the potential to re-emerge into wider human populations **(Figure 1) (Li et al., 2013).** Sylvatic cycles are maintained between non-human primates and *Aedes spp.* mosquitoes, though the mosquitoes involved in this cycle typically differ from those prevalent in urban cycles **(Zeller et al., 2016)**. In Africa, sylvatic cycles are present in forest or savannah areas where non-human primates are found. In Asia, CHIKV is believed to be primarily maintained mainly in urban cycles, though non-human primates with CHIKV antibodies have been identified in some areas **(Tsetsarkin and Weaver, 2011).** In addition to non-human primates, CHIKV has also been isolated from other mammals, such as bats, which may potentially act as reservoirs **(Zeller et al., 2016).**

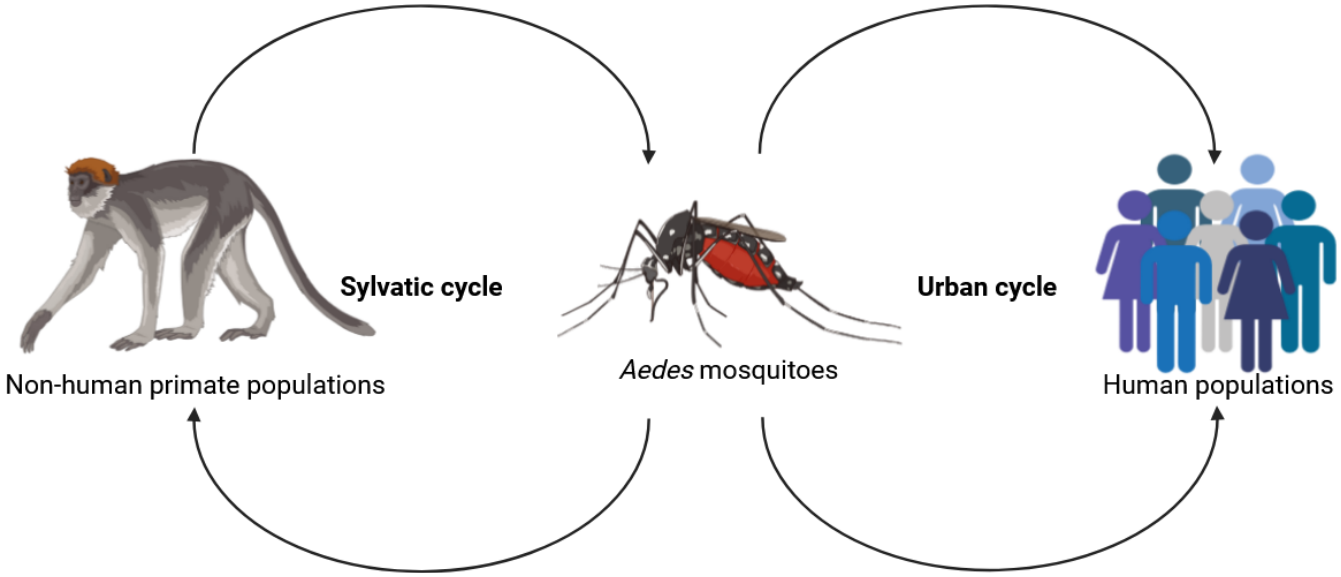


Figure 1: Schematic representing the sylvatic and urban cycles of CHIKV. *CHIKV is transmitted through female Aedes mosquitoes when taking a bloodmeal. Sylvatic cycles are maintained between non-human primates and mosquitoes. Spillover events leading to transmission of CHIKV to human populations through the bite of a CHIKV infected mosquito can result in large outbreaks. Urban cycles are maintained by humans and mosquitoes in urban environments and allow CHIKV to endemically infect human populations* ***(Adapted from (Singh and Unni, 2011))****.*

1.2.2 Pathogenesis of CHIKV

While CHIKV induces a highly cytopathic effect in mammalian cells, resulting in acute infection, the virus is asymptomatic in mosquitoes with a comparatively moderate cytopathic effect **(Li et al., 2013).** This allows the virus to persistently infect mosquitoes, which can then transmit the disease to new mammalian hosts through a bloodmeal.

Though associated with low mortality, Chikungunya fever is often associated with high morbidity **(Hapuarachchi et al., 2010).** Alongside the febrile symptoms of Chikungunya fever, which include fever, rash and fatigue, the virus can cause severe joint and skeletomuscular pain, which can persist for months or years **(Echavarria-Consuegra et al., 2023)**. In fact, the word Chikungunya can be translated to mean "bends upwards" from Makonde, the local language of Tanzania where the disease was first recorded. This name refers to the contorted figures of sufferers of the disease, which result from the painful arthritic symptoms **(Zeller et al., 2016).**

The incubation period of CHIKV from the initial mosquito bite to first disease symptoms is between 1 and 12 days **(Zeller et al., 2016)**. An average of 20% of infections are asymptomatic, though these can still be highly viraemic and infect mosquitoes during bloodmeals **(Freire et al., 2022)**. CHIKV initially infects dermal fibroblasts before spreading via the bloodstream to several different tissues including the liver, joints and muscle **(Schwartz and Albert, 2010).**

Though most symptomatic cases see improvements after the first two weeks of disease onset, the arthritic symptoms can become chronic, causing long term morbidity, even after acute illness has passed. Joint and skeletomuscular symptoms can last months or years after initial infection, and mainly affect the peripheral joints **(Carvalho et al., 2017).** Up to 50% of symptomatic infections will develop chronic

symptoms **(Delang et al., 2014)** and result in high morbidity for the patient, affecting both their quality of life and ability to work. Children suffering from the disease are additionally likely to miss long periods of school **(Rezza and Weaver, 2019)**. The wider communities of those infected with CHIKV can therefore be affected by outbreaks both socially and economically for years after the initial outbreak due to the chronic nature of the disease. Chronic symptoms can further impact and individual's financial situation by requiring drugs for pain treatment and additional support care, in addition to a reduced ability to work **(Rezza and Weaver, 2019).** The exact cause of chronic arthralgia symptoms remains unclear but may be linked to the severity of the initial acute infection **(Wada et al., 2017).** Despite the long-term suffering CHIKV can cause, the disease is rarely fatal, with an estimated mortality rate of approximately 1 in 1000 infections **(Rathore et al., 2013),** however the death toll may be underreported due to the disease mainly affecting low and middle income countries **(Rezza and Weaver, 2019)**. CHIKV is often misdiagnosed due to similarities to other mosquito-borne illnesses such as dengue **(Tiozzo et al., 2025).** Severe symptoms such as encephalitis are possible, and mainly affect the elderly or very young children **(Wada et al., 2017).** Severe complications of CHIKV infection in pregnant women can additionally result in miscarriage, death in early infancy or neurological issues in the child **(Rezza and Weaver, 2019)**.

1.2.3 CHIKV lineages and epidemiology

There are three CHIKV genotypes that have been categorised by phylogenetic analysis; West African, Asian, and Eastern, Central, South African (ECSA) **(Paingankar and Arankalle, 2014).** The Asian lineage diverged from the West African between 100-150 years ago, likely due to an infected traveller moving to a previously unaffected area, creating a bottleneck effect **(Hyde et al., 2015).** An

additional lineage of the ECSA genotype known as the Indian Ocean Lineage (IOL) has also been identified as a specific cause for concern, due to its role in several outbreaks since the re-emergence of CHIKV.

The first recorded outbreak of CHIKV in Tanzania had an estimated rate of incidence of 23% in the population between July 1952 and May the following year **(Zeller et al., 2016).** Though sporadic outbreaks of CHIKV by West African and Asian lineages occurred between its first discovery in 1952 and its re-emergence in the mid-2000s, the disease was not believed to be of great concern due to its low mortality and the small scale of these outbreaks **(Simon et al., 2011)**. The first major outbreak of CHIKV was identified on Lamu Island off the coast of Kenya in 2004, infecting an estimated 70% of the island's population at its peak **(Zeller et al., 2016).** This strain of CHIKV spread along the coast and to several other islands including La Reunion, causing a large outbreak that infected an estimated third of the overall population and resulted in a high incidence of severe symptoms such as encephalitis **(Krejbich-Trotot et al., 2011).** The virus then disseminated to several other tropical areas including Southeast Asia and the Indian Ocean islands. This outbreak was the first time the new clade of ECSA (IOL) was detected **(Abraham et al., 2015).**

Increased travel and globalisation have also contributed to CHIKV's dissemination to new areas, with the virus becoming endemic in several regions where it had not been detected prior to its re-emergence in 2004. For example, CHIKV was first confirmed in the Americas at the beginning of the 2010s, with the first case ever reported in the Caribbean in 2013 **(Zeller et al., 2016).** Within the next year, CHIKV cases had been reported throughout South and Central America, with autochthonous cases confirmed as far as the Southern United States. Over 1.2 million autochthonous cases were reported throughout the Americas between 2013 and

2014 **(Zeller et al., 2016),** though this had reduced to 101,570 cases in 2020 **(Reyes-Gastellou et al., 2021).** This exemplifies the rapid rate at which CHIKV can spread through naïve populations. Additionally, CHIKV was first reported in Europe in 2007, with 217 cases detected in northeastern Italy. These isolates were identified as the IOL and were found circulating in local *Ae. albopictus* populations. Limited outbreaks have additionally been reported in France and other southern European countries such as Croatia **(Simon et al., 2011). Figure 2** demonstrates the distribution of CHIKV in 2020, with cases of CHIKV in over a hundred countries.

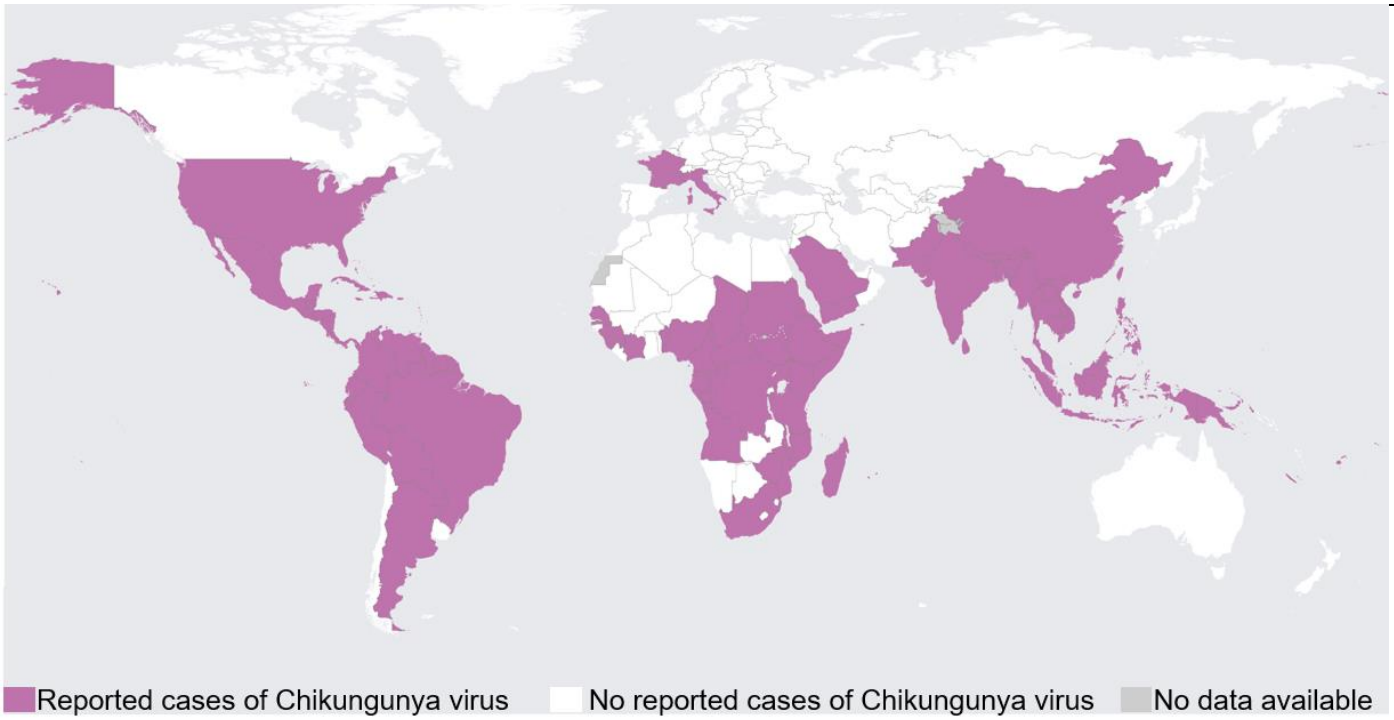


Figure 2: Map representing the global distribution of CHIKV. *Countries that have reported cases of local CHIKV transmission are shown in purple, while countries that have no confirmed cases of autochthonous CHIKV cases are shown in white* ***(Adapted from (WHO, 2020).***

1.2.4 Increased fitness of the Indian Ocean Lineage in *Ae. albopictus*

One main difference between the IOL outbreak in the mid-2000s and previous outbreaks of CHIKV was a difference in the vector transmitting the disease. Previous outbreaks were mainly transmitted by *Ae. aegypti* vectors **(Goertz et al., 2018)** while the La Reunion outbreak was mainly transmitted by *Ae. albopictus* mosquitoes **(Zeller et al., 2016).** IOL has increased replication efficiency in *Ae. albopictus* vectors due to a mutation in the sequence of the E1 protein, which changed the alanine at position 226 to a valine, resulting in an increase in fitness in these mosquitoes. Virus entry into cells is often mediated by cholesterol **(Carey et al., 2019),** with the rate of CHIKV fusion previously reported as being correlated with cholesterol concentration in a cell **(Sousa et al., 2020)**. The A226V mutation results in less dependence on cholesterol, as position 226 is located close to the peptide fusion region of E1 that interacts with cholesterol during entry **(Figure 3) (Zeller et al., 2016).** As A226V decreased E1's dependence on cholesterol, it increased the efficiency of CHIKV infection in *Ae. albopictus* cells, whose membranes have a low cholesterol composition **(Schwartz and Albert, 2010)**. The increased efficiency in CHIKV entry consequently led to an increase in replication, with higher virus titres detected in mosquito salivary glands in IOL isolates than in strains encoding wildtype (WT) E1 **(Zeller et al., 2016).** This in turn increased viral dissemination and transmission, causing wider scale outbreaks than had previously been seen in CHIKV **(Wada et al., 2017).**

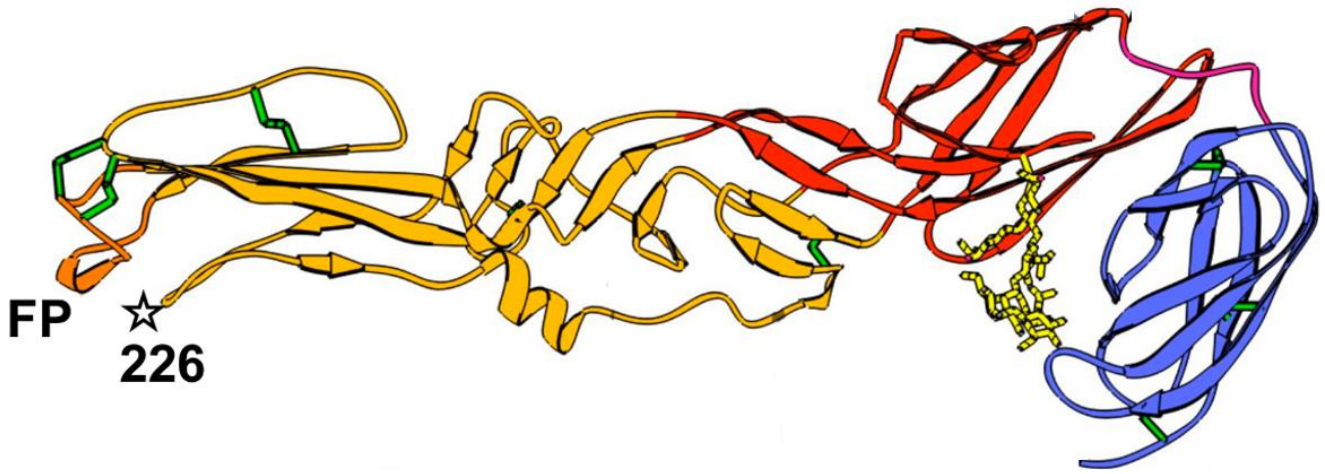


Figure 3: 3D schematic representation of CHIKV E1 protein**. Predicted CHIKV E1 structure, modelled using data of the CHIKV E1 sequence and the crystal structure of SFV E1. Fusion peptide that interacts with cholesterol during entry is coloured orange and labelled as FP. Proximity of position 226 (labelled as a star) where alanine was mutated to a valine in IOL isolates is demonstrated. **Adapted from (Schuffenecker et al., 2006).

The A226V mutation that increased the fitness of CHIKV in *Ae. albopictus* has had wider implications for the transmission of the virus. *Ae. albopictus* has a larger geographic range than *Ae. aegypti* and more aggressive biting behaviour. Despite A226V increasing CHIKV's fitness in *Ae. albopictus* mosquitoes, the mutation has no negative effect on fitness in *Ae. aegypti* **(Kril et al., 2021)**. With climate change potentially expanding the range of both major mosquito vectors into regions that are currently more temperate, the potential of CHIKV to spread to previously unaffected areas has increased.

Isolates from an outbreak of IOL in Malaysia were reported to encode another beneficial adaptation for CHIKV **(Fu et al., 2019)**. Unlike previously described mutations of CHIKV, which were categorised in structural proteins, the isolates from this 2008 outbreak were found to encode a mutation in non-structural protein (nsP)

4. This mutation substituted the arginine at position 83 of nsP4 to a serine and resulted in an increase of virus fitness in mammalian cells with no observable effect in mosquito cells **(Fu et al., 2019)**. Though the frequency of nsP4 R83S rose throughout the population, isolates of the IOL encoding WT nsP4 were still being reported in 2013, indicating that this mutation did not completely outcompete the WT strain **(Fu et al., 2019)**. The mutation was located in the N-terminus of nsP4, where the protein binds to other nsPs and host cell proteins. As the mutation only enhanced infection in mammalian cells, this may indicate that the adaptation is beneficial when interacting with mammalian host cell protein.

1.2.5 Treatments and Prevention of CHIKV

While several compounds have been found to inhibit CHIKV *in vitro,* there are no specific commercially available antivirals for CHIKV, or indeed any other alphavirus, and so treatment of Chikungunya fever is focused on relieving the symptoms via painkillers and anti-inflammatories to limit arthritic symptoms **(Echavarria-Consuegra et al., 2023)**. A live attenuated vaccine Ixchiq was approved in February 2024 for clinical use by the FDA in the United States, for people over the age of 18 that are at risk of CHIKV infection **(Weber et al., 2024).** However, this vaccine has yet to be implemented in the most high-risk countries where the disease is endemic, and adverse effects after receiving the vaccine have been reported, with up to a third of people reporting mild symptoms such as headache or fatigue and 1.9% reporting severe symptoms **(Schneider et al., 2023).**

Control of mosquito populations by methods such as insecticides, which can target both larval and adult mosquitoes, or removal of standing water in urban areas, which removes mosquito breeding grounds, can limit the spread of the disease **(Freire et**

al., 2022). Personal bite prevention strategies such as the use of mosquito nets and mosquito repellents or wearing clothing that covers areas such as arms and legs can be employed to protect individuals from mosquito bites **(Hucke et al., 2021).**

Additionally, the release of genetically modified mosquitoes, such as sterile male mosquitoes, has been employed in several areas including Australia, Mexico and Vietnam to limit mosquito populations and consequently the transmission a range of mosquito-borne diseases, including CHIKV **(Nguyen et al., 2015) (Sanchez-Aldana-Sanchez et al., 2023).** The introduction of *Wolbachia* infected males can limit the number of offspring being produced by a mosquito population due to cytoplasmic incompatibility, in which the infected males inhibit egg development in uninfected females **(Tortosa et al., 2008)**. However, the release of genetically modified or *Wolbachia* infected mosquitoes can be limited by the cost of rearing the insects, in addition to negative public perceptions of their release into the environment **(Hucke et al., 2021).**

Though vector control can prevent new infections by interrupting the lifecycle from vector to mammalian host, these methods do nothing to support individuals who are already suffering from acute or chronic symptoms of CHIKV infection. Therefore, there is a need to identify targets for antiviral treatment, in the hopes of treating the debilitating symptoms of CHIKV.

1.3.1 CHIKV genome organisation

CHIKV is an enveloped, positive sense, single stranded RNA virus **(Fox and Diamond, 2016)**. The genome of CHIKV is ~11.8 kb in length, unsegmented, and is composed of two open reading frames (ORFs), which are separated by a non-coding intergenic junction region **(Khan et al., 2002)**. Two untranslated regions (UTRs) flank

the genome, with a poly(A) tail present at the 3' end and a type-0 -m7GpppG cap at the 5' of the genome **(Figure 4) (Jungfleisch et al., 2022)**.

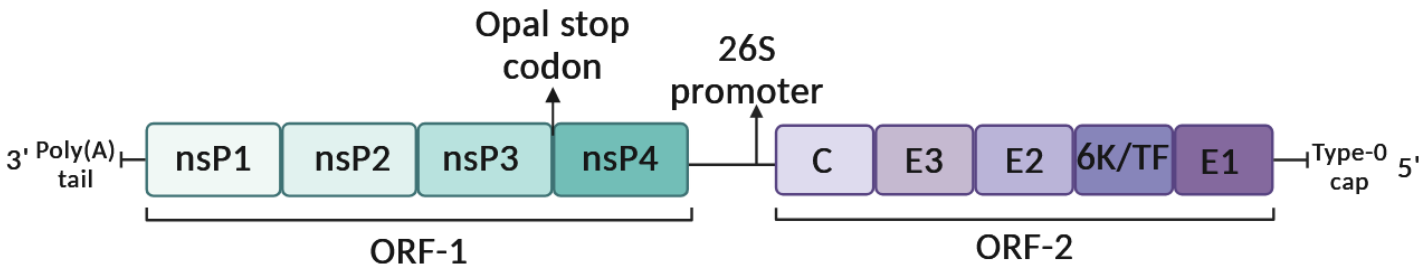


Figure 4: Schematic representation of the CHIKV genome. *The CHIKV genome is ~11.8 kb in length, containing two ORFs. ORF-1 encodes four nsPs. An opal stop codon present in some CHIKV isolates is found upstream of the cleavage site between nsP3 and nsP4. The 26S promoter in the intergenic region allows the expression of sub-genomic RNA, from which ORF-2 can be translated from. ORF-2 encodes five structural proteins, Capsid protein (C), three envelope proteins (E1-3) and 6K or Trans frame (TF) protein which is expressed as a result of translational frameshifting. Poly(A) tail and type-0 cap flank the genome at the 3' and 5' respectively.*

CHIKV's ORFs are translated from the genome as polyproteins, which are then processed by proteolytic cleavage to produce their respective mature proteins **(Figure 5) (Van Huizen and McInerney, 2020).** The 5' ORF (ORF-1) encodes the nsPs and makes up approximately two thirds of the overall genome **(Cristea et al., 2006),** while the 3' ORF (ORF-2) encodes the viral structural proteins **(Martins et al., 2023).**

ORF-1 is translated directly from the genome, expressing four nsPs **(Foy et al., 2013)**, all of which are involved in replicating the viral genome by forming the replication complex (RC) **(Rausalu et al., 2016)**. Translation of the nsPs as a polyprotein is essential for the formation of the RC and for guiding the localisation of

the RC during the initial replication stages **(Lello et al., 2021).** An opal stop codon is present in some isolates of CHIKV, six codons upstream of the nsP3/nsP4 cleavage site. This codon prevents the translation of nsP4 except in the event of translational readthrough **(Figure 5) (Li et al., 2023).** ORF-2 is translated from sub-genomic RNA, which is synthesised from the negative strand RNA intermediate, and so is translated later in the replication cycle. Structural proteins are translated from ORF-2 to a far higher concentration than nsPs **(Van Huizen and McInerney, 2020).** As with nsPs, structural proteins are translated as a polyprotein and then proteolytically cleaved **(Abdelnabi et al., 2015).** The structural proteins expressed include capsid (C) protein, envelope proteins E1, E2 and E3, and 6K protein **(Issac et al., 2014).** Structural proteins form the new virion, containing a positive sense genome**,** which buds from the host cell as infectious progeny virions **(Garmashova et al., 2007).**

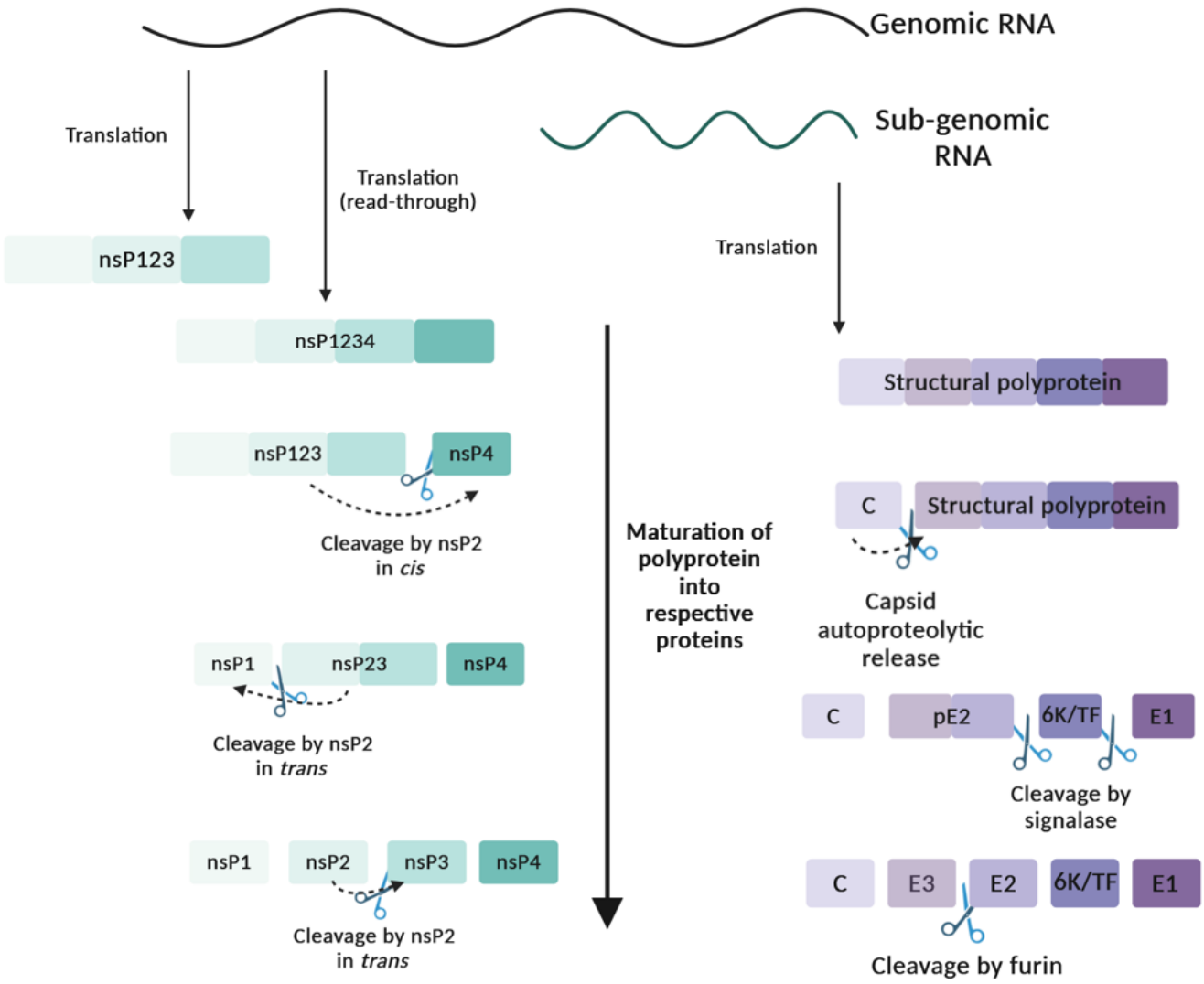


Figure 5: Schematic representation of the translation of ORFs into polyproteins and cleavage into individual proteins. *nsPs are translated from genomic RNA, with translation producing nsP123 and nsP1234 by readthrough. nsP4 is first cleaved from the polyprotein, followed by nsP1, then cleavage between nsP2 and nsP3. Structural proteins are translated from sub-genomic RNA. Capsid is first autocleaved, 6K and E1 are cleaved by signalase in the ER, followed by pE2 cleavage by furin.* ***Adapted from (Kril et al., 2021).***

1.3.2 nsP1

nsP1 is the first protein encoded by the CHIKV genome **(Figure 4)**. It is a ~60 kDa protein and possesses methyltransferase and guanylyl transferase domains that are

responsible for capping the viral genome at the 5' end **(Freire et al., 2022)**. Amino acid residues conferring capping activity are conserved across the alphavirus genus, as the presence of the type-0 cap is critical for ORF-1 translation **(Prosser et al., 2023).** In addition to capping CHIKV genomic RNA, nsP1 anchors the RC to the host plasma membrane during genome replication and directs other nsPs to the membrane **(Matkovic et al., 2019)**. Membrane association is thought to be mediated by palmitoylation of cysteine residues within the nsP1 C-terminus, which more closely associates the RC to the membrane **(Carey et al., 2019)**. Palmitoylation is critical in CHIKV, with deletions of palmitoylated nsP1 residues completely inhibiting replication of the virus, as palmitoylation is involved in forming the correct interactions of nsP1 with other nsPs, in addition to binding to the membrane **(Utt et al., 2019)**. nsP1 binding to the plasma membrane is known to play a role in spherule production, with attachment of nsP1 triggering rearrangement of the plasma membrane **(Kril et al., 2021).**

1.3.3 nsP2

nsP2 is a ~90 kDa protein possessing RNA helicase, phosphatase, and protease activity **(Matkovic et al., 2019)**. The N-terminus of nsP2 possesses the helicase and nucleoside triphosphatase activity **(Skidmore and Bradfute, 2023)**, in addition to RNA phosphatase which initiates capping of the viral genome by removing gamma-phosphate at the 5' of the viral RNA **(Carey et al., 2019).** The C-terminus possesses cysteine protease, which is responsible for processing of the nsP polyprotein by cleavage. nsP2 additionally suppresses host cell transcription in Old World alphaviruses (including CHIKV) by removing the host RNA polymerase II catalytic subunit Rpb1, which stops host mRNAs from being transcribed while viral protein synthesis continues **(Carey et al., 2019).**

1.3.4 nsP3

nsP3 is an ~60 kDa protein and is essential for genome replication **(Verma et al., 2023).** nsP3 helps to drive the assembly of the RC, contains ADP-ribose binding and hydrolase domains **(Martin et al., 2024)**, and interacts with a range of host cell proteins **(Freire et al., 2022)**. This results in some nsP3 being sequestered to the membrane for involvement in replication, while other nsP3 is distributed throughout the cytoplasm to participate in nsP3-host protein interactions.

nsP3 has three domains: the alphavirus unique domain, the macrodomain and the hypervariable domain (HVD). The alphavirus unique domain and macrodomain are both found in the N-terminus of nsP3 and are well conserved across the genus, while the HVD in the C-terminus has high sequence variation **(Dominguez et al., 2021)**. This allows the HVD, which is involved in protein-protein interactions, to bind to a multitude of host protein motifs **(Goertz et al., 2018).** Some nsP3-host protein interactions are important for efficient replication of CHIKV, for example disruption of the interaction between nsP3 and G3BP1 and G3BP2 inhibits CHIKV replication **(Kril et al., 2021).**

1.3.5 nsP4

CHIKV nsP4 is an ~70 kDa protein that acts as the RNA dependent RNA polymerase (RdRp) for CHIKV, transcribing new viral RNA **(Thoka et al., 2018)** As the RdRp, nsP4 is the most highly conserved protein across the alphavirus family **(Ahola and Merits, 2016).** nsP4 is the first protein cleaved by nsP2 from the nsP1234 polyprotein, so that the initial RC can be formed **(Figure 5) (Spuul et al., 2011).**

In addition to polymerase activity, nsP4 has terminal adenylytransferase (TATase) activity which maintains the viral genome poly(A) tail **(Freire et al., 2022)**. TATase

function is essential for replication, as a conserved 3' sequence element that is critical for replication exists next to adenylate residues of the poly(A) tail, therefore, the poly(A) tail must be maintained and is repaired by nsP4 to ensure replication can occur **(Pietila et al., 2017).** TATase and polymerase domains are contained in the C-terminus, while the 150 amino acid long N-terminus is largely disordered for interactions with viral and host cell proteins **(Pareek et al., 2022)**. While the N-terminus domain of nsP4 is conserved across the alphavirus genus, it has low homology with other virus families **(Chen et al., 2017).**

1.3.6 Capsid

C protein is the first protein encoded by ORF-2 **(Figure 4).** The C-terminus contains the protease activity for autoproteolytic release of C protein from the structural polyprotein **(Carey et al., 2019)**. This leads to free C protein being released into the cytosol and the exposure of the endoplasmic reticulum (ER) localisation site on the remaining polypeptide, where it will be processed. The C-terminus of C protein is highly conserved across the alphavirus genome, while the N-terminus has high variability **(Carey et al., 2019)**. The N-terminus binds to the viral RNA during assembly, leading to encapsidation of the viral genome. In addition to encapsidation, C protein also plays a role in regulating protein synthesis during infection **(Singh et al., 2018).**

1.3.7 Envelope protein

CHIKV has three envelope proteins known as E1, E2 and E3. E1 and E2 form the glycoprotein spikes in the mature virion, of which there are 80 spikes, each consisting of three E1-E2 heterodimers **(Freire et al., 2022)**. E1 is involved in the fusion of the virus with the cell membrane during entry, while E2 is involved in

anchoring the virion to the cell surface and interacting with host entry receptors **(Lee and Chu, 2015)**. E2 is initially expressed as a precursor protein with E3 (pE2, **Figure 5) (Carey et al., 2019)**. This precursor protein is required for the correct formation of the mature glycoprotein spikes, preventing the disassociation of E1 and E2, as well as hiding the E1 fusion loop when the virus is not fully matured, preventing premature fusion at the cell surface **(Singh et al., 2018)**. E3 is not incorporated into mature CHIKV virions.

1.3.8 6K/Trans-frame protein

6K is a small protein of ~6 kDa **(Carey et al., 2019)**, that is known to be involved in several mechanisms of CHIKV infection, including virus budding and assembly. The protein is involved in regulating host ion channels through interactions of its N-terminus, and has additionally been reported to aid the budding of new virions from the cell by permeabilising the host cell plasma membrane **(Singh et al., 2018).** Ribosomal frameshifting during the translation of ORF-2 gives rise to the production of the trans-frame (TF) protein **(Carey et al., 2019)**. TF protein is also a very small protein of approximately 8 kDa. The N-terminus of TF protein and 6K is identical, but the C-terminus is modified as a result of frameshifting **(Singh et al., 2018)**. Like 6K, TF protein is involved in assembly and the management of host cell ion channels. Despite their role in the viral lifecycle, deletion of 6K and TF do not completely inhibit CHIKV replication, though viral titre of infected cells is reduced when 6K and TF proteins are not present **(Skidmore and Bradfute, 2023).**

1.4 CHIKV lifecycle

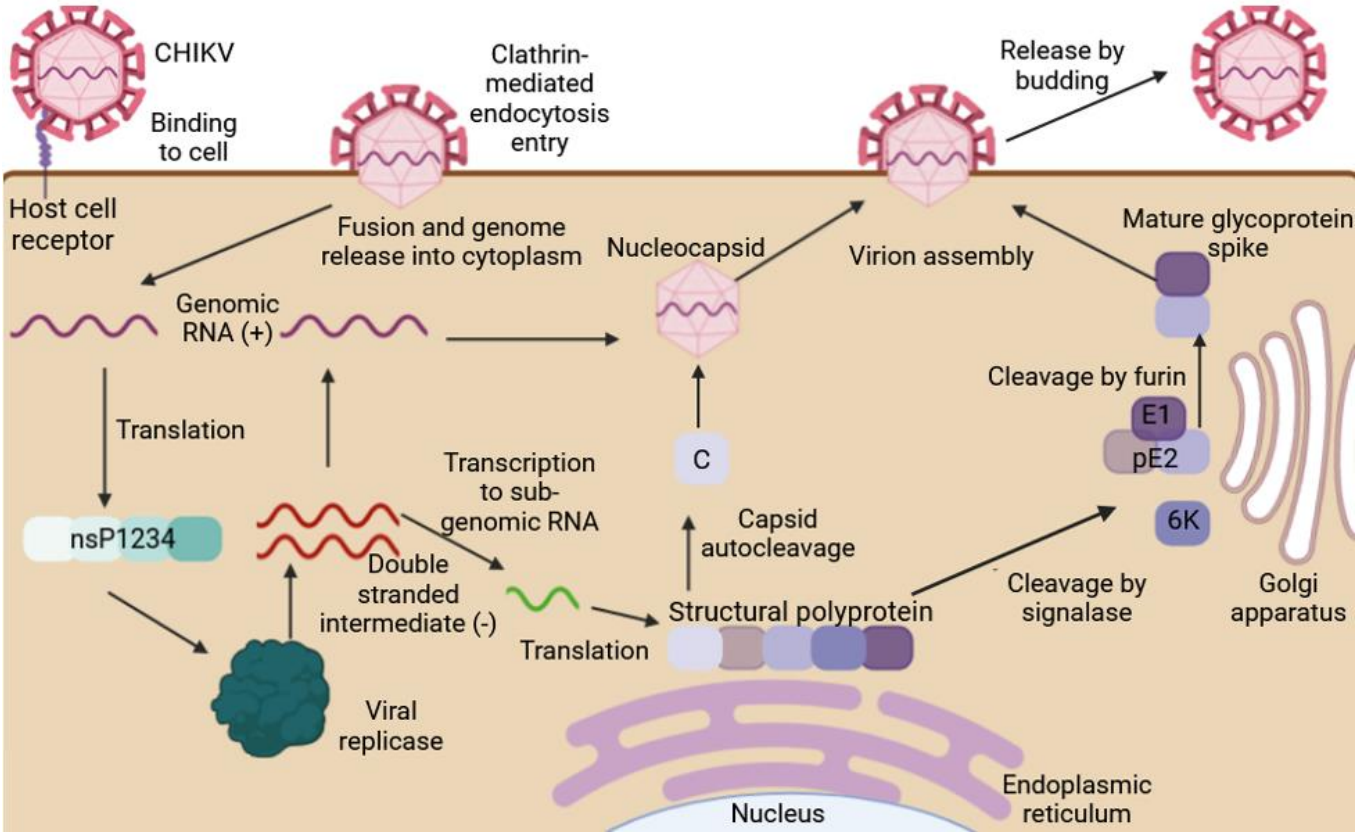
1.4.1 Entry

CHIKV shows a broad range of tropism for both mammalian and mosquito cell lines **(Basavappa et al., 2022).** Entry to the cell is dictated by attachment of the virus via binding of the protein E2 to host cell receptors. One example of a cell surface protein that CHIKV uses as an entry receptor is the matrix remodelling associated surface protein Mxra8, which has been shown to be a CHIKV receptor in mammalian cells **(Tan et al., 2022b).** Once attached, the virus initiates clathrin-mediated entry into the cell **(Echavarria-Consuegra et al., 2023)**. CHIKV has also been shown to enter muscle and Vero cells by macropinocytosis, showing that the virus is capable of entering cells by alternative methods. However, while inhibition of clathrin-mediated entry greatly decreases viral titre, inhibition of macropinocytosis has no significant effect on titre, indicating that endocytosis is the major form of entry of CHIKV **(Reis et al., 2022)**. Clathrin-mediated endocytosis involves the enveloping of the attached virion into a vesicle, which is then transported into the cytoplasm. Once inside the endosome, changes in pH result in conformational changes of the E1-E2 heterodimer, leading to the previously buried E1 fusion loop to become exposed **(Carey et al., 2019)**. When exposed, the fusion loop binds to the endosomal membrane, resulting in the virus being released into the cytoplasm **(Echavarria-Consuegra et al., 2023)**. C protein then binds to 60S ribosomal subunit in the cytoplasm, resulting in the disassembly of the nucleocapsid and release of the genomic RNA.

1.4.2 Early replication events

As viruses do not possess their own translation machinery, they must use host machinery for viral protein translation. CHIKV is a positive-sense RNA virus and so once the genome has been released into the cell cytoplasm, ORF-1 can be directly translated by the host **(Figure 6A)**. The type-0 cap and poly(A) tail of the viral genomic RNA mimic host mRNA, allowing translation to occur **(Hyde et al., 2015)**. The presence of the opal stop codon upstream of nsP4 in some CHIKV isolates results in the majority of translation products being the polyprotein nsP123 in these isolates **(Barton et al., 1991)**. However, as a result of readthrough, between 10-20% of translation events will result in the translation of polyprotein nsP1234 **(Rupp et al., 2011).** nsP4 is rapidly released from the polyprotein via protease action of nsP2, producing the active RdRp. nsP4 can be cleaved from the polyprotein in both *cis* and *trans* but is preferentially cleaved in *cis* **(Prosser et al., 2023).** nsP4 interacts with nsP123, forming the initial RC **(Figure 7A).** This initial RC is responsible for the synthesis of full-length negative strand RNA from the viral genome template **(Figure 6A)**. UTRs on viral RNA act as promoters, binding to active nsP4 to initiate genome replication of both positive and negative RNA **(Hyde et al., 2015).**

A)


CHIKV
Binding to cell
Host cell receptor
Clathrin-mediated endocytosis entry
Fusion and genome release into cytoplasm
Genomic RNA (+)
Translation
nsP1234
Double stranded intermediate (-)
Transcription to sub-genomic RNA
Translation
Viral replicase
Nucleocapsid
C
Capsid autocleavage
Structural polyprotein
Virion assembly
Release by budding
Mature glycoprotein spike
Cleavage by furin
E1
pE2
6K
Cleavage by signalase
Golgi apparatus
Endoplasmic reticulum
Nucleus

B)

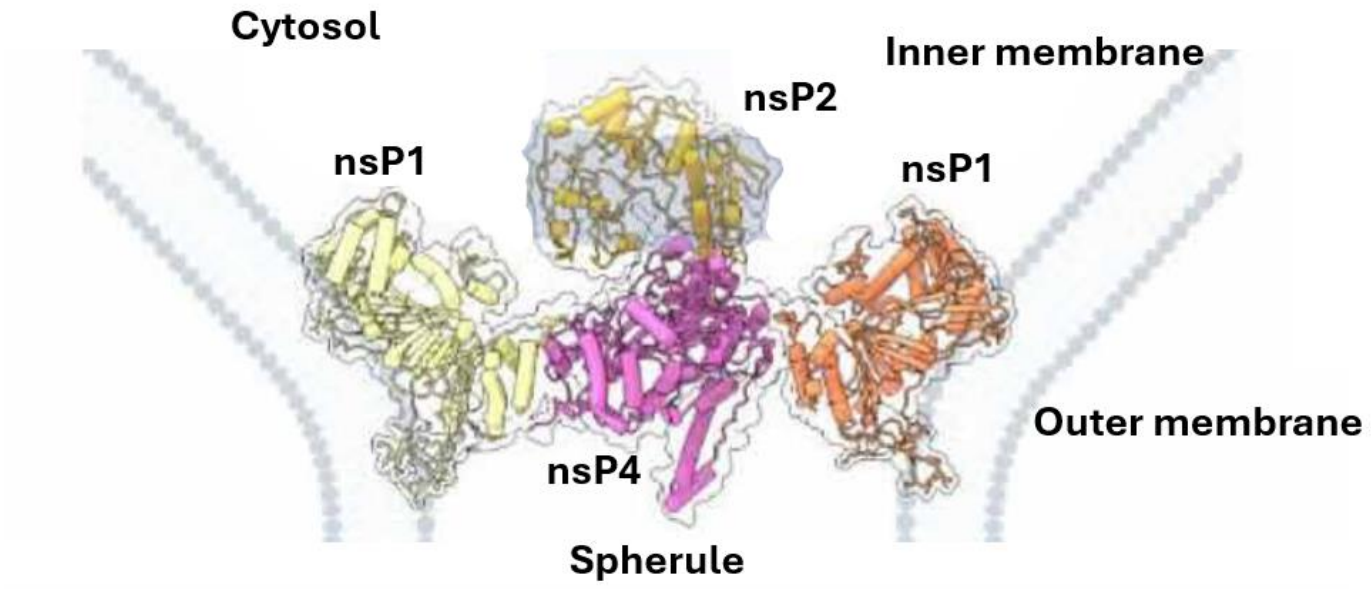

Cytosol
Inner membrane
nsP2
nsP1
nsP1
Outer membrane
nsP4
Spherule

Figure 6: Schematic representation of CHIKV replication cycle. A) *Representation of full CHIKV lifecycle. Virion attaches to host cell receptor on plasma membrane, triggering entry by clathrin-mediated endocytosis. Fusion of viral protein E1 with endosomal membrane results in release of nucleocapsid into the cytosol, which disassembles and releases the viral genome. Genomic RNA is translated directly into the nsP polyprotein, which undergoes processing to form the viral replicase. Double stranded RNA intermediates produced from positive RNA genome can synthesise new full length genomic RNA or produce sub-genomic RNA. Sub-genomic RNA is translated into the structural polyprotein, from which capsid (C) protein is released by autoproteolysis. E1 and 6K are released via cleavage by host signalases. E1 binds to the precursor protein pE2, which both undergo post translational modification in the Golgi apparatus. E3 is cleaved by host protein furin, leaving the E1-E2 glycoprotein spike. Genomic RNA is encapsidated by capsid proteins, forming the nucleocapsid. E1-E2 spikes recruit the nucleocapsid to the surface, and the new mature virion buds from the cell. Budding results in the CHIKV virion being enveloped* ***Adapted from (Freppel et al., 2024, Peinado et al., 2022). B)*** *Location of CHIKV RC within cell. CHIKV RC is situated at the host cell membrane, in membrane invaginations known as spherules. nsP2 and nsP4 sit within nsP1 ring, with nsP1 anchoring RC to the cell membrane at the spherule neck. A cytoplasmic ring, believed to be constructed of nsP3 and host factors is additionally associated with the nsP1 ring (not pictured).* ***Adapted from (Tan et al., 2022a).***

CHIKV genome replication takes place in membrane derived structures known as spherules **(Figure 6B) (Echavarria-Consuegra et al., 2023)**. These structures are around 50 nm in width and connect to the plasma membrane by a “neck-like” structure, where the RC is located **(Spuul et al., 2011).** Spherules are invaginations of the membrane and enhance viral replication by helping the virus to evade innate immune detection, as spherules prevent dsRNA intermediates being recognised by pattern recognition receptors in the cytoplasm **(Tan et al., 2022a)**. Spherules form in infected cells as early as two hours post infection (hpi), the same time that replication has also been reported to begin in CHIKV **(Reis et al., 2022)**. Spherules will

eventually become internalised, though RNA synthesis will continue after internalisation. CHIKV spherules have been reported to remain at the plasma membrane for longer than other alphaviruses **(Tan et al., 2022a)**, with internalised CHIKV spherules shown to contain C protein. This contrasts other alphaviruses infections in which spherules are internalised at an earlier timepoint in their lifecycle, as these spherules contain no structural proteins **(Reis et al., 2022)**.

nsP1 can additionally be released from the nsP123 polyprotein in *trans* **(Singh et al., 2018)** to form a short-lived nsP1-nsP23-nsP4 which preferentially transcribes negative strand RNA but is additionally capable of synthesising positive strand genomic RNA **(Figure 7B) (Shin et al., 2012)**. nsP23 is rapidly cleaved into its respective proteins in *trans,* in order to form the mature RC **(Lello et al., 2021)***.* Negative strand intermediate CHIKV genomic RNA can then be used as a template for amplification of positive strand genomic RNA and the 26S sub-genomic RNA, with the latter acting as a template for translation of the structural proteins **(Cristea et al., 2006).**

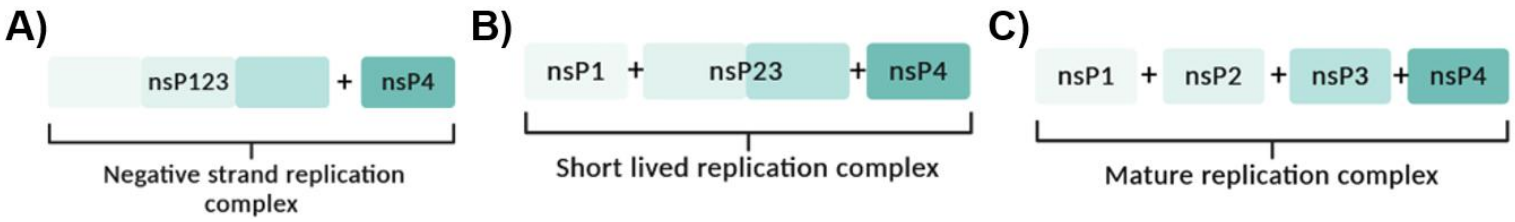


Figure 7: Schematic representation of the compositions of the three different viral replication complexes. A) *Initial negative strand replication complex consisting of nsP123 and free nsP4* ***B)*** *Short lived replication complex that preferentially replicates the negative strand, consisting of nsP23 and free nsP1 and nsP4* ***C)*** *Mature, long lived replication complex composed of four fully cleaved and matured nsPs.*

1.4.3 Switch to positive RNA strand replication

Once the concentration of nsP123 reaches a high level, typically around 4 hpi, replication will switch from negative strand to positive strand synthesis. nsP123 is cleaved into its' respective proteins by nsP2, and the four nsPs interact to form the mature RC **(Figure 7C) (Matkovic et al., 2019)**. The mature RC will then replicate full-length genomic RNA for encapsidation into new virions, in addition to transcribing sub-genomic mRNA via the 26S promoter at the ORF-1/ORF-2 junction region. Late-stage RCs are more stable than the initial RCs, with a longer half-life **(Spuul et al., 2011)**, resulting in positive strand RNA being synthesised to higher levels than negative strand intermediates as positive strand synthesis continues throughout the CHIKV lifecycle while negative strand synthesis ceases when the mature RC forms **(Lulla et al., 2012)**. Excess nsPs will form non-replicative complexes at the plasma membrane as a means of regulating viral replication **(Tan et al., 2022a).**

1.4.4 Assembly

C protein is rapidly cleaved from the structural polyprotein by autoproteolysis **(Echavarria-Consuegra et al., 2023)**. Free C protein interacts with genomic RNA by arginine, lysine, and proline residues in its N-terminus domain, which catalyses the oligomerisation of C proteins to form the nucleocapsid and encapsidate a singular viral genome **(Figure 6A) (Carey et al., 2019)**. C protein recognises signals in the genome region encoding nsP1 and in the 5' end of the genome so that only full-length genomic RNA is packaged into virions **(Leung et al., 2011, Weiss et al., 1989).** The nucleocapsid is made up of 240 individual C proteins and around 40nm wide **(Reis et al., 2022).**

The majority of translation events produce the structural polyprotein E3-E2-6K-E1 after C protein cleavage, but a minor product E3-E2-TF can also be produced as a result of ribosomal frameshifting (around 10-18% of translation events, **(Singh et al., 2018))**. The E3 N-terminus will traffic the remaining polyprotein to the ER **(Echavarria-Consuegra et al., 2023)**. At the ER, signalases will cleave the polyprotein resulting in the production of pE2 and either 6K and E1 or TF **(Figure 6A) (Carey et al., 2019)**. E1 and pE2 bind together to form a heterodimer that undergoes post-translational modifications and conformational changes inside the ER lumen **(Figure 6A)**. These heterodimers then trimerise, after which the host protease furin cleaves pE2 in the *trans*-Golgi, resulting in the formation of the mature E1-E2 heterodimers **(Carey et al., 2019)**. Mature spikes are membrane associated and recruit the nucleocapsid to the cell surface **(Echavarria-Consuegra et al., 2023),** with E2 association with the C protein hydrophobic domain resulting in the virion budding from the cell **(Kril et al., 2021).** As the virus buds, it acquires a lipid envelope from the host cell plasma membrane, which is studded with the

glycoprotein spikes **(Carey et al., 2019)**. In addition to budding from the flat cell membrane surface, **Reis et al. (2022)** demonstrated in a study that CHIKV can be released from the cell in a process similar to budding, by inducing the cell membrane to form protrusions that release a singular virion. Additionally, it was reported in the same study that virions may be able to acquire an envelope internally, without release from the cell. Around 80% of the cells observed released virus particles by either budding or protrusions. The mature released virion is an icosahedral shape with *T=4* symmetry **(Singh et al., 2018)** and around 70 nm wide **(Lee et al., 2013).**

1.5 nsP4

1.5.1 Structure of CHIKV nsP4

The polymerase activity of CHIKV nsP4 is contained in the C-terminus region of the protein, while the N-terminus region is disordered to interact with other proteins including the other viral nsPs **(Pareek et al., 2022)**. Alphavirus nsP4 has been reported as being monomeric by size exclusion chromatography and gel filtration data **(Chen et al., 2017, Tan et al., 2022b, Tomar et al., 2006).**

Mutations to the C-terminus of SINV nsP4 have been shown to affect polymerase activity. For example, a single nucleotide change has been shown to the prevent the RC from switching from negative strand replication to positive strand replication **(Sawicki et al., 1990).** Not a lot is known about the function of the N-terminus of alphavirus RdRps, although it is known that it is not directly involved in RNA synthesis. N-terminus truncated versions of nsP4 still possess polymerase activity – albeit limited. However, *de novo* RNA synthesis does not occur without the presence of the N-terminus due to the lack of interactions with other nsPs, and while TATase activity of nsP4 is functional without the N-terminus, it is more efficient when the N-

terminus is present, despite being less stable than truncated versions where the N-terminus has been removed **(Tan et al., 2022b).** The N-terminal domain is essential for nsP4 to be functional within the RC, as a result of the other nsPs interacting with nsP4 via this domain **(Tomar et al., 2006).**

Viral RdRps typically possess a right-hand fold structure possessing seven motifs, A-G, that form the three main domains of the protein: the palm, fingers and thumb **(Pareek et al., 2022).** The palm domain, formed from motifs A-E, is the centre of the polymerase catalytic site and possesses the conserved GDD motif **(Pareek et al., 2022)**. The catalytic site is formed from two aspartate residues from motif C and third aspartate from motif A **(Reyes-Gastellou et al., 2021)**. Together, they process incoming nucleotides during RNA synthesis and control magnesium coordination **(Reyes-Gastellou et al., 2021)**. Motifs F and G, on the other hand, make up the finger domains which are involved in conformational changes of the RdRp, and so are highly flexible **(Martin et al., 2024)**.

Predicted structural computational models of CHIKV nsP4 have been reported several times, with the typical palm, finger, and thumb domains of viral RdRps **(Freire et al., 2022, Ghildiyal et al., 2019, Pareek et al., 2022, Reyes-Gastellou et al., 2021, Martins et al., 2023).** However, attempts to experimentally resolve the 3D structure of CHIKV nsP4 have been unsuccessful thus far. In fact, it was only in 2022 that the first 3D structures of any alphavirus nsP4 was solved. **Tan et al. (2022b)** first published crystal structures for the catalytic domains of SINV and RRV nsP4 **(Figure 8)**. The N- terminus was poorly resolved in RRV, with several disordered regions that were unable to be resolved, in particular the RNA binding tunnel. Further mass spectrometry analysis allowed for the RRV N-terminus structure to be categorised when recombinantly expressed alone. The difficulty in analysing this structure was in

part due to the N-terminus' "intrinsic flexibility" **(Tan et al., 2022b).** RRV is more closely related to CHIKV, as they are both classified in the SFV complex and have a higher similarity on an amino acid level than SINV and CHIKV **(79% and 73% respectively (Chen et al., 2017)),** so is more likely to possess a similar structure to CHIKV than SINV. A second study attempting to solve the structure of the CHIKV RC was published later that year, however despite successful expression and purification of nsP1-3, they were unable to purify nsP4 **(Tan et al., 2022a)**. The study instead utilised ONNV nsP4, which has a 91% sequence similarity to CHIKV RdRps on an amino acid level **(Martin et al., 2024)**. Though chimeric RCs combining nsP123 and nsP4 from two different alphaviruses are not always functional, ONNV nsP4 has previously been shown to produce a functional RC with CHIKV nsP123 with replication and transcription being detected to high levels **(Lello et al., 2021).** ONNV RdRp was therefore used as a substitute for CHIKV nsP4, due to CHIKV nsP4's poor solubility. This study revealed that nsP4 was positioned in the centre of the CHIKV RC **(Figure 6B)**. **Tan et al. (2022a)** were additionally able to resolve the N-terminus domain of ONNV nsP4 in the context of the RC, despite having difficulty resolving this region in RRV and SINV nsP4s expressed without the other members of the RC. Together these two studies indicated that free alphavirus nsP4 has a weakly functioning, partially unfolded conformation that possesses high plasticity, and then becomes fully activated when incorporated into the RC **(Tan et al., 2022a)**.

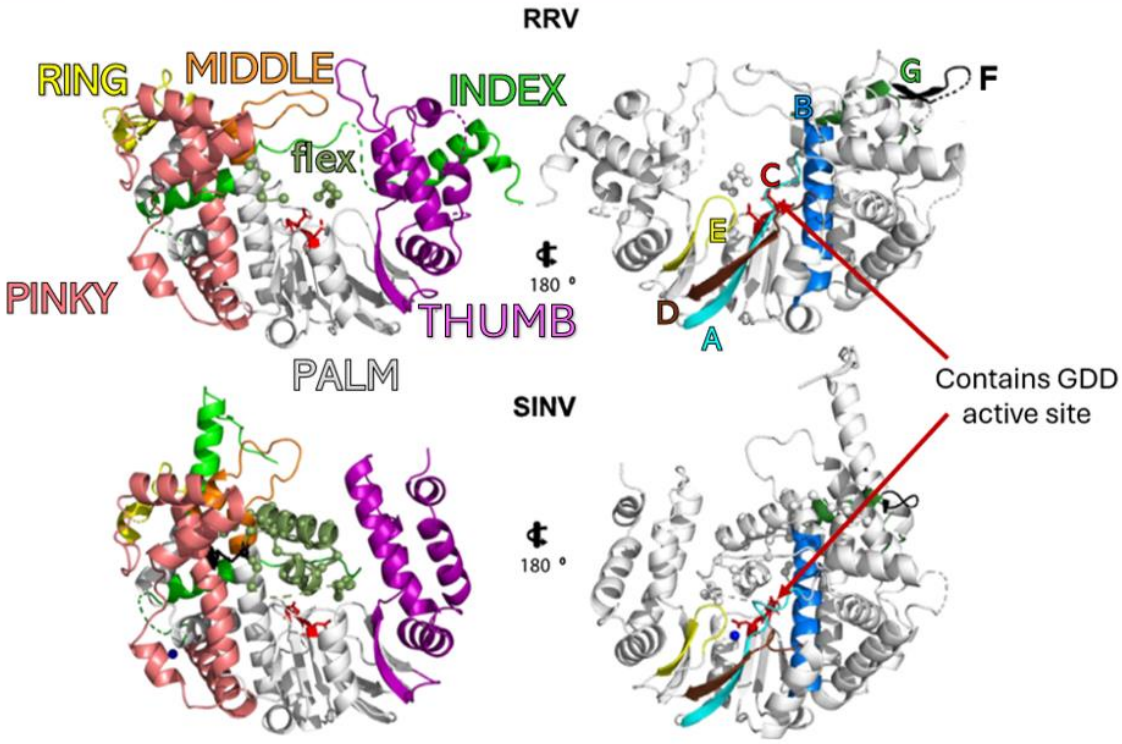


Figure 8: Schematic representation of crystal structures of RRV (top) and SINV (bottom) catalytic cores. *On left: RdRp domain differentiated and labelled by colour: index in green, middle in orange, ring in yellow, pinky in coral pink, palm in grey, thumb in magenta. On right: Right hand fold motifs are differentiated and labelled by colour: A is teal, B is dark blue, C is red, D is brown, E is yellow, F is black, and G is green. Motifs are shown in front right hand (left) and rear view (right). Red arrow points to Motif C, which possesses the catalytic site. Flex region (olive) represents missing residues (83 residues in RRV, 39 residues in SINV). RRV nsP4 has a resolution of 2.8 Å, SINV nsP4 has a resolution of 1.9 Å.* ***Adapted from (Tan et al., 2022b).***

Due to the limited information available on alphavirus RdRps, especially before the structures of the first alphavirus polymerases were reported, CHIKV nsP4 was previously been compared to distantly related arboviruses such as dengue virus (DENV) and Zika virus (ZIKV) in order to provide insights into its structure and compare its similarity to other arboviruses **(Reyes-Gastellou et al., 2021).** As the GDD catalytic site is highly conserved and mutations to this site are lethal to

replication, structures of motifs A-C where the GDD site is found were highly conserved across CHIKV, DENV and ZIKV, with the GDD site being perfectly aligned between viruses. Motifs D and E in CHIKV were also predicted to be structurally similar to DENV and ZIKV RdRps **(Reyes-Gastellou et al., 2021)**. However, this was disputed by **Tan et al. (2022b),** as the first resolved alphavirus nsP4 structures indicated that Motif D in RRV and SINV formed an elongated β-strand in contrast to the β-strand loop usually found in viral RdRps. Motifs F and G were not aligned between any of the arboviruses discussed in the study, possessing different structures in each virus. Additionally, unlike the other two arboviruses, the predicted CHIKV RdRp structure did not possess a priming loop, indicating structural differences between CHIKV nsP4 and RdRps of distantly related viruses **(Reyes-Gastellou et al., 2021)**.

1.5.2 Complications in studying nsP4.

When attempting to study alphavirus nsP4 protein, several complications have hindered the ability for experimental work to be carried out. First is the abundance of nsP4 during experimental work. As only ~10% of ORF-1 translation events produce the full length nsP1234 polyprotein, there is a lower concentration of nsP4 in comparison with the other viral nsPs. Increasing the ratio of nsP123 to nsP4 *in vitro* increases the rate of negative strand replication, indicating that the opal stop codon regulates CHIKV replication by preventing the expression of the full nsP polyprotein **(Lello et al., 2021).** Surplus nsP4 in the cell has been reported as detrimental to late-stage CHIKV replication, leading to a decrease in the levels of structural proteins and mature virions produced despite an increase in RNA synthesis **(Kaur et al., 2020)**. Though CHIKV isolates without the opal stop codon have been identified, the lack of the opal stop codon does not have a detrimental effect on the overall virus

fitness **(Ahola and Merits, 2016).** Isolates that do not possess an opal stop codon do not produce increased levels of mature nsP4 because of additional regulatory methods of controlling nsP4 levels, such as the N-end rule pathway **(Lello et al., 2021)**.

Protein degradation by the N-end rule pathway is initiated by a signal in the N-terminus that destabilises the protein, marking it for conjugation with ubiquitin which results in degradation by a host protease **(de Groot et al., 1991).** nsP4 possesses a completely conserved tyrosine residue in its N-terminus **(Chen et al., 2017, Shirako and Strauss, 1998),** which acts as a destabilising residue and marks the protein for degradation by the N-end rule pathway **(Pareek et al., 2022)**. This results in surplus nsP4 being rapidly degraded by the host cell. As a result of degradation, SINV nsP4 has been reported as having a short half-life both during infection *in vitro* and in cell lysates **(de Groot et al., 1991)**. This destabilising residue is essential for the proper formation of the viral RC, with its removal being deleterious for the virus **(Chen et al., 2017, Rubach et al., 2009).** However, nsP4 appears to become stabilised once complexed within the RC, preventing degradation and increasing its half-life **(de Groot et al., 1991)**. The concentration of overexpressed CHIKV nsP4 in human embryonic kidney cells has been shown to be regulated by the host protein RACK1 which, which modulates degradation of nsP4 by the N-end rule pathway **(Yan et al., 2024)**. Between the opal stop codon and the degradation of alphavirus RdRps by host cell pathways, the tight regulation of nsP4 concentrations in infected cells can make studying the protein challenging.

Another issue when studying nsP4 are difficulties in expressing and purifying the recombinant nsP4. Solubility of recombinant nsP4 has been reported as an issue in several alphavirus species, with challenges in separating nsP4 from insoluble cell

debris when lysing and releasing the protein from the cell membrane **(Rubach et al., 2009)**. Additionally, nsP4 has been reported to aggregate during attempts to express and isolate the recombinant protein **(Tan et al., 2022b).**

Recombinant SINV nsP4 was the first alphavirus RdRp to be successfully expressed and purified **(Tomar et al., 2006).** Initially, only a truncated version of SINV nsP4, missing 97 amino acid residues in the N-terminus, was able to be solubilised. This protein was the only one out of 10 truncated version of nsP4 that was successfully isolated **(Tomar et al., 2006). Tomar et al. (2006)** also attempted to isolate the full length RdRp but were unsuccessful. They hypothesised that this failure was the result of nsP4 degradation by the N-end rule pathway. While truncated versions of nsP4 can be useful in studying the core RdRp, by analysis of its polymerase inhibitors through high throughput screenings and analysis of TATase activity, the missing N-terminus region prevents protein-protein interactions forming and so have limited use **(Ahola and Merits, 2016).** Functional RCs, for example, can only be formed as a result of the correct series of interactions between the nsPs, the viral RNA and host factors **(Lello et al., 2021),** which are unable to occur without the N-terminus domain**.** The first successful isolation of functional, full-length alphavirus nsP4 was reported by **Rubach et al. (2009),** with the addition of SUMO tag to stabilise SINV RdRp during isolation which was then cleaved following purification**.**

Limited reports of CHIKV nsP4 expression and purification using a bacterial expression system in *Escherichia coli* have been reported **(Martins et al., 2023)**, but isolating functional recombinant nsP4 is often still a limiting factor in the study of the protein. Additionally, though an active, truncated form of CHIKV nsP4 had previously been reported by **Chen et al. (2017),** a later study by **Lello et al. (2021)** described their inability in express and purify both full length and the truncated form previously

reported. This emphasises the difficulties faced when attempting to purify CHIKV nsP4. Furthermore, the instability of nsP4 in eukaryotic cells due to ubiquitin degradation creates challenges in isolating the protein from these cells, as expressed protein can have a very short half-life prior to purification **(Ahola and Merits, 2016).**

Despite ONNV and CHIKV having a 91% similarity in amino acid sequence, they exhibit wide biological differences including the infection of different mosquito hosts. While CHIKV is transmitted by *Aedes spp.* mosquitoes, ONNV is transmitted by the species *Anopheles* **(Khan et al., 2002).** This difference in vector can result in disparities in the way the viruses interact with their hosts, as they will have adapted different mechanisms for evading their host's immune systems. The N-terminus domain specifically has been proposed to have adapted to interact with host factors due to its proximity to the protein surface and its lack of conserved structure between virus families **(Jia and Gong, 2019, Tan et al., 2022b).** In addition to alphavirus nsP4 having previously being reported as possessing lower polymerase activity than other species of distantly related arboviruses, polymerase activity varies across the alphavirus genus. For example, the rate of polymerase activity of recombinantly expressed SINV nsP4 is significantly higher than the rate of CHIKV nsP4. **(Tan et al., 2022b).** These differences occur across alphavirus polymerases despite being the most highly conserved protein, therefore there is a need to study each alphavirus individually, to account for these differences.

The difficulty in expressing and isolating functional recombinant CHIKV nsP4 has also meant that antibodies against nsP4 have been difficult to raise. Commercial antibodies for CHIKV nsP4 have only recently become available, and due to the poor

solubility, low abundance and potential degradation of the protein in cell lysate, the available antibodies are not very efficient.

1.5.3 nsP4 interactions

nsP4 is known to interact with the other CHIKV nsPs and cellular proteins during replication. For example, nsP4 has been shown to interact with the nsP2 N-terminus, which possesses helicase properties, for efficient RNA unwinding **(Rana et al., 2014)**. Conversely, free nsP4 has low polymerase activity, which increases when it is stabilised by and incorporated into the RC **(Tan et al., 2022a).**

In addition to interactions with other viral proteins, nsP4 interacts with several host factors during infection. Alphavirus infections in mammalian cells have a lytic effect, and so the virus must evade or exploit the host response to efficiently complete its lifecycle **(Abraham et al., 2015)**. Many host proteins have been identified as interactors of CHIKV nsPs, therefore it is logical that nsP4 would also interact with host cell proteins **(Freire et al., 2022)**. However, due to the various difficulties working with nsP4 outlined above, studies investigating the CHIKV nsP4 interactome remain limited. Host cell proteins that have previously been identified as interactors of CHIKV nsP4 include heat shock protein 90 (HSP90), TMEM45B and eukaryotic initiation factor α (eIF2α).

The gene encoding TMEM45B is stimulated by interferons and is therefore upregulated by the innate host response to CHIKV infection **(Brehin et al., 2009)**. TMEM45B is a transmembrane protein that has been shown to play a role in mammalian cell proliferation **(Yan et al., 2022).** While TMEM45A and TMEM45C appear to have no effect on alphavirus replication, TMEM45B inhibits replication of SINV and CHIKV, with downregulation of TMEM45B resulting in recovery of viral

RNA replication **(Yan et al., 2022).** TMEM45B was consistently immunoprecipitated with nsP1 and nsP4, and colocalised with both nsPs in both SINV and CHIKV. When expressed alone, recombinant nsP4 was dispersed throughout the cell, but strongly colocalised with TMEM45B when they were co-expressed. **Yan et al. (2022)** hypothesised that TMEM45B inhibited replication by preventing interactions between nsP1 and nsP4, thereby disrupting formation of the RC.

TMEM45B demonstrates how model alphaviruses can give insight into CHIKV interactions, as both SINV and CHIKV nsP4 interact with and are inhibited by TMEM45B in mammalian cells. However, as biological mechanisms can vary widely between alphavirus species, CHIKV still must be studied individually as not all interactions found in the model species will be applicable to CHIKV.

A common host pathway targeted by viruses during infection is the unfolded protein response (UPR) **(Abraham et al., 2015)**. The UPR is induced in cells when under stress. One pathway of the UPR involves PKR-like ER kinases (PERK) and is therefore known as the PERK pathway **(Scholte et al., 2015)**. PERK is activated by dimerising and becoming phosphorylated **(Rathore et al., 2013)**. Active PERK then goes on to phosphorylate eIF2α by phosphorylation, leading to a shutdown of protein synthesis **(Scholte et al., 2015)**. Additionally, PERK activates C/EBP homologous protein (CHOP) which can result in cell death **(Fros et al., 2015b).**

When the UPR is activated as a result of CHIKV infection, CHIKV nsP4 binds to eIF2α blocking phosphorylation, despite active PERK being detected. **(Rathore et al., 2013)**. Phosphorylation of eIF2α is not detected in CHIKV infected cells until 48 hpi, with recovery of the PERK pathway hypothesised to be due to the degradation of nsP4. Additionally, CHOP was only detected at very low levels, and did not

increase during infection **(Rathore et al., 2013)**. This interaction between nsP4 and eIF2α is therefore beneficial to CHIKV as it allows host protein synthesis machinery to continue, which is necessary for the CHIKV lifecycle as the virus is not capable of translating its own proteins.

The relationship of nsP4 and eIF2α in CHIKV differs from other alphaviruses, exemplifying why protein interactions of CHIKV nsP4 must be studied individually. For example, in the model alphavirus SINV, nsP4 has no effect on eIF2α phosphorylation, and so the PERK pathway proceeds uninterrupted during SINV infection. This demonstrates that although model viruses can be useful for comparisons of host responses in CHIKV, like in the case of TMEM45B above, protein-protein interactions during alphavirus infection are not always universal.

While studying the effects of CHIKV on different pathways of the UPR, **Rathore et al. (2013)** reported that the levels of UPR chaperones were being upregulated during infection. HSPs regulate cellular homeostasis, acting as chaperones that regulate and recycle misfolded or unfolded proteins. HSP90 is one of the more highly expressed families of HSPs and are highly conserved across mammalian cells **(Rathore et al., 2014).** Two isoforms of HSP90 are found in the cytoplasm; HSP90α which is upregulated when the cell experiences stress and HSP90β which is continuously expressed by the cell **(Rathore et al., 2014).** HSP90 has been reported as having a role during the infection of several viruses, such as the stabilisation of the RC in influenza viruses **(Rathore et al., 2013).** Other HSPs such as HSP70 have additionally been shown to promote CHIKV replication during infection **(Paingankar and Arankalle, 2014).**

HSP90 inhibitors were shown to reduce CHIKV virus titre in a dose dependent manner **(Das et al., 2014, Rathore et al., 2014).** However, this inhibitory effect was reduced when inhibitory molecules were introduced at later timepoints in infection, with the introduction at 0 hpi reducing viral titre significantly more than at 4,6 or 8 hpi. This indicates that HSP90 plays a role in early CHIKV infection **(Das et al., 2014).** HSP90 inhibitors have additionally been shown to reduce viraemia and concentration of cytokines in the bloodstream in CHIKV infected mice, demonstrating the proviral effect of HSP90 during CHIKV infection **(Rathore et al., 2014).**

Three of the four nsPs encoded by CHIKV have been reported to interact with HSP90: nsP2, nsP3 and nsP4 **(Das et al., 2014, Rathore et al., 2014).** The interaction of nsP4 with HSP90 was demonstrated by **Rathore et al. (2014)** through mass spectrometry analysis, which indicated nsP4 interacted with both HSP90 isoforms while nsP3 only interacts with HSP90β. Moreover, nsP4 was co-precipitated HSP90α and vice versa, again indicating an interaction between the two proteins.

When HSP90α levels were inhibited by siRNA during CHIKV infection, the levels of viral RNA produced were significantly decreased, suggesting that the interaction between HSP90α and nsP4 promotes viral replication. **Rathore et al. (2014)** suggested that as HSP90α was mainly found in the cytoplasm, it's interaction with nsP4 was pre-RC formation, helping to stabilise nsP4 and produce the early RC.

1.5.4 Importance of studying nsP4

CHIKV is capable of causing widespread epidemics with infection associated with very high morbidity, with both acute and chronic CHIKV patients reporting high levels of pain and low mobility **(Tiozzo et al., 2025)**. With the threat of CHIKV spreading further into temperate areas and naïve populations, due to increased globalisation

and higher temperatures as a result of global warming, and the associated social and economic burden of CHIKV outbreaks, discovery of therapeutic treatments is crucial **(Paingankar and Arankalle, 2014)**. Chronic CHIKV symptoms have been hypothesised to result from the severity of the initial infection, therefore treatments to prevent severe disease from occurring during the acute stages could prevent long lasting symptoms from developing**.**

Antivirals are often designed to target RdRps as they are completely essential to the viral lifecycle **(Pareek et al., 2022)**. However, antivirals can also be targeted towards key protein-protein interactions that are detrimental towards virus replication when inhibited. Furthermore, host proteins are often preferable as drug targets as antiviral resistance is less likely to develop **(Carey et al., 2019)**. Consequently, a greater understanding of CHIKV nsP4 could not only provide insights into the mechanisms of CHIKV replication and translation, but also help to identify novel targets for drug treatments **(Echavarria-Consuegra et al., 2023).** The limited knowledge of the nsP4 interactome during CHIKV infection is the result of both difficulties studying CHIKV nsP4 and a lack of interest in CHIKV until its re-emergence in the mid-2000s. Developing antivirals against CHIKV nsP4 host interactions may also provide broad spectrum treatment for additional alphaviruses, as they often target similar pathways **(Carey et al., 2019)**, although as previously mentioned the interactions of nsP4 can differ across the genus.

1.5.5 Solutions to difficulties studying nsP4

It is clear that a new system for studying nsP4 is needed in order to effectively analyse the protein and its interactions, particularly in conjunction with its role in genome replication during early infection events.

Attempts to ligate an affinity tag in the SINV nsP4 C-terminus have been reported as unsuccessful **(Rubach et al., 2009)** However, ligation of a SUMO-tag in the N-terminus stabilised SINV nsP4 so the full-length protein could be isolated via the bacterial expression system **(Rubach et al., 2009)**. The choice of lysis buffer employed during cell lysis has also been indicated as playing an important role in the solubility of nsP4 when attempting to isolate the protein **(Chen et al., 2017).**

Insertion of a Strep-FLAG tag into the N-terminus domain of SINV has previously been shown to have a relatively low decrease in RdRp function and has therefore been hypothesised as a good candidate for use in studying nsP4 **(Tan et al., 2022b)**. The ligation of a tag into the N-terminus not being deleterious to polymerase function, unlike tags inserted into the C-terminus, is likely due to the fact that the N-terminus is not directly involved in RNA synthesis. One suggested use of the Strep-FLAG tag inserted into SINV nsP4 was the investigation of nsP4 interactors via co-immunoprecipitation (Co-IP), by targeting the FLAG tag in place of using anti-nsP4 antibodies.

In order to study interactions of CHIKV nsP4 during early genome replication events, the majority of experimental work in this study was conducted using a CHIKV sub-genomic replicon (SGR). The CHIKV SGR expresses the four viral nsPs and encodes a firefly luciferase reporter gene in ORF-2, resulting in luciferase being expressed by translation of sub-genomic RNA. Structural proteins are not produced by the SGR, as the genes encoding these proteins have been replaced with the luciferase reporter gene. Use of the CHIKV SGR allows for the study of early replication events, in isolation to other CHIKV lifecycle stages. Interactions of CHIKV nsP4 during the early replication events of the virus lifecycle using the CHIKV SGR have not previously been reported. In addition to a version of the CHIKV SGR that

encoded the WT versions of the four nsPs **(see Chapter 3)**, a second CHIKV SGR which encoded recombinant nsP4 tagged with Strep II and three tandem FLAG tags was engineered in this study, to facilitate the future study of nsP4 and its host interactome, as described in **Chapter 4.**

1.6 Aims

This project aimed to enhance the ease of studying the nsP4 proteome interactome, during CHIKV translation and genome replication. To address this, three main aims were followed:

1: Optimisation of lysis buffer conditions for the release and detection of nsP4 from infected cell culture during CHIKV replication.

2: Development of a FLAG tag system to detect nsP4 during CHIKV sub-genomic replicon replication.

3: Validation of the new FLAG tag nsP4 sub-genomic replicon system by luciferase assays and western blotting.

Chapter 2: Materials and Methods

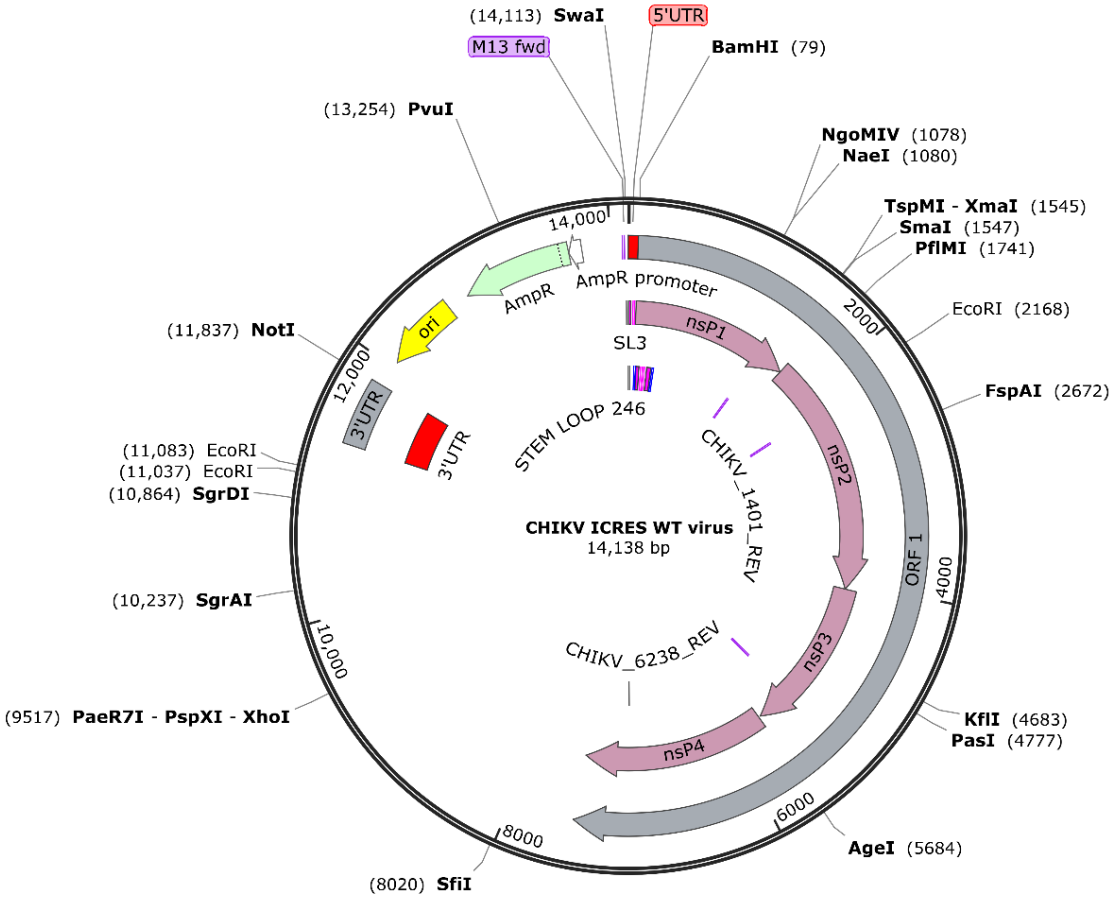
2.1 cDNA infectious clone and sub-genomic replicon plasmids

All plasmids utilised in this study are shown in **Figure 9.** Two infectious cDNA CHIKV-ICRES clones, derived from the CHIKV ECSA lineage **(Pohjala et al., 2011),** were utilised in this project. SP6-ICRES1 encoded WT versions of all CHIKV viral proteins (referred to as WT virus), while SP6-ICRES-nsP4Strep3F was a recombinant encoding nsP4 tagged at the N-terminus with Strep tag II and three tandem FLAG tags (referred to as nsP4-FLAG) **(Figure 9A and 9B respectively)**.

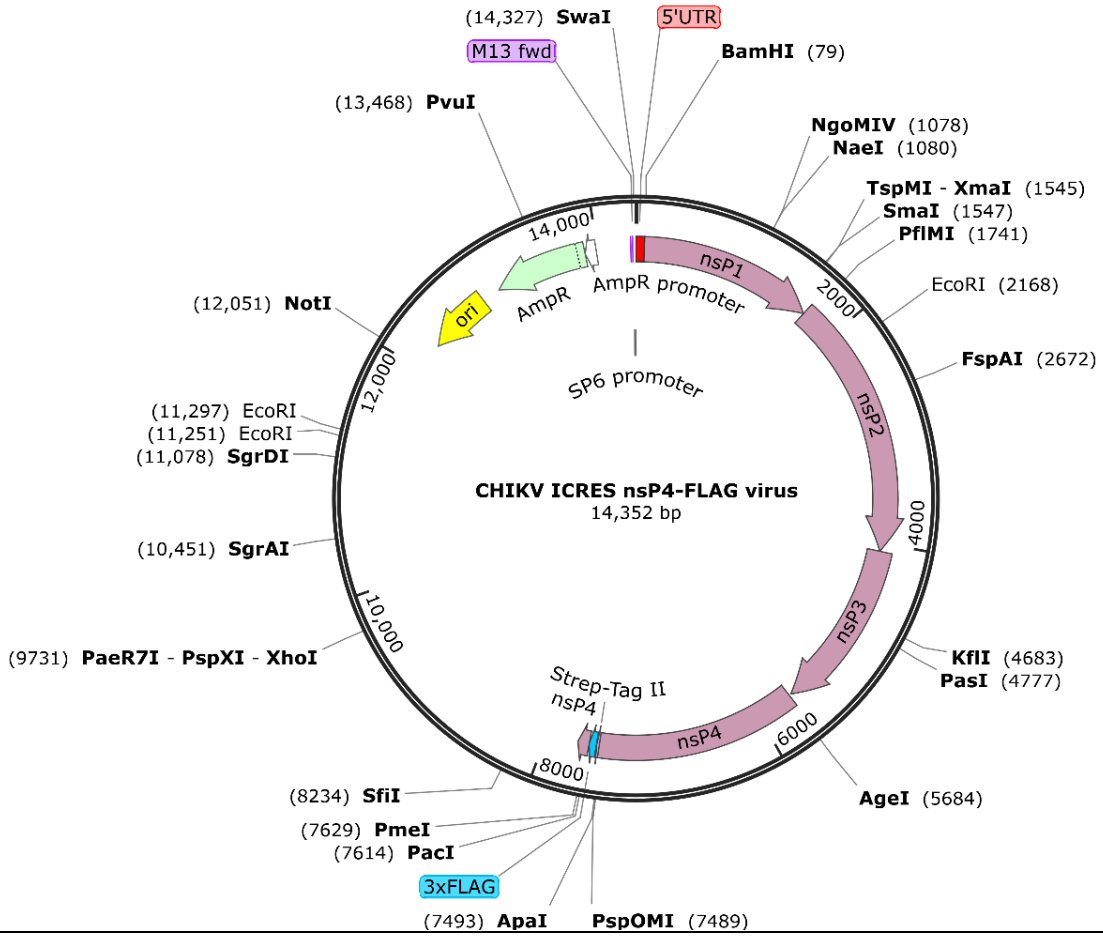
The CHIKV SGR, referred to as the WT SGR throughout **(Figure 9C)**, encodes a firefly luciferase marker in ORF-2 in place of structural protein encoding genes **(Pohjala et al., 2011).** A second SGR was produced as part of this project, which incorporated the nsP4 N-terminal Strep II and FLAG tags from nsP4-FLAG and is subsequently referred to as nsP4-3XF SGR **(Figure 9D)**. The cloning strategy for this is described in **sections 2.6-2.9.**

Both ICRES clones and the WT SGR were kind gifts from the Andres Merits lab.

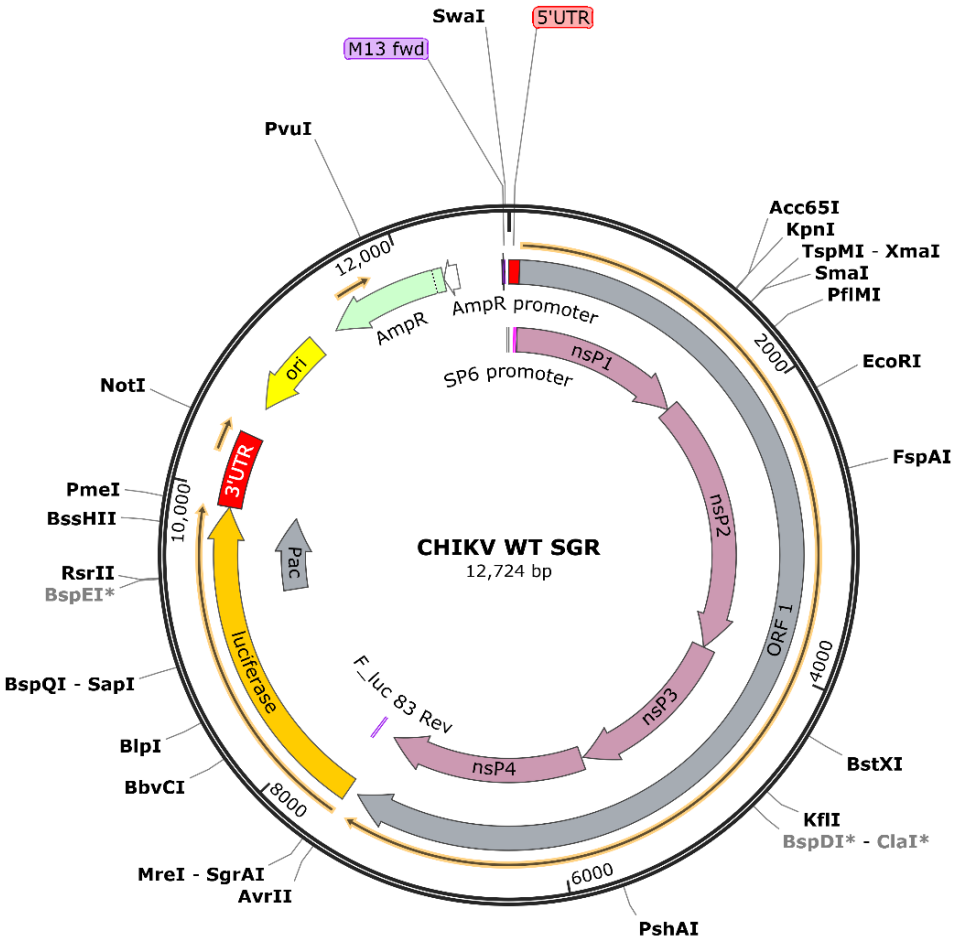
A)


(14,113) SwaI
5'UTR
M13 fwd
BamHI (79)
(13,254) PvuI
NgoMIV (1078)
NaeI (1080)
TspMI - XmaI (1545)
SmaI (1547)
PflMI (1741)
14,000
AmpR
AmpR promoter
nsP1
2000
EcoRI (2168)
(11,837) NotI
ori
12,000
SL3
3'UTR
3'UTR
STEM LOOP 246
FspAI (2672)
(11,083) EcoRI
(11,037) EcoRI
(10,864) SgrDI
CHIKV_1401_REV
nsP2
CHIKV ICRES WT virus
14,138 bp
ORF 1
4000
(10,237) SgrAI
10,000
CHIKV_6238_REV
nsP3
(9517) PaeR7I - PspXI - XhoI
KflI (4683)
PasI (4777)
nsP4
6000
8000
AgeI (5684)
(8020) SfiI

B)


(14,327) SwaI
5'UTR
M13 fwd
BamHI (79)
(13,468) PvuI
NgoMIV (1078)
NaeI (1080)
TspMI - XmaI (1545)
SmaI (1547)
PflMI (1741)
14,000
nsP1
AmpR
AmpR promoter
2000
EcoRI (2168)
(12,051) NotI
ori
12,000
SP6 promoter
FspAI (2672)
nsP2
(11,297) EcoRI
(11,251) EcoRI
(11,078) SgrDI
CHIKV ICRES nsP4-FLAG virus
14,352 bp
4000
(10,451) SgrAI
nsP3
10,000
(9731) PaeR7I - PspXI - XhoI
KflI (4683)
PasI (4777)
Strep-Tag II
nsP4
nsP4
6000
8000
AgeI (5684)
(8234) SfiI
(7629) PmeI
(7614) PacI
3xFLAG
(7493) ApaI
PspOMI (7489)

C)


SwaI
5'UTR
M13 fwd
PvuI
Acc65I
KpnI
TspMI - XmaI
SmaI
PflMI
EcoRI
FspAI
NotI
PmeI
BssHII
RsrII
BspEI*
BspQI - SapI
BlpI
BbvCI
MreI - SgrAI
AvrII
BstXI
KflI
BspDI* - ClaI*
PshAI
12,000
2000
4000
6000
8000
10,000
AmpR
AmpR promoter
ori
SP6 promoter
nsP1
nsP2
nsP3
nsP4
ORF 1
3'UTR
Pac
luciferase
F_luc 83 Rev
CHIKV WT SGR
12,724 bp

D)

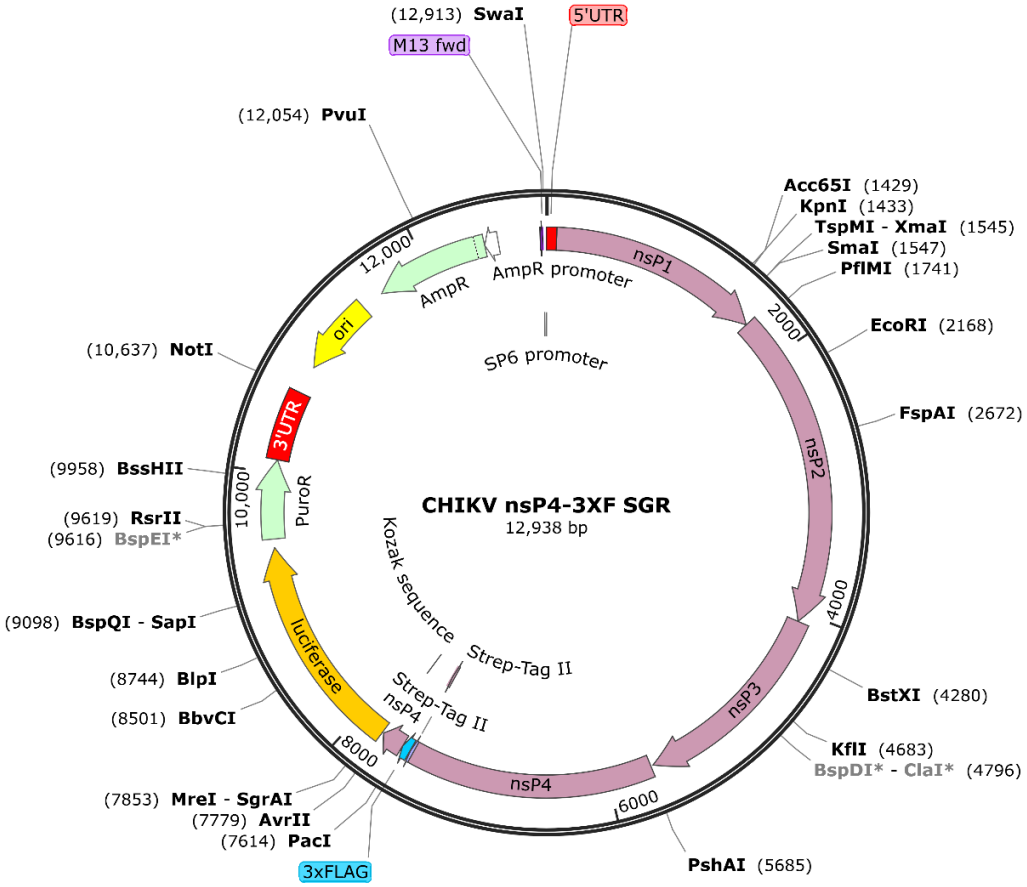

(12,913) SwaI
5'UTR
M13 fwd
(12,054) PvuI
Acc65I (1429)
KpnI (1433)
TspMI - XmaI (1545)
SmaI (1547)
PflMI (1741)
EcoRI (2168)
FspAI (2672)
(10,637) NotI
(9958) BssHII
(9619) RsrII
(9616) BspEI*
(9098) BspQI - SapI
(8744) BlpI
(8501) BbvCI
(7853) MreI - SgrAI
(7779) AvrII
(7614) PacI
3xFLAG
BstXI (4280)
KflI (4683)
BspDI* - ClaI* (4796)
PshAI (5685)
12,000
2000
4000
6000
8000
10,000
AmpR
AmpR promoter
ori
SP6 promoter
nsP1
nsP2
nsP3
nsP4
3'UTR
PuroR
luciferase
Kozak sequence
Strep-Tag II
Strep-Tag II
nsP4
CHIKV nsP4-3XF SGR
12,938 bp

Figure 9: Schematic representation of the CHIKV cDNA plasmid constructs used throughout this study. *Full length map of:* ***A)*** *CHIKV ICRES WT, the wild type virus* ***B)*** *CHIKV ICRES nsP4-FLAG, the virus producing functional nsP4 tagged with a Strep II tag and three FLAG tags* ***C)*** *CHIKV WT SGR, which encodes the four WT non-structural proteins in ORF-1 and a luciferase reporter in place of the structural proteins in ORF-2 and* ***D)*** *CHIKV nsP4-3XF SGR, which encodes four non-structural proteins including a nsP4 tagged with a Strep II tag and three FLAG tags and the luciferase marker in ORF-2.*

2.2 Primers

Primers (Merck Life Science Limited) used in the cloning and sequencing of the CHIKV nsP4-3XF SGR are shown in **Table 1.** Primers were reconstituted in nuclease free water and stored at -20°C.

Primer	Sequence
Forward	5’ CTACGACGTGGATCAGAGAAG 3’
Reverse	5’ TATATCCTAGGTAGCTGATTAGTGTTTAGA 3’

Table 1: *Forward and reverse primer sequences used in cloning and sequencing of CHIKV nsP4-3XF SGR.*

2.3 Cell culture

RD, BHK-21 and Huh7 cells were maintained in a humidified 37°C incubator with 5% CO_2. Mammalian cells were maintained in complete media, consisting of Dulbecco’s Modified Eagle Medium (DMEM, Sigma Aldrich) supplemented with 100 IU penicillin/mL, 100 µg streptomycin/mL and 10% foetal bovine serum (FBS, Invitrogen). To passage, cells at 70-80% confluency were washed with phosphate buffered saline (PBS), incubated with trypsin-EDTA (Sigma Aldrich) for 5 minutes, diluted with complete media and split to the appropriate concentration.

C6/36 *Ae. albopictus* mosquito cells were maintained in a 28°C incubator. Mosquito cells were maintained in Lebowitz's media, supplemented with 10% FBS, 10% tryptose phosphate broth, 100 IU penicillin/mL and 100 µg streptomycin/mL. To passage, the cells were washed with PBS, then harvested by scraping into 5ml media, before diluting to the appropriate concentration.

2.4 Gel electrophoresis

DNA and RNA size and quality were analysed by agarose gel electrophoresis. Native agarose DNA gels were made up of 0.8% agarose in 1X TAE buffer (40 mM Tris, 20 mM acetate, 1 mM EDTA) with the addition of 1:10000 SybrSafe DNA gel stain (Invitrogen). DNA samples were mixed with 6X DNA gel stain (NEB) before native agarose gel electrophoresis in 1X TAE buffer, against a 1kb Plus DNA ladder (NEB), at 80 V for 60 minutes.

RNA integrity and size was analysed by denaturing agarose formaldehyde MOPS gel electrophoresis. A 1% agarose gel in MOPS buffer (40mm MOPS, 10 mM sodium acetate, 1mM EDTA) was prepared, with the addition of 1.7% formaldehyde (v/v) and 1:10000 SYBRSafe DNA gel stain. RNA was mixed with 2X RNA loading dye (Invitrogen) and heated at 55°C for 10 minutes. Samples were run against a Millenium RNA Marker (Ambion) at 80 V for 60 minutes.

Agarose gels were visualised by UV transillumination on a GelDoc Go System (BioRad).

2.5 Plasmid transformations

1 µg of plasmid DNA was mixed with *Mix and Go* competent DH5α *E. coli* cells (Zymo Research Corporation) and incubated for 5 minutes before spreading on an agar plate supplemented with 0.1 µg/mL ampicillin and incubated at 37°C overnight.

Colonies were picked, inoculated into 5 mL LB broth and incubated at 37°C for 16-18hrs while being shaken at 180 rpm. Plasmid DNA was extracted from colonies using GeneJET Plasmid Miniprep or Maxiprep kits (Thermo Fisher) as per the manufacturer's instructions and analysed by native agarose gel electrophoresis and nanodrop spectroscopy.

2.6 Plasmid digestion

For linearisation of CHIKV infectious virus and SGR plasmids, 10 µg of plasmid was incubated with 2 µl NotI-HF and 10X CutSmart buffer (NEB) overnight at 37°C and analysed by native agarose gel electrophoresis.

For restriction digestions of the SGRs, 1 µg of plasmid was incubated with 1 µl EcoRI and AvrII (NEB) each and 10X CutSmart buffer for four hours at 37°C, then analysed by native agarose gel electrophoresis.

2.7 DNA purification

Linearised plasmids were purified using Monarch® PCR & DNA Cleanup Kit (5 µg) (NEB) as per the manufacturer's instructions.

Plasmid digest products were separated by native agarose gel electrophoresis and the desired products were then excised from the gel under blue light. Excised bands were purified by Monarch® DNA Gel extraction kit, as per the manufacturer's instructions.

2.8 Touchdown PCR

PCR was performed using nsP4-FLAG virus plasmid DNA as a template to amplify the sequence to be used in the cloning for the nsP4-3XF SGR. The forward primer was designed to attach upstream of the native EcoRI site in ORF-1. The reverse primer was designed to attach to a sequence downstream of nsP4 and engineered to include an AvrII site in the final product.

Eight reactions were set up, as shown in **Table 2.** Reactions using Q5 enzyme were made up as followed: 5 µl 5X Q5 reaction buffer, optional 5 µl 5X Q5 High GC enhancer, 0.5 µl 10mM dNTPs, 1.25 µl each 10 µM forward and reverse primer, 250 ng template DNA, 0.25 µl Q5 DNA polymerase and nuclease free water up to 25 µl. Reactions using Phusion HiFi enzyme were made up as followed: 4 µl 5X Phusion Hifi reaction buffer or 4 µl 5X Phusion GC buffer, 0.4 µl 10mM dNTPs, 1 µl each 10 µM forward and reverse primer, 250 ng template DNA, 0.6 µl DMSO, 0.2 µl Q5 DNA polymerase and nuclease free water up to 20 µl.

Touchdown PCR was performed using the conditions: 98°C for 10s, 98°C for 10s, 65°C for 30s (with a decrease in temperature by 1°C every cycle until the final temperature of 55°C), 72°C for 10s per kb, for 30 cycles, then 72°C for 10 minutes.

The product was analysed by native agarose gel electrophoresis, digested with AvrII and EcoRI as previously described, and purified using a Monarch® PCR & DNA Cleanup Kit (5 µg) (NEB), for use in cloning.

Reaction	Enzyme	Template (nsP4-FLAG)	Buffer
I	Q5	Linear plasmid	Q5 buffer
II	Q5	Linear plasmid	Q5 buffer & GC enhancer
III	Q5	Uncut plasmid	Q5 buffer
IV	Q5	Uncut plasmid	Q5 buffer & GC enhancer
1	Phusion	Linear plasmid	Hifi buffer
2	Phusion	Linear plasmid	GC buffer
3	Phusion	Uncut plasmid	Hifi buffer
4	Phusion	Uncut plasmid	GC buffer

Table 2: *Makeup of eight different reactions tested to amplify nsP4-FLAG sequence by touchdown PCR.*

2.9 Ligation

Prior to ligation, the WT SGR backbone was CIP treated with Quick CIP (NEB) for 15 minutes at 37°C.

The ligation reaction was assembled using T4 ligase buffer (NEB), T4 DNA ligase (NEB), vector (WT SGR backbone) and insert DNA (PCR product) and nuclease-free water up to 20 µl. Vector and insert DNA were combined at a ratio of 1:3, with the total DNA in the reaction being ≤ 100ng. The reaction was incubated at 16°C overnight and transformed into bacteria as previously described. Colonies were picked, inoculated into LB broth, and mini-prepped. Plasmid DNA was then digested using AvrII and EcoRI and analysed by native agarose gel electrophoresis.

2.10 Sequencing of PCR products

Development of the nsP4-3XF SGR was confirmed by analysis by Sanger sequencing and whole plasmid sequencing by NG3 sequencing (Eurofins Genomics). Sequencing results were analysed using SnapGene software, version 8.0.1.

2.11 RNA *in vitro* transcription

Capped RNA for the transfection of the SGRs and virus infection was produced using an mMessage mMachine SP6 Transcription kit (Thermo Fisher). 1 µg linearised DNA was added to 2 µl 10X reaction buffer, 10 µl NTP/CAP, 2 µl Enzyme Mix, 1 µl RNasin® Plus ribonuclease inhibitor (Promega) and nuclease free water up to 20 µl. The mixture was incubated at 37°C for 4 hours, before the addition of 1 µl Turbo DNase (Thermo Fisher) for 15 minutes. Samples were purified using RNA Clean & Concentrator spin columns (Zymo).

2.12 RNA transfection into mammalian and mosquito cells

Cells were seeded at the appropriate concentration ($2.5x10^5$ mammalian cells per well on 6 well plates, $1x10^5$ mammalian cells or $2x10^5$ mosquito cells per well on 12 well plates), in complete media and grown overnight to 80% confluency. Before transfection, media was removed, cells were washed once with PBS and in incubated in opti-MEM (Thermo Fisher Scientific) at 37°C for 10 minutes.

Cells were transfected using Lipofectamine 2000 (Thermo Fisher Scientific) as per the manufacturer's instructions. For example, when using 12 well plates, 2 µl Lipofectamine 2000 was incubated in 98 µl opti-MEM per well, as was 1 µg RNA (or equivalent volume of DEPC water for the mock transfection) with opti-MEM made up to 100 µl per well. After a five-minute incubation, the RNA and Lipofectamine

mixtures were combined, incubated at room temperature for 25 minutes, then spread evenly dropwise over each well. Cells were incubated at 37°C for 4 hours, washed with PBS and replaced with media. Cells were incubated at 37°C until harvested at the desired time point (4, 24 or 48 hours).

Cells were lysed with a range of lysis buffers, including passive lysis buffer (PLB, Promega), Glasgow lysis buffer (GLB, 10 mM PIPES pH 7.2, 120 mM KCl, 30 mM NaCl, 5 mM $MgCl_2$, 1% Triton X-100 and 10% glycerol), NET lysis buffer (20 mM Tris, 100 mM NaCl, 1mM EDTA, 0.1% Triton X-100), RIPA lysis buffer (25 mM Tris HCl pH 7.6, 150 mM NaCl, 1% NP-40, 1% sodium deoxycholate, 0.1% SDS) and Pierce IP lysis Buffer (Thermo Fisher, 25 mM Tris-HCl pH 7.4, 150 mM NaCl, 1 mM EDTA, 1% NP-40, 5% glycerol). All buffers excluding PLB were supplemented with cOmplete, EDTA-free Protease Inhibitor Tablets (Roche).

2.13 Luciferase assay

Luciferase signal was measured using the Luciferase Assay System (Promega). 20 µl of samples lysed in PLB was mixed with 100 µl Reporter Lysis Buffer and measured using the FLUOStar Omega microplate reader and analysed using Optima (BMG LABTECH). Statistical analysis was performed using GraphPad Prism, version 10.4.1.

2.14 Cell electroporation with viral RNA

All BSL-3 infectious virus work was carried out by trained members of the Tuplin lab (Yuqian Liu, Zhaoxia Pang and Samuel Dobson). 400 µl of cell suspension ($3x10^6$ cells/ml in DEPC PBS) was mixed with 1 µg virus RNA in a cuvette. The cuvette was electroporated at 260 V, the cells were resuspended in 10 ml complete media and

seeded into a T75 flask. The supernatant was removed at 48 hpi, the cells were washed once with PBS, then lysed with PLB.

2.15 Plaque assay

Cells were seeded onto a 6 well plate at a concentration of 4 x10^5 cells/well and grown overnight. Cells were washed in PBS, each well was infected with virus (10-fold dilutions of 10^{-1} to 10^{-6}) and incubated at 37°C for 1 hour before removal of the virus. Cells were washed with PBS and overlayed with 1 ml 1.6% Methyl Cellulose and 1 ml complete media. Cells were then incubated at 37°C for 72 hours to allow plaque formation. The supernatant was removed from the wells and 1 ml 4% formaldehyde was added to each well to incubate for 15 minutes. Cells were then stained with 0.25% crystal violet for 30 minutes, then gently washed with water until plaques could be counted. Viral titre was calculated using the formula below:

$$Viral\ titre\ (PFU/ml) = \frac{number\ of\ plaques}{dilution\ factor\ x\ volume\ (ml)}$$

2.16 SDS-PAGE and western blotting

Cell lysate samples were separated by SDS-PAGE. Proteins were run through a two-part gel: a 5% stacking gel (30% acrylamide, 1M Tris-HCl pH 6.8, dH_2O, 10% SDS, 10% ammonium persulphate (APS) and TEMED) and 10% resolving gel (30% acrylamide, 1.5 M Tris-HCL pH 8.8, dH_2O, 10% SDS, 10% APS and TEMED). Samples were mixed with 5X SDS loading buffer (4% SDS, 40% glycerol, 10% β-mercaptomethanol, 120 mM Tris-HCL pH 6.8 and bromophenol 59 blue) and boiled for 10 minutes at 96°C. Samples were loaded into the gel and run against 3 µl Color Prestained Protein Standard Broad Range (NEB). The gel was run in 1X SDS-PAGE running buffer (50 mM Tris, 380 mM glycine, 0.1% SDS) at 155V for 90 minutes.

Transfer of the proteins to a PVDF membrane was achieved by assembling the gel and methanol soaked PVDF membrane between two pieces of filter paper soaked in Towbin transfer buffer ((5 mM Tris, 192 mM glycine and 20% methanol). The membrane assembly was then loaded into a Transblot-Turbo (BioRad) and transferred at 25V for an hour. Once transferred, the membrane was blocked in 5% skim milk in TBS Tween (TST) (25 mM Tris, 137 mM NaCl, 0.1% Tween-20) for 30 minutes and washed once in TBST and incubated in the primary antibody (diluted to the appropriate concentration in TBST, **Table 3)** at 4°C overnight.

The membrane was washed three times in TBST and incubated in the fluorescent secondary antibody (diluted 1:10000 in TST) for 1-2 hours. The membranes were washed three additional times in TBST before being imaged by a Li-Cor Odyssey SA and analysed via Empiria Studio (version 3.2). A list of primary and secondary utilised in this study is shown in **Table 3.**

Antibody	Origin	Dilution
Rabbit α nsP4	GeneTex	1:1000-1:5000
Mouse α β-actin	Sigma Aldritch	1:2000
Rabbit α β-actin	BioRad	1:1000
Rabbit α nsP4	Andres Merits	1:200
Rabbit α nsP3	Andres Merits	1:2000
Mouse α FLAG	Sigma Aldritch	1:1000
Mouse α HSP90	Abcam	1:2000
Rabbit α TMEM45B	Abcam	1:1000
Donkey α Rabbit 680RD	Li-Cor	1:10000
Donkey α Rabbit 800W	Li-Cor	1:10000
Donkey α Mouse 800W	Li-Cor	1:10000
Goat α Mouse 680RD	Li-Cor	1:10000

Table 3: *Origin of antibodies employed for detection of protein by western blot, and dilution used.*

2.17 Co-immunoprecipitation of nsP4

BHK-21 cells were seeded into 6 wells plated and transfected with nsP4-3XF SGR RNA as previously described**.** Cells were washed twice in PBS, lysed with 75 µl IP lysis buffer per well, scraped and rocked at 4°C for 30 minutes. Cells were then spun down at 4°C for 15 minutes at maximum speed, and the supernatant was removed from the cell debris.

FLAG-specific magnetic agarose beads were utilised for isolation of nsP4-3XF, and the any proteins bound in a complex with nsP4 (ThermoFisher). Briefly, 50 µl bead slurry was added to 300 µl cell lysate and incubated for 2 hours while rotating. The beads were then placed on a magnet, the supernatant was removed and stored as the flow-through sample at -20°C. The beads were resuspended and washed in 500 µl IP lysis buffer twice and once in 100 µl PBS. After washing, the beads were resuspended in 50 µl 1X SDS loading buffer, boiled for 10 minutes at 96°C, then replaced on the magnet to move the supernatant to a new tube (Co-IP sample). Co-IP samples were analysed by western blotting.

Chapter 3: Optimisation of an anti-nsP4 antibody for Co-immunoprecipitation.

3.0 Introduction

The low solubility and stability of recombinant CHIKV nsP4 have hindered the study of the protein, its interactions with host cell proteins, and the development of treatments targeting CHIKV RdRp. For example, the inability to express and purify recombinant CHIKV nsP4 means that no structure has been able to be experimentally determined. As such our current structural knowledge is based on computationally derived CHIKV nsP4 molecular structures and comparisons to related alphaviruses. This lack of structural information has hindered the development of CHIKV antivirals by rational design that target elements of the nsP4 structure.

Viruses will interact with numerous host proteins during infection, that control different aspects of their replication cycle. Elucidating these protein-protein interactions between viruses and the host cell can help uncover the mechanisms that regulate virus replication during infection and provide insight into the adaptations that viruses have evolved to evade the host antiviral response. These interactions can in turn be exploited as antiviral drug targets.

Protein interactions can be identified by Co-IP, in which the protein of interest is purified from the cell by binding to a specific antibody. Following purification, proteins bound in complex to the target protein can be identified by western blot or quantitative mass spectrometry. The production of antibodies targeting CHIKV nsP4 has been hindered by the challenges discussed earlier of expressing and isolating nsP4. Due to the nsP4's poor solubility and difficulties in expressing the protein, anti-

CHIKV nsP4 antibodies have only very recently become available for use in scientific research. As a result of this, very few host cell interactions of nsP4 during CHIKV replication have been identified and studied. Only one study has been published describing selection of antibodies against CHIKV nsPs, with the antibodies only being raised against one antigen, due to limitations expressing and isolating the proteins **(Kumar et al., 2015).** Some studies of nsP4 interactions have reported using non-commercial antibodies raised against nsP4 **(Rathore et al., 2013, Rathore et al., 2014, Scholte et al., 2015),** and commercially available anti-nsP4 antibodies have recently come to market. Recently, Genetex have marketed a polyclonal anti-CHIKV nsP4 antibody, but to date no published studies have reported use of this antibody. Additionally, the insertion of a protein tag, which is often used as an alternative approach to investigating protein interactions, has only recently been achieved in CHIKV nsP4. The inability to produce a tagged version of nsP4, in addition to the lack of efficient anti-nsP4 antibodies available, has hindered the study of the CHIKV nsP4 host interactome.

The initial aims of this project were to identify novel protein interactions of CHIKV nsP4 in mammalian cells during early virus replication events through the isolation of nsP4 protein complexes by Co-IP, which would then be analysed by tandem-mass-tag mass spectrometry (TMT-MS). In order to perform the Co-IP experiments, an anti-nsP4 antibody would first have to be optimised.

Co-IPs isolate target protein complexes by incubating the cell lysate of infected cells with an antibody against the target protein, and then purifying the complexes by binding the antibody to magnetic agarose beads. The beads can then be washed to remove unbound proteins from the sample, so that the target protein and its interactome can then be identified and quantified by TMT-MS.

TMT-MS involves the use of isobaric tags, which when intact have the same mass but have unique masses when ionised. Multiple samples can be analysed at once by using different tags for each sample, which are then mixed and analysed parallelly **(Zhang and Elias, 2017).** Intensities of each tag variant are analysed to quantify the abundance of each peptide present in the samples, with highly abundant proteins indicating a possible interaction with nsP4. These potential interactions would then undergo further experimentation to validate the potential interaction. Validation of potential interactions would include inhibition of the interacting protein by either siRNAs, shRNA or CRISPR-Cas9 to investigate the effect of the interaction on CHIKV replication. Additionally, immunofluorescent assays (IFAs) can be performed to determine if nsP4 and the potential interacting protein colocalise at different timepoints in the lifecycle of CHIKV. Co-IPs using antibodies against the potential interacting protein could also be performed, and then analysed by western blotting to determine if nsP4 is immunoprecipitated alongside the interactor, reinforcing the likelihood of an interaction between the two proteins.

3.1.0 Determining the most efficient antibody for nsP4 detection

The majority of this project was carried out using CHIKV SGRs to study nsP4 during early replication events of the virus replication cycle. Previous work categorizing CHIKV nsP4 interactions, as outlined above in **Section 1.5.3,** have been carried out in cells expressing nsP4 *in trans,* but use of the CHIKV SGR allows for nsP4 interactions to be captured when nsP4 is within the context of the RC. The WT SGR expresses all four nsPs but does not express any structural proteins due to the replacement of the genes encoding the structural proteins in ORF-2 with a firefly luciferase reporter gene **(Figure 10).** Production of the structural CHIKV proteins is dispensable for CHIKV RNA replication, and the use of SGRs allows CHIKV

translation and genome replication (referred to as SGR replication) to be studied in isolation to entry, structural protein expression, assembly and the release of mature virions. Cell lysates transfected with the SGR can be assayed for luciferase signal, to ensure transfected cells have high levels of SGR replication, as expression of the firefly luciferase marker is dependent on the expression of ORF-2 via sub-genomic RNA expression.

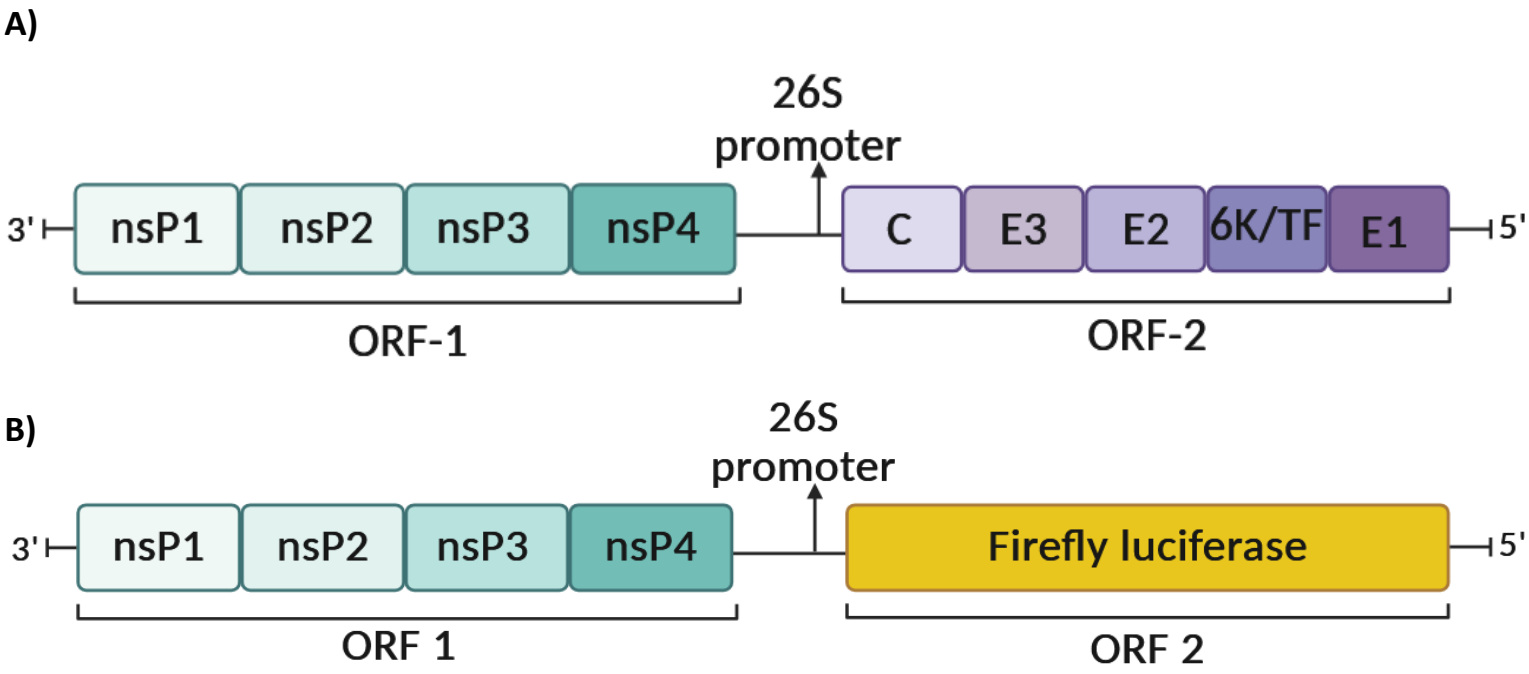


Figure 10: Schematic representation comparing the CHIKV genome and WT CHIKV SGR. A) *CHIKV genome encoding the four nsPs from ORF-1 and five structural proteins from ORF-2.* ***B)*** *The CHIKV SGR, which is derived from a CHIKV isolate of the ECSA lineage. The WT SGR produces WT versions of the nsPs from ORF-1, but does not encode structural proteins, as ORF-2 has been replaced with a firefly luciferase reporter gene.*

To optimise anti-nsP4 antibodies, attempts were made to detect nsP4 by western blotting to determine the antibody efficiency. Lysates of cells that had been transfected with the WT SGR were analysed by SDS-PAGE, then transferred to a membrane which was incubated with an anti-nsP4 antibody. Incubation of the

membrane with a fluorescently linked secondary antibody was then conducted to detect nsP4, as the secondary antibody binds to the primary anti-nsP4 antibody.

RD cells were initially chosen for study of nsP4 interactions using the WT SGR. RD cells are derived from human muscle cells, which CHIKV infects in humans, and so these cells are a biologically relevant model. Moreover, RD cells are highly permissive for infection by CHIKV and the CHIKV SGR, making them a useful tool for the study of CHIKV **(Roberts et al., 2017, Sun et al., 2024).**

3.1.1 Results

3.1.1.1 Genetex antibody did not detect nsP4 from RD cell lysates

Due to the availability of the commercial Genetex antibody, it was initially chosen for optimisation of nsP4 detection. The Genetex antibody was tested for its ability to detect nsP4 by western blot analysis. RD cells were transfected with the WT SGR in 12 well plates, lysed at 24 hours post transfection (hpt) and assayed for luciferase signal to confirm SGR replication. Once SGR replication was confirmed, the lysates were analysed by western blotting to test for nsP4 detection.

The membranes were initially incubated in a 1:5000 dilution of Genetex anti-nsP4, as per the manufacturer's recommendation for western blotting **(Figure 11A)**. However, as no band for nsP4 was observed, different concentrations of the Genetex antibody were then tested. These dilutions went up to 1:1000 but yielded no positive results for nsP4 detection **(Figure 11B).** As the loading control, β-actin, was clearly detected from the lysates, and replication of the SGR *in vitro* was confirmed by luciferase assaying, it was concluded that the Genetex anti-nsP4 antibody could not efficiently detect CHIKV nsP4 from cell lysates. As nsP4 could not be detected using the Genetex antibody, alternative antibodies targeting nsP4 were then considered.

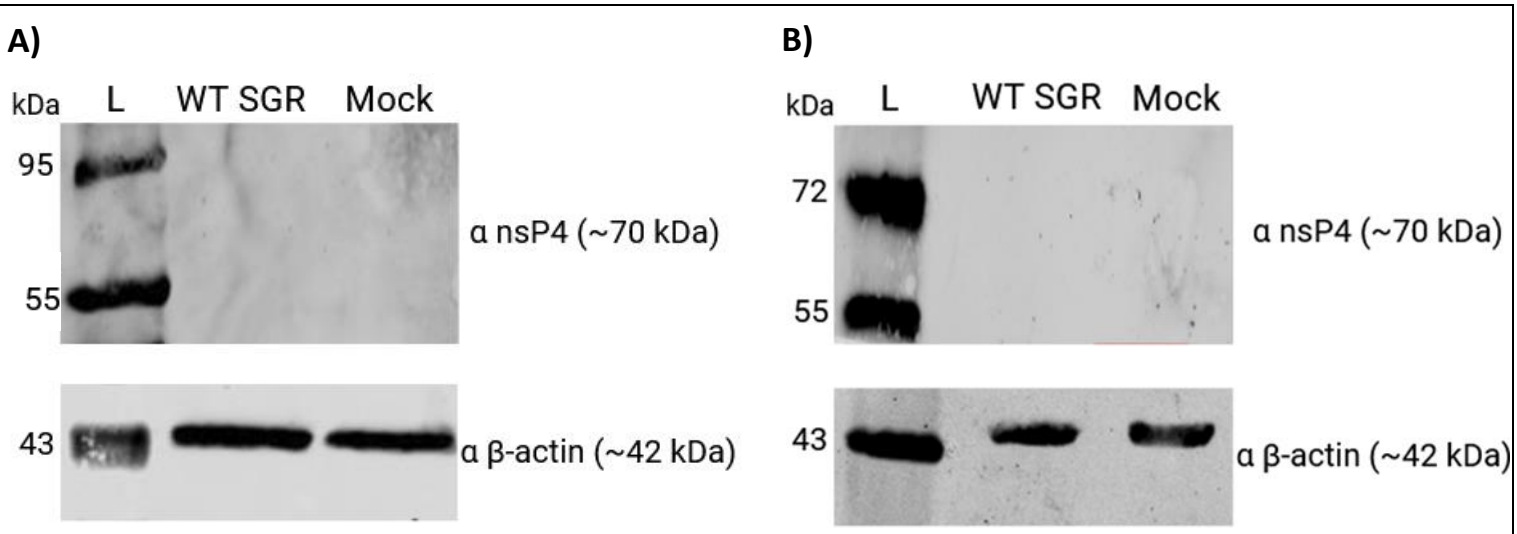


Figure 11: Western blots testing the Genetex anti-nsP4 antibody. *Western blots analysing the efficiency of Genetex anti-nsP4 antibody to detect nsP4 (~70 kDa). RD cells were transfected with WT SGR and harvested at 24 hpt using PLB. Mock lysates were transfected with DEPC water in place of SGR RNA as a negative control for nsP4 detection. Detection of β-actin (~42 kDa) was used as a loading control.* ***A)*** *RD cell lysates incubated in anti-nsP4 antibody at a dilution of 1:5000 (top) and anti-β-actin (bottom).* ***B)*** *RD cell lysates incubated anti-nsP4 antibody at a dilution of 1:1000 (top) and anti β-actin (bottom). nsP4 was not detected from cell lysates at either dilution.*

3.1.1.2 The Merits anti-nsP4 antibody did not consistently detect nsP4 from RD cell lysates

An anti-nsP4 antibody made in house by the Andres Merits lab was identified from published studies on nsP4 interactions **(Rathore et al., 2013, Rathore et al., 2014, Scholte et al., 2015)**. This antibody was acquired from the Merits lab, with Prof Andres Merits indicating that while this antibody was the most efficient available at detecting nsP4 (in comparison to commercial antibodies) it was still an inefficient antibody **(personal communication)**.

The Merits anti-nsp4 antibody was tested at a concentration of 1:200 on recommendation from Prof Andres Merits **(Figure 12A)**. In addition to RD cells lysed with PLB, a sample lysed with IP lysis buffer was additionally tested as it was hypothesised that nsP4 detection by western blotting could be hindered by release of the protein from the RC associated cell membrane during lysis. PLB and IP lysis buffer are both non-ionic lysis buffers, but PLB uses Triton-X as a detergent while IP lysis buffer uses NP-40. While a band the approximate size of nsP4 could be detected from transfected RD cells lysed with PLB, no equivalent band was detected in the transfected RD cells lysed with IP lysis buffer. However, a faint band was detected in the untransfected negative control sample (mock sample) lysed in IP lysis buffer at the approximate size of nsP4 **(Figure 12A)**. This band was hypothesised to be a non-specific interaction of the antibody with a host protein of a similar size to nsP4 or the result of contamination of the mock lysate from a neighbouring lane.

Transfection of RD cells was then repeated and tested for nsP4 detection by western blotting, to confirm the previous positive result could be replicated using the Merits anti-nsP4 antibody. As PLB was positive for nsP4 detection previously, the cells were lysed in this buffer. However, despite being able to be detected previously by this antibody, nsP4 was not observed in this repeat **(Figure 12B)**.

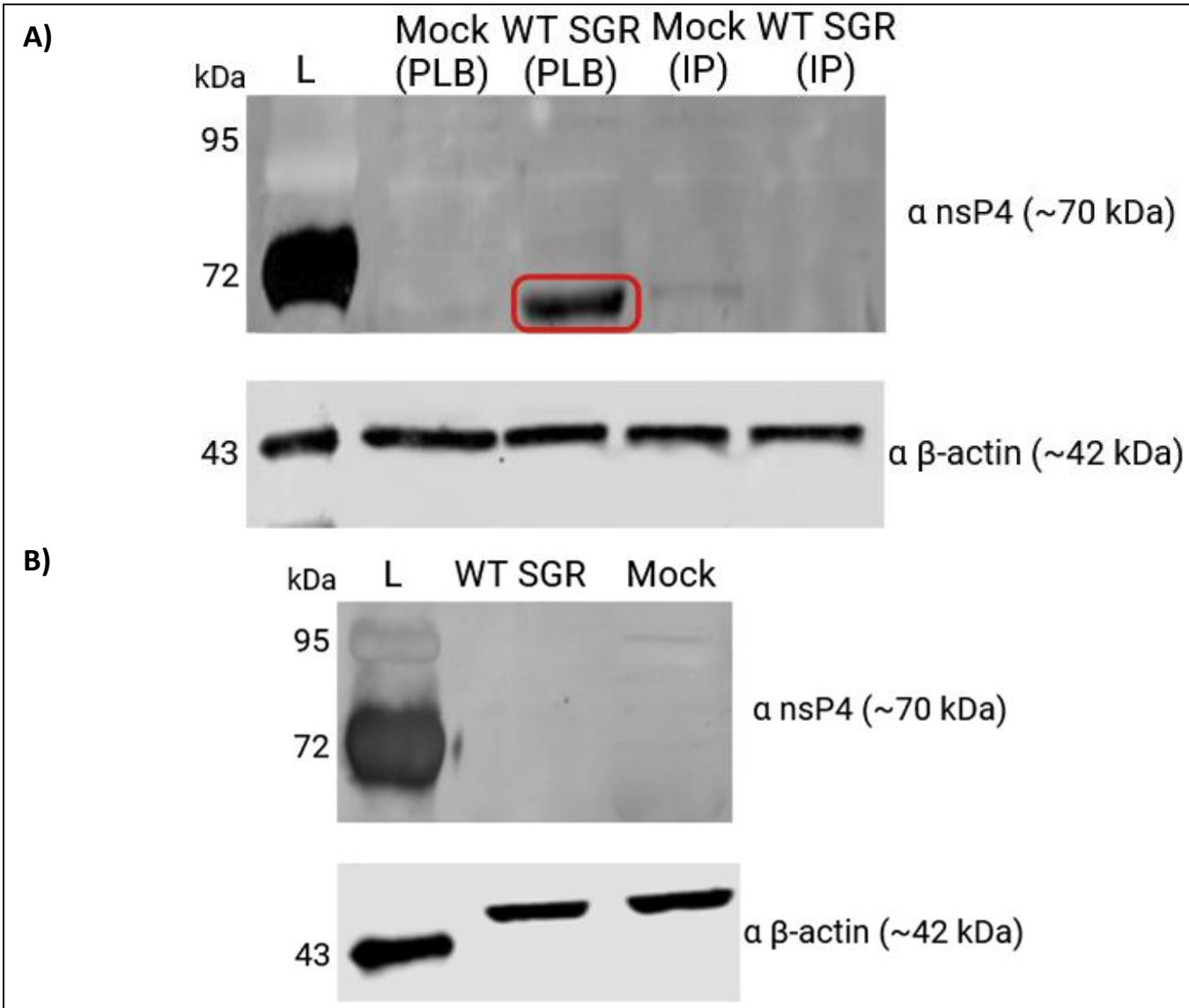


Figure 12: Western blots testing the Merits anti-nsP4 antibody. *Western blot analysing efficiency of anti-nsP4 antibody produced by the Merits lab to detect nsP4 (~70 kDa). RD cells were transfected with WT SGR and harvested at 24 hpt. Mock lysates were transfected with DEPC water in place of SGR RNA as a negative control for nsP4 detection. Detection of β-actin (~42 kDa) was used as a loading control.* ***A)*** *RD cell lysates lysed with either PLB or IP lysis buffer. Blots were incubated in anti-nsP4 antibody (top) and anti-β-actin (bottom). nsP4 was detected in SGR transfected cells lysed in PLB. Faint band around nsP4 size was also detected in Mock IP lysis buffer lane.* ***B)*** *RD cell lysates lysed with PLB. Blots were incubated in anti-nsP4 antibody (top) and anti-β-actin (bottom). nsP4 was not detected in this repeat.*

A potential reason for the inconsistency in nsP4 detection was hypothesised to be insufficient levels of SGR replication in the transfected RD cells. The CHIKV SGR has previously been shown to replicate to higher levels in BHK-21 cells **(Roberts et al., 2017).** Consequently, it was theorised that if BHK-21 cells were used in place of RD cells, nsP4 would be able to be expressed to a higher level and thus could be more consistently detected by the Merits anti-nsP4 antibody. Alternatively, it was thought that as nsP4 is bound to the cell membrane via the RC during genome replication, the lysis buffers used in these experiments may not be sufficiently stringent to completely release nsP4 from the RC associated membrane during lysis. If nsP4 was not released from the membrane, when cell debris was removed from the lysate, nsP4 would also be removed from the samples. To investigate this hypothesis, five different lysis buffers were analysed to determine if nsP4 could be consistently detected when cells were lysed with a particular lysis buffer. Additionally, detection of nsP4 was compared in RD, BHK-21 and Huh7 cells (which like RD cells are a biologically relevant model for the study of CHIKV). From this, it could then be determined if the use of a particular lysis buffer or cell type resulted in the efficient and replicable detection of nsP4 using the Merits anti-nsP4 antibody. Using this cell line or lysis buffer, the Merits anti-nsP4 antibody could then be optimised for nsP4 detection by western blotting and later downstream applications - including Co-IPs to isolate nsP4 protein complexes that would then be sent for TMT-MS.

3.2.0 Optimisation of cell line and lysis buffer for CHIKV nsP4 detection by western blot analysis

The CHIKV SGR is known to replicate at different rates in different cell lines **(Roberts et al., 2017).** RD cells are biologically relevant model for the study of CHIKV genome replication, originating from a human muscle rhabdomyosarcoma

cell. However, the CHIKV SGR replicates to higher titres in other cell lines. BHK-21 cells, which originate from baby hamster kidney cells, replicate the CHIKV SGR to a very high level. As they are not human in origin, the proteome is not the most biologically relevant when investigating protein-protein interactions that can then be exploited for drug treatments in humans. Huh7 cells are an alternative biologically relevant human cell line to RD cells, originating from a hepatocyte carcinoma cell, and are widely used as a model in CHIKV studies **(Sun et al., 2024, Müller et al., 2019).** However, Huh7 cells have been shown to replicate the SGR to a lower level than RD or BHK-21 cells. These three cell lines were selected, transfected with the WT SGR, and lysed with five different lysis buffers to examine if one cell type or lysis buffer permitted the detection of nsP4 more consistently.

As nsP4 is membrane bound, low stringency lysis buffers may not release the protein from the membrane, which would permit the protein from being detected by western blotting. The type of lysis buffer used has previously been reported to be important when solubilising nsP4 **(Chen et al., 2017).** PLB was originally employed to lyse transfected RD cells, as it is compatible with assaying for luciferase signal, to ensure sufficient replication of the SGR has taken place before western blotting. However, PLB may not be the most effective buffer for release of proteins from cell membrane as it has a relatively low stringency. RIPA buffer has strong detergent agents and was therefore hypothesised as likely to release nsP4 from the RC associated membrane, however the high stringency may also disrupt protein interactions, making it less ideal for downstream applications. Alternatively, IP lysis buffer is a modified recipe of RIPA buffer containing no SDS, so may be better for downstream applications, as it is non-ionic and would be less likely to disrupt protein complexes. GLB was also chosen as an alternative between the low stringency PLB

and high stringency RIPA buffers. Lastly, NET lysis buffer was selected for testing, as it was identified in a paper studying the interaction of nsP4 with eIF2α **(Rathore et al., 2013).**

3.2.1 Results

In an attempt to increase the concentration of nsP4 in the lysates, 6 well plates were transfected with SGR RNA, instead of 12 well plates, though the volume of lysis buffer was kept at 100 µl. The samples lysed in PLB were assayed for luciferase signal to ensure SGR replication and translation had occurred in these cells. Detection of nsP3 was also included to confirm CHIKV SGR replication in non-PLB lysed samples, which are not compatible with the luciferase assay system. Viral lysate was produced from RD cells infected with an infectious recombinant CHIKV that encoded nsP4 tagged with a Strep II tag and three FLAG tags in the N-terminus (nsP4-FLAG virus) and was obtained as a positive control for the detection of CHIKV nsPs. This lysate was chosen over WT virus lysate due to the increased availability of nsP4-FLAG viral lysate.

3.2.1.1 RD cells

A band of the approximate size of nsP4 (~70 kDa) was detected in all transfected RD cells, regardless of lysis buffer used. nsP3 was also detected from several transfected lysates, but despite having a predicted size of ~60 kDa, the protein was observed around the 72 kDa protein marker. This size difference has been reported for nsP3 previously and is due to post translational modifications **(Teppor et al., 2021).**

The viral lysate had the highest band intensity detected for both nsP4 and nsP3 on both blots analysing the RD cell lysates **(Figure 13A and Figure 13B),** although the

band for nsP3 on the first blot **(Figure 13A)** was notably weaker than on the second **(Figure 13B).** Interestingly, two bands were detected in the nsP4-FLAG virus lysate lane when incubated with the Merits anti-nsP4 antibody, with a larger band at the 72 kDa protein marker and a second band just below the marker, which was of similar size to the bands detected in the samples transfected with the WT SGR **(Figure 13)**. The larger band was consistent with nsP4-FLAG, indicating a size shift of nsP4 from the WT at ~70 kDa (seen in the SGR transfected lysates) to version encoding the Strep II and FLAG tags at ~76 kDa. However, it is unknown what the second smaller band was.

RD cells lysed with GLB had the highest level of band intensity detected for both nsP4 and nsP3 out of the samples transfected with the WT SGR, with a similar level of intensity detected in the viral lysate as the GLB lysed RD cells **(Figure 13A)**. This indicates that GLB is the most efficient lysis buffer at releasing the nsPs, including nsP4, from the RC associated cell membrane. Surprisingly, the band intensity of nsP4 was higher than nsP3 in this sample **(Figure 13A)**, as low levels of nsP3 would indicate a lower level of SGR replication and therefore a weaker signal detected for nsP4. The second highest level of band intensity detected for both nsP3 and nsP4 was observed in the RD cell lysates lysed with RIPA buffer, the most stringent buffer used in this experiment **(Figure 13B).**

Though a band the approximate size of nsP4 was observed in all other transfected cell lysates, the bands were much less intense than the bands detected from the viral lysate or transfected cells lysed with GLB or RIPA buffer. The transfected RD sample lysed with NET lysis buffer only had weak levels of signal detected for both nsP4 and nsP3 **(Figure 13A),** indicating that it is not an efficient lysis buffer for the release of the membrane associated nsPs from the cell, despite having previously

been used in a paper studying the interactions of nsP4 in HEK293 cells **(Rathore et al., 2013).** Additionally, though a band the approximate size of nsP4 could be detected from transfected RD cells lysed with IP lysis buffer, nsP3 was not observed from this same sample **(Figure 13B).** As previously seen **(Figure 12A),** a band the approximate the size of nsP4 was detected in the mock sample lysed with IP lysis buffer, and additionally in the mock sample lysed with RIPA buffer. The band of similar size to nsP4 in the mock samples lysed with IP and RIPA lysis buffers was not detected in the mock samples lysed with PLB, GLB or NET lysis buffer **(Figure 13A).** The presence of the band the size of nsP4 in the IP lysed transfected RD cells, despite the lack of nsP3 signal, may indicate that the band observed is non-specific binding of the Merits anti-nsP4 antibody, comparable to the band seen in the untransfected IP lysed sample, as the lack of nsP3 indicates a lack of sufficient SGR replication in this sample **(Figure 13B).**

Samples lysed with PLB were assayed for luciferase signal, which indicated that replication and translation of the SGR had occurred. Although not all samples could be assayed to confirm replication, the lack of signal for nsP4 and nsP3 in the PLB lysate, in which SGR replication was confirmed, could indicate that the release of nsPs in lysates with low signal intensities detected for nsP4 and nsP3 was not sufficient **(Figure 13A)**. However, nsP4-FLAG infected cells were also lysed with PLB, and both nsPs were able to be efficiently detected from this lysate **(Figure 13)**. This suggests that the nsP4-FLAG replicates more efficiently than the CHIKV SGR in RD cells, resulting in the higher concentration of nsPs in the virus lysate, which can then be more easily detected by anti-nsP3 and the Merits anti-nsP4 antibodies, indicating that both nsP4 concentration and release from the RC associated membrane play a role in nsP4 detection by western blotting.

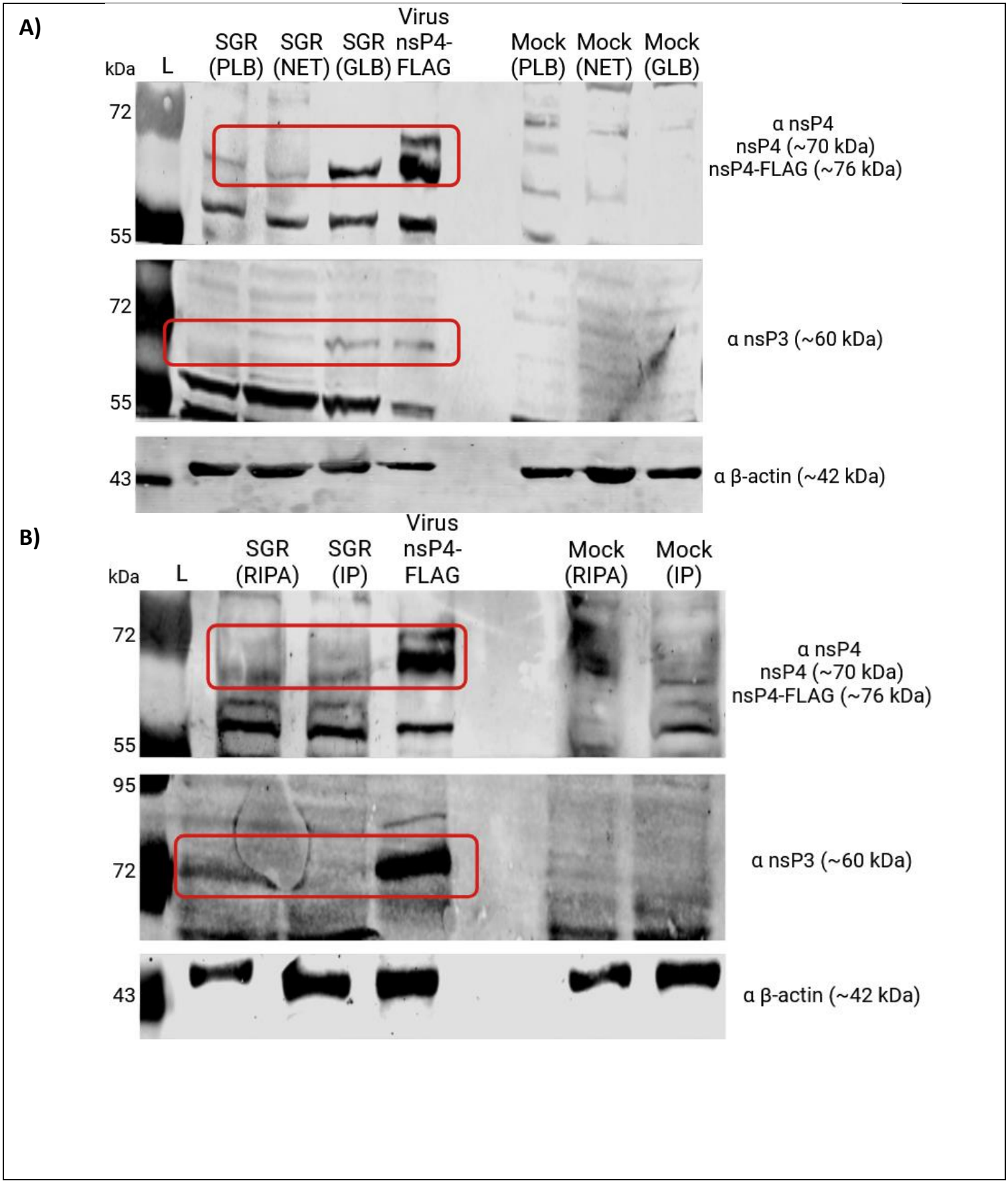

A)
kDa
L
SGR (PLB)
SGR (NET)
SGR (GLB)
Virus nsP4-FLAG
Mock (PLB)
Mock (NET)
Mock (GLB)
72
55
α nsP4
nsP4 (~70 kDa)
nsP4-FLAG (~76 kDa)
72
55
α nsP3 (~60 kDa)
43
α β-actin (~42 kDa)
B)
kDa
L
SGR (RIPA)
SGR (IP)
Virus nsP4-FLAG
Mock (RIPA)
Mock (IP)
72
55
α nsP4
nsP4 (~70 kDa)
nsP4-FLAG (~76 kDa)
95
72
α nsP3 (~60 kDa)
43
α β-actin (~42 kDa)

Figure 13: Western blots comparing efficiency of nsP4 detection from RD lysates, lysed using a variety of lysis buffer. *Western blots analysing the efficiency of nsP4 detection (~70 kDa) from RD cell lysates when lysed with a variety of lysis buffers. RD cells were transfected with WT SGR and harvested at 24 hpt using either PLB, NET lysis buffer, GLB, RIPA or IP lysis buffer. Mock lysates were transfected with DEPC water in place of SGR RNA as a negative control for nsP detection. RD cells infected with CHIKV nsP4-FLAG and harvested at 24 hpi in PLB was used as the nsP4-FLAG viral lysate. nsP3 (~60 kDa) was used a positive control for SGR replication, nsP4-FLAG lysate was used as a positive control for nsP detection, and β-actin (~42 kDa) was detected as a loading control.* ***A)*** *RD cells lysed with PLB, NET lysis buffer and GLB and tested for nsP4 (top), nsP3 (middle) and β-actin (bottom).* ***B)*** *RD cells lysed with RIPA or IP lysis buffer, and tested for nsP4 (top), nsP3 (middle) and β-actin (bottom). nsP4 was detected in all SGR transfected lysates and in the viral lysate. nsP3 was detected in all SGR and viral lysates, with the exception of the sample lysed with IP lysis buffer.*

3.2.1.2 BHK-21 cells

Unlike the RD cell samples transfected with the WT SGR **(Figure 13),** nsP4 was not detected in all transfected BHK-21 lysates, despite the WT SGR replicating to a higher level in BHK-21 cells than RD cells **(Figure 14).** Additionally, the level of nsP4 and nsP3 band intensity in the nsP4-FLAG lysate was very weak on the first blot comparing the BHK-21 samples **(Figure 14A).** This is divergent from the pattern seen on the other blots comparing nsP4 detection from samples lysed with different buffers **(Figures 13-15),** as all other blots, including the second blot analysing the BHK-21 samples **(Figure 14B),** had the highest signal intensity for nsP4 and nsP3 from the nsP4-FLAG lysate. As the viral lysate used for all lysis buffer experiments came from the same sample, the lack of signal in this blot cannot be attributed to a lack of nsP4 expression or release from the RC associated membrane, as the protein was strongly detected on the other blots. Instead, the reduced signal in nsP4 detection in the nsP4-FLAG sample may be attributed to the inefficiency of the Merits

anti-nsP4 antibody, which has previously been shown to not replicably detect nsP4 from cell lysates **(Figure 12).**

As with RD cells, BHK-21 cells lysed with PLB had a band of low signal intensity for both nsP4 and nsP3, despite high levels of SGR replication being confirmed by luciferase assay **(Figure 14A).** This reinforces that PLB is not a sufficiently stringent lysis buffer for the release of nsPs from the RC associated membrane in SGR transfected cells.

Clear bands representing nsP4 were observed in transfected BHK-21 cells lysed with NET, IP, and RIPA lysis buffers **(Figure 14)**. Indeed, the sample lysed with NET lysis buffer was observed to have had a higher intensity than bands for both nsP4 and nsP3 than those observed in the viral nsP4-FLAG lysate on the first blot **(Figure 14A).** The high band intensity observed in the transfected BHK-21 cells lysed with NET lysis buffer was inconsistent with the result of the RD sample lysed with NET lysis buffer, as that sample was observed to have weak signal intensity for both nsP4 and nsP3 and was therefore assumed to inefficiently release the nsPs from the cell membrane **(Figure 13A)**. The increase in signal intensity may result from the increased level of SGR replication in BHK-21 cells in comparison to RD cells.

Consistent with the previous results **(Figure 12A and 13B)**, a band of similar size to nsP4 was found in the mock samples lysed with RIPA and IP lysis buffers **(Figure 14B).** Though the band of approximate size to nsP4 in the transfected RD sample lysed with IP lysis buffer was presumed to be non-specific binding due to the lack of nsP3 signal **(Figure 13B)**, it is not possible to conclude if the band found ~70 kDa in the BHK-21 sample lysed with IP lysis buffer was the result of non-specific binding from this data as nsP3 was efficiently detected in the BHK-21 sample **(Figure 14B)**.

Moreover, this demonstrates that the identity of the band the size of nsP4 also cannot be confidently identified in any sample lysed with either IP or RIPA lysis buffer from these experiments **(Figure 13-15)**.

nsP4 was not observed in BHK-21 cells lysed with GLB **(Figure 14A)**, despite RD cells lysed with this buffer having the strongest signal for nsP4 out of all the RD cell samples tested **(Figure 13)**. As nsP3 was detected from the BHK-21 sample lysed with GLB, the absence of signal for nsP4 was not the result of a lack of SGR replication in these cells **(Figure 14A)**. The lack of nsP4 detection may instead be the result of degradation of nsP4 by the host proteome *in vitro,* or an issue with the efficiency of the Merits anti-nsP4 antibody used, as the nsP4-FLAG viral lysate had a reduced level of nsP4 detected on this blot **(Figure 14A)**, despite being strongly detected from the other blots performed **(Figures 13, 14B and 15)**.

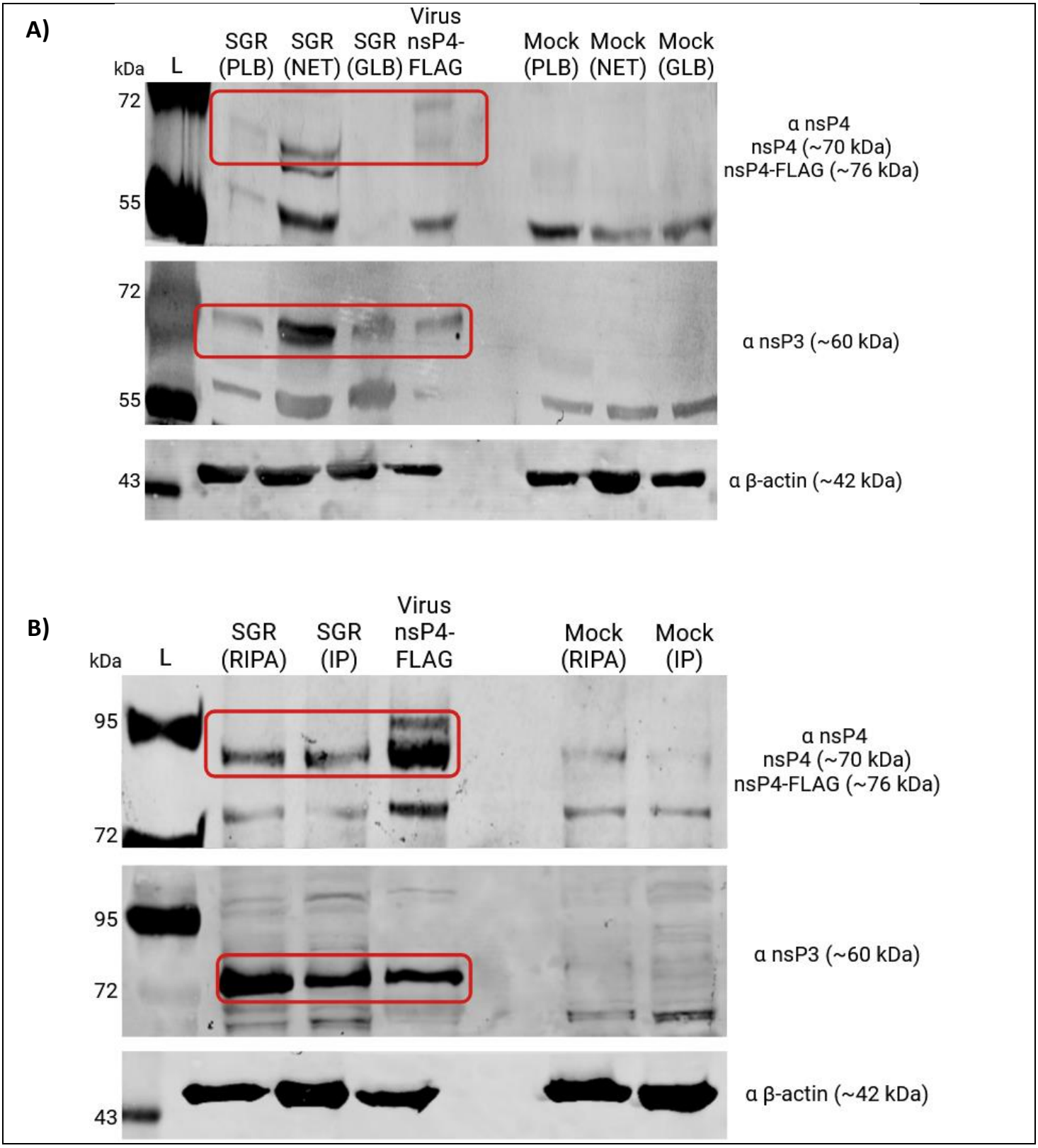

A)
kDa
L
SGR (PLB)
SGR (NET)
SGR (GLB)
Virus nsP4-FLAG
Mock (PLB)
Mock (NET)
Mock (GLB)
72
55
α nsP4
nsP4 (~70 kDa)
nsP4-FLAG (~76 kDa)
72
55
α nsP3 (~60 kDa)
43
α β-actin (~42 kDa)
B)
kDa
L
SGR (RIPA)
SGR (IP)
Virus nsP4-FLAG
Mock (RIPA)
Mock (IP)
95
72
α nsP4
nsP4 (~70 kDa)
nsP4-FLAG (~76 kDa)
95
72
α nsP3 (~60 kDa)
43
α β-actin (~42 kDa)

Figure 14: Western blots comparing efficiency of nsP4 detection from BHK-21 lysates, lysed using a variety of lysis buffer. *Western blots analysing the efficiency of nsP4 detection (~70 kDa) from BHK-21 cell lysates when lysed with a variety of lysis buffers. BHK-21 cells were transfected with WT SGR and harvested at 24 hpt using either PLB, NET lysis buffer, GLB, RIPA or IP lysis buffer. Mock lysates were transfected with DEPC water in place of SGR RNA as a negative control for nsP detection. RD cells infected with CHIKV nsP4-FLAG and harvested at 24 hpi in PLB was used as the nsP4-FLAG viral lysate. nsP3 (~60 kDa) was used a positive control for SGR replication, nsP4-FLAG lysate was used as a positive control for nsP detection, and β-actin (~42 kDa) was detected as a loading control.* ***A)*** *BHK-21 cells lysed with PLB, NET lysis buffer and GLB and tested for nsP4 (top), nsP3 (middle) and β-actin (bottom).* ***B)*** *BHK-21 cells lysed with RIPA or IP lysis buffer, and tested for nsP4 (top), nsP3 (middle) and β-actin (bottom). nsP3 was detected in all SGR transfected lysates and in the viral lysate. nsP4 was detected in all SGR transfected and viral lysates, with the exception of the sample lysed with GLB.*

3.2.1.3 Huh7 cells

As with the blots comparing the RD cell samples **(Figure 13),** nsP4-FLAG viral lysate had the highest level of band intensity detected for nsP4 and nsP3 on both blots comparing the Huh7 cell samples **(Figure 15)**.

Though the level of SGR replication is lower in Huh7 than in BHK-21 cells, a higher level of signal intensity for nsP4 was detected in transfected Huh7 cells lysed with PLB **(Figure 15A)** than in BHK-21 cells lysed with PLB **(Figure 15B)**. The higher level of signal observed in PLB lysed cells in both RD and Huh7 cells over BHK-21 cells, which replicates the SGR to higher levels than either human cell line, reinforces the inefficient and unreliable release of the nsPs by PLB. **(Figures 13A, 14A and 15A)**.

Transfected Huh7 cells lysed with NET lysis buffer were observed to have bands for both nsP4 and nsP3, though the signal intensity of these bands was not particularly strong **(Figure 15A)**. The signal intensity of the band found in the sample lysed with GLB was similar to that observed in the NET buffer lysed cells, despite no nsP3 being detected in the GLB sample, which would imply that insufficient levels of SGR replication had taken place **(Figure 15A)**. As there was not a non-specific band the approximate size of nsP4 present in the mock sample lysed with GLB, the band found in the transfected sample was assumed to be nsP4, despite the lack of nsP3 signal. Moreover, the cells lysed in PLB, which were confirmed to have sufficient levels of SGR replication by luciferase assay, had a similar level of signal intensity for nsP4, which suggests that SGR replication was occurring in the cells lysed with GLB **(Figure 15A)**. As cells lysed with GLB or NET lysis buffer only inconsistently able to detect nsP4 and nsP3 by western blotting across the three cell types tested, these buffers were determined to be insufficient for the reliable release of nsP4 from cell lysates for the study of nsP4 interactions **(Figures 13A, 14A and 15A)**.

Transfected Huh7 samples lysed with RIPA and IP lysis buffer were both observed to have low levels of signal intensity for both nsP3 and nsP4 **(Figure 15B)**. Additionally, the non-specific band of similar size to nsP4 was observed in the untransfected mock samples lysed with RIPA and IP lysis buffers.

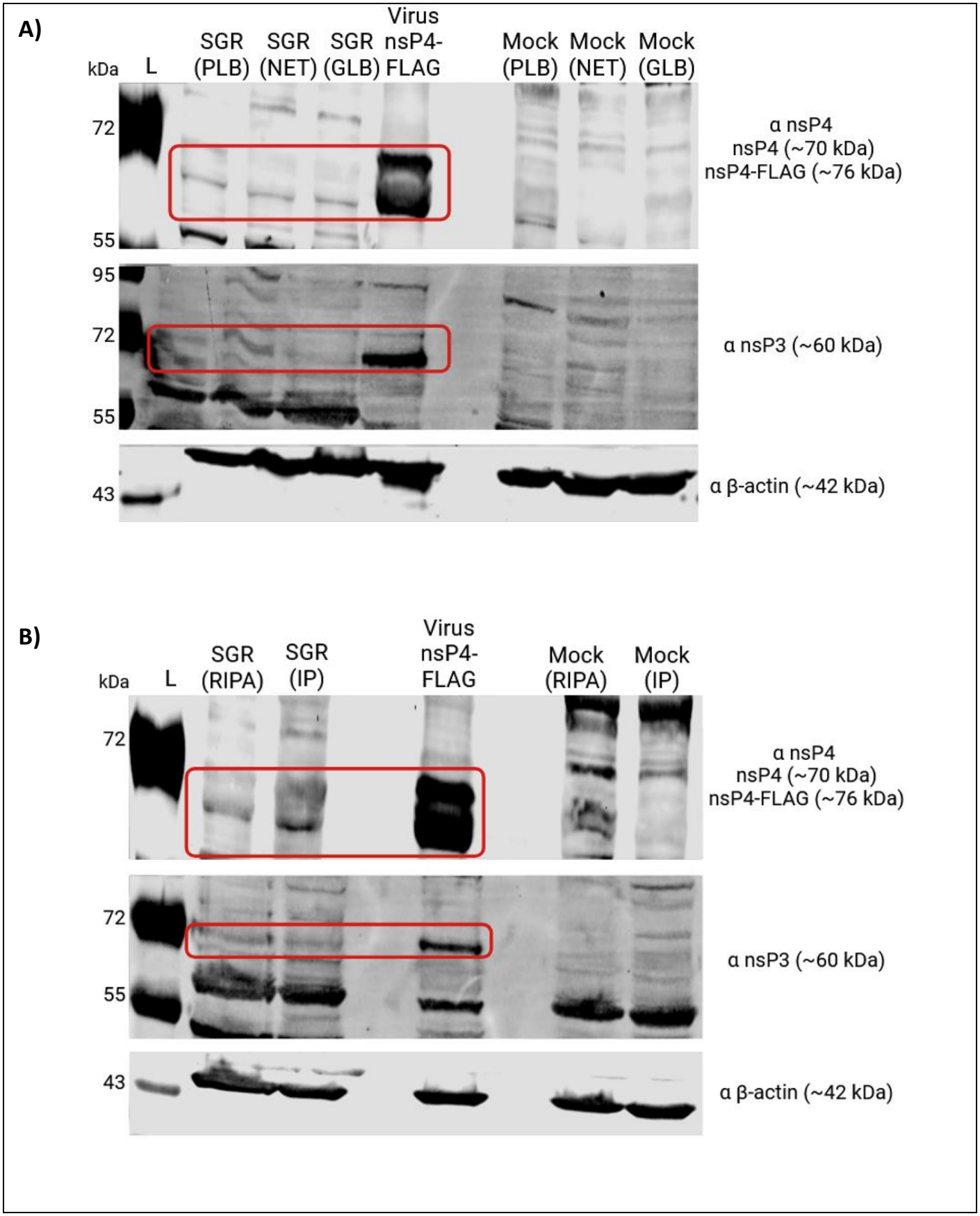

A)
kDa
L
SGR (PLB)
SGR (NET)
SGR (GLB)
Virus nsP4-FLAG
Mock (PLB)
Mock (NET)
Mock (GLB)
72
55
95
72
55
43
α nsP4
nsP4 (~70 kDa)
nsP4-FLAG (~76 kDa)
α nsP3 (~60 kDa)
α β-actin (~42 kDa)
B)
kDa
L
SGR (RIPA)
SGR (IP)
Virus nsP4-FLAG
Mock (RIPA)
Mock (IP)
72
72
55
43
α nsP4
nsP4 (~70 kDa)
nsP4-FLAG (~76 kDa)
α nsP3 (~60 kDa)
α β-actin (~42 kDa)

Figure 15: Western blots comparing efficiency of nsP4 detection from Huh7 lysates, lysed using a variety of lysis buffer. *Western blots analysing the efficiency of nsP4 detection (~70 kDa) from Huh7 cell lysates when lysed with a variety of lysis buffers. Huh7 cells were transfected with WT SGR and harvested at 24 hpt using either PLB, NET lysis buffer, GLB, RIPA or IP lysis buffer. Mock lysates were transfected with DEPC water in place of SGR RNA as a negative control for nsP detection. RD cells infected with CHIKV nsP4-FLAG and harvested at 24 hpi in PLB was used as the nsP4-FLAG viral lysate. nsP3 (~60 kDa) was used a positive control for SGR replication, nsP4-FLAG lysate was used as a positive control for nsP detection, and β-actin (~42 kDa) was detected as a loading control.* ***A)*** *Huh7 cells lysed with PLB, NET lysis buffer and GLB and tested for nsP4 (top), nsP3 (middle) and β-actin (bottom). Huh7 cells lysed with RIPA or IP lysis buffer, and tested for nsP4 (top), nsP3 (middle) and β-actin (bottom). nsP4 and nsP3 were detected in all SGR transfected lysates and in the viral lysate.*

3.2.1.4 Summary of Results

While PLB and NET lysis buffer were observed to have bands the size of both nsP4 and nsP3 in all three cell lines **(Figure 13-15),** the bands observed were often of very low intensity. This may be the result of insufficient release of nsP4 from the cell membrane, which could cause difficulties studying nsP4 and its interactome in further studies.

The use of RIPA buffer also resulted in bands of the same size as nsP4 and nsP3 being observed in all three cell lines **(Figures 13-15)**. However, a band the same size of nsP4 was additionally observed in all mock sample lysed with RIPA buffer, and therefore it could not confidently be said that RIPA released nsP4 in all three experiments. IP buffer was also observed to possess a band of similar size to nsP4 in the mock lysates, and therefore nsP4 could not confidently be identified in the cell lysates lysed with this buffer **(Figures 13-15)**. This was especially obvious in the RD cell sample lysed with IP lysis buffer, as a band the size of nsP4 was observed in the transfected RD cell sample despite no nsP3 being detected **(Figure 13B)**.

While the RD cell sample lysed with GLB produced a band of strong intensity the size of nsP4 **(Figure 13A)**, it only produced a weak band in Huh7 cells, and a band

for nsP4 was not observed in BHK cells lysed with this buffer **(Figure 14A)**. Moreover, in Huh7 cells lysed with GLB, nsP3 was not observed **(Figure 15A)**.

The results of each buffers ability to release nsP4 and nsP3, by observation of a band for each protein by western blotting, is summarise in the table below **(Table 4)**.

Cell Type	**nsP4 positive**	**nsP4 negative**	**nsP3 positive**	**nsP3 negative**
RD	PLB, NET, GLB+, IP, RIPA+	N/A	PLB, NET, GLB, RIPA+	IP
BHK	PLB, NET+, IP+, RIPA+	GLB	PLB, NET+, GLB, RIPA+, IP+	N/A
Huh7	PLB, NET, GLB, IP, RIPA	N/A	PLB, NET, RIPA, IP	GLB

Table 4: *Summary of western blots attempting to determine the most efficient lysis buffer for nsP4 release and detection* ***(Figures 13-15)****. (Buffer+) indicates a strong band was observed.*

3.3 Discussion

The production of antibodies capable of targeting CHIKV nsP4 has been hindered due to several problems purifying both endogenous and recombinantly expressed protein, including low concentrations of nsP4 within the cell due to the opal stop codon, instability of the protein as a result of degradation, and the fact that full-length recombinant nsP4 is insoluble when expressed in prokaryotic cells. As a result of this, few antibodies against nsP4 are available, and the ones that are available are not efficient at targeting nsP4.

Two antibodies were utilised to determine which was more efficient at detecting CHIKV nsP4 in lysates from RD cells transfected with the WT SGR. Despite testing different dilutions, nsP4 was not detected by the Genetex antibody in RD cells lysed with PLB **(Figure 11)**. As studies involving the use of an anti-nsP4 antibody produced by Merits lab have previously been published, the lab was contacted and kindly gifted their antibody for use in this project. Although this antibody was able to detect nsP4 from RD cell lysates once **(Figure 12A),** the antibody did not efficiently and reproducibly detect nsP4 from SGR transfected RD cells **(Figure 12B).**

The inconsistency in nsP4 detection by the Merits anti-nsP4 antibody was hypothesised to be the result of incomplete release of the RC from the cell membrane during lysis, which would limit the detection of nsP4, as release of nsP4 from the RC associated host membrane has previously been described as an issue when studying the protein **(Tomar et al., 2006).** Additionally, the level of SGR replication in RD cells was also theorised to play a role in nsP4 detection by western blotting. However, it is difficult to make conclusions about the most efficient method of lysing cells for release of nsP4 from the RC associated cell membrane, or the cell

type that should be used when studying nsP4 from this current study **(Figures 13-15)**. No method of lysing cells consistently detected nsP4, or indeed nsP3, by western blot analysis.

Detection of nsP3 was employed as a method of confirming replication and expression of the WT SGR from lysates that could not be assayed for replication by luciferase signal expression, due to incompatible lysis buffers being used. However, nsP3 was also inconsistently detected, with nsP3 only being detected very weakly in SGR transfected RD cells lysed with PLB despite SGR replication being confirmed by luciferase assay. While the lack of signal detected in PLB lysed samples could be attributed to the inefficient release of nsPs from the RC associated membrane during lysis, bands the approximate size of nsP4 were detected from the lysates of some cells, despite nsP3 not being detected from the same sample, including IP lysis buffer lysed RD cells and GLB lysed Huh7 cells **(Figures 13B and 15A)**. Moreover, bands the size of nsP4 were found in some samples with weak levels of signal intensity for nsP3 (such as IP buffer lysed Huh7 cells, **Figure 15B)**, then conversely nsP4 would be absent in samples observed to have strong bands for nsP3, such as the BHK-21 cells lysed with GLB **(Figure 14A)**. This indicates that the detection of nsP3 is not a good indicator for nsP4 expression, as one can be detected without the other.

The CHIKV anti-nsP antibodies employed were both unspecific, demonstrating lots of background that resulted in many unspecific bands being detected by western blotting **(Figures 13-15).** This could be attributed to the antibodies not being optimised for an efficient level of detection of the target proteins and minimal levels of background. Determination of the optimal dilution of an antibody depends on the specificity of the antibody used and the levels of the target antigen in the sample.

Polyclonal antibodies, like the nsP4 antibodies utilised in this study, are less specific than monoclonal antibodies, but have the advantage of being able to bind to multiple epitopes and are therefore more likely to detect the target in the sample. This amplification of nsP4 detection when using these polyclonal antibodies for western blotting can be seen as an advantage, as nsP4 is produced at lower concentrations than other CHIKV proteins. However, as they are less specific there is a higher chance of background on blots or unspecific binding to other proteins.

The bands of a similar size to nsP4, that were observed only in the mock samples lysed with RIPA and IP buffers **(Figures 12A, 13B, 14B and 15B),** could be the result of these buffers releasing or solubilising a host protein that the other buffers tested did not. The presence of this band in the mock samples may suggest that nsP4 is not being detected from WT SGR transfected cells that were lysed by RIPA or IP lysis buffer, as the band detected at the size of nsP4 in the SGR transfected samples may be the unspecific band observed in the mock samples. For example, though the SGR transfected RD lysate did not detect nsP3, which would imply that replication did not occur in these cells, a band was observed at the approximate size of nsP4 **(Figure 13B)**. The identity of the band in the SGR transfected samples and mock samples could be analysed by excising the band from the SDS-PAGE gel and performing mass spectrometry to identify the bands observed in these samples, to confirm if the band observed at the approximate size of nsP4 in the transfected cells lysed with either RIPA or IP lysis buffers was nsP4.

Cells lysed with RIPA buffer most consistently detected nsP3 and nsP4 from cell lysates, however there are several issues when using this buffer. Firstly, the level of signal intensity detected for nsP4 in both RD and Huh7 cells is still relatively weak **(Figures 13B and 15B)**, which poses a problem when studying the nsP4 host

interactome as they are still not efficiently detected in biologically relevant cells. Secondly, as RIPA buffer is incompatible with the luciferase assay system, it is not possible to confirm SGR replication in cells lysed with this buffer. While nsP3 was employed in this study as a method to confirm SGR replication in lysates that could not undergo luciferase assay, there are issues with the use of nsP3 to determine replication in cells, as outlined above. Additionally, RIPA lysis solubilises a band the approximate size of nsP4 in untransfected lysates, which the Merits anti-nsP4 antibody will also non-specifically bind to **(Figures 13B, 14B and 15B)**. This means that nsP4 cannot be confidently detected from cells lysed with RIPA buffer, even in the presence of nsP3, as any band detected at the size of nsP4 may be unspecific binding to a host cell protein. Moreover, RIPA buffer is not an ideal buffer for downstream applications, as its high stringency can disrupt protein interactions during lysis.

As more infectious CHIKV nsP4-FLAG lysate was available during these experiments than WT virus lysate, it was used as a positive control in the lysis buffer optimisation experiments. The virus lysate produced two distinct bands when incubated with the Merits anti-nsP4 antibody. As the larger band was consistent with the approximate size of nsP4-FLAG (~76 kDa), it was assumed that the larger band was nsP4. Interestingly, the larger band had a much stronger signal when detected by anti-FLAG antibodies in comparison to the second smaller band, but the smaller band was also detected by anti-FLAG antibodies **(Appendix Figure 1)**. This could suggest that the smaller band represented a version of nsP4 that was altered or cleaved in a way that decreased the efficiency of nsP4-FLAG to be able to be detected by anti-FLAG antibodies. Alternatively, the second band may have represented a non-specific host protein of a similar size to WT nsP4.

Though insufficient lysis and SGR replication can partially explain the inconsistency of nsP4 detection, the efficiency of the Merits anti-nsP4 antibody itself may also take some responsibility in the inability of nsP4 to be reliably detected by western blotting. For example, though nsP4-FLAG had the highest level of signal intensity when detecting nsP4 for the large majority of blots performed **(Figures 13-15)**, the first blot analysing the BHK-21 samples poorly detected nsP4 despite the same sample being used for all blots **(Figure 14B)**. As the lack of signal cannot be attributed to poor replication, insufficient release, or *in vitro* degradation as none of these issues arose on the other blots, the lack of signal may be attributable to the antibody itself.

The unspecific nature of nsP4 antibodies would provide issues for downstream applications, even if the detection of nsP4 by western blotting could be optimised. From the high amount of background detected during western blotting **(Figures 13-15)**, it is logical to assume that use of the Merits anti-nsP4 antibody in Co-IP would also capture many non-specific proteins. Consequently, it was concluded from these results that the antibodies currently available for nsP4 are not efficient enough to proceed with the original Co-IP aim of this project. It was decided that a novel system to detect and isolate nsP4 was to be engineered. In this novel system, nsP4 would be tagged with a Strep II tag and three tandem FLAG tags, so that detection of nsP4 by western blotting and isolation of nsP4 protein complexes by Co-IP could be performed by targeting the FLAG tags rather than nsP4 directly.

Chapter 4: Engineering and validating a novel system for studying the nsP4 host interactome by development of a sub-genomic replicon expressing recombinant nsP4 tagged with Strep II and FLAG

4.0 Introduction

An infectious clone of CHIKV, derived from the ECSA lineage, was identified from a published study **(Scholte et al., 2015).** This virus encoded recombinant nsP4 that was tagged with a Strep II tag and three tandem FLAG tags in the N-terminus. In order to study nsP4 during early CHIKV replication events, an SGR that encoded the FLAG tagged nsP4, from the virus reported by **Scholte et al. (2015),** was to be engineered **(Figure 16)**. Once validated, this SGR encoding nsP4 tagged with the Strep II and FLAG tags (nsP4-3XF) could be optimised for use in Co-IPs and TMT-MS to allow the identification of novel host-nsP4 protein interactions.

Previous attempts to tag nsP4 have had deleterious effects on CHIKV replication. For example, it has been shown that alphavirus nsP4 cannot be tagged in the C-terminus as it is deleterious to the polymerase function, preventing virus replication **(Rubach et al., 2009).** While not essential for polymerase activity, the N-terminal domain of nsP4 improves replication efficiency. The dispensability of the nsP4 N-terminus makes it a more suitable target for the insertion of a peptide tag than the C-terminus. Additionally, the N-terminus is unstructured and flexible in comparison to the more rigid hand structure of the C-terminus, meaning that insertion of a peptide tag in this region would be less likely to disrupt an essential structure **(Tan et al., 2022b)**. However, certain mutations of the N-terminus have also been shown to effect replication efficiency, so consideration must be taken when deciding the

placement of the tag to ensure sufficient polymerase activity of the tagged recombinant protein. For example, location of a Strep-FLAG tag in a region which resulted in the disruption of the formation of a flexible loop in RRV nsP4 completely inhibited polymerase function of the protein **(Tan et al., 2022b)**.

The first reported isolation of recombinant expressed alphavirus nsP4 used a SUMO tag placed in the N-terminus to stabilise the protein and prevent degradation by the N-end rule pathway **(Rubach et al., 2009).** Constructs expressing CHIKV nsP4 tagged with GFP have also previously been described and allowed Co-IPs and IFAs to take place by targeting GFP to isolate and detect nsP4 **(Rathore et al., 2013, Rathore et al., 2014)**. However, these constructs only expressed nsP4-GFP in *trans* and there are no reports of a virus or SGR expressing nsP4-GFP which would be required to study the nsP4 interactome during active CHIKV genome replication. Additionally, GFP tags are larger than other available peptide tags such as FLAG (~27 kDa and ~1 kDa respectively) and are therefore less likely to maintain native protein folding than smaller peptide tags.

The creation of FLAG tagged alphavirus nsP4 has previously been reported. Insertion of FLAG tags in the N-terminus of SINV nsP4 at both residues 92 and 109 were reported by **Tan et al. (2022b).** The insertion of these tags was reported to result in a relatively low decrease in SINV replication efficiency. However, ligation of the FLAG tag into the same residues in RRV nsP4 significantly inhibited polymerase activity **(Tan et al., 2022b).** This emphasises the importance of peptide tag placement in alphavirus nsP4, especially in the RdRp N-terminus, which is the least conserved region of nsP4 in the alphavirus genus, and so peptide tags placed in the same residues can have widely different effects on protein function between different alphaviruses **(Tan et al., 2022b).**

Protein tags have a variety of uses, including as a target in Co-IPs, for isolation of proteins for which specific antibodies are not available. **Tan et al. (2022b)** suggested that their SINV construct encoding a recombinant nsP4 tagged Strep-FLAG tagged nsP4 could be used for immunoprecipitations of nsP4 protein complexes, to characterise novel interactions of nsP4. Indeed, FLAG-specific beads are available for Co-IPs of FLAG tagged proteins, in addition to indirect isolation using FLAG antibodies and agarose beads.

The SGR containing nsP4-3XF was engineered by cloning the sequence encoding the tagged nsP4 from nsP4-FLAG virus into the existing WT SGR **(Figure 16).** Development of this new system would allow the study of the nsP4 host interactome, without the involvement of anti-nsP4 antibodies, by targeting the FLAG tags instead. To ensure that this new system could be used in the study of nsP4 host interactions, the level of SGR replication was tested by luciferase assay to ensure that sufficient levels of replication and translation were occurring. The detection of nsP4-3XF by anti-FLAG antibodies in place of the inefficient Merits anti-nsP4 antibody was then analysed by western blotting. Lastly, a preliminary Co-IP experiment was performed, to investigate if the new system could be used to isolate protein complexes involving nsP4 in the early replication events.

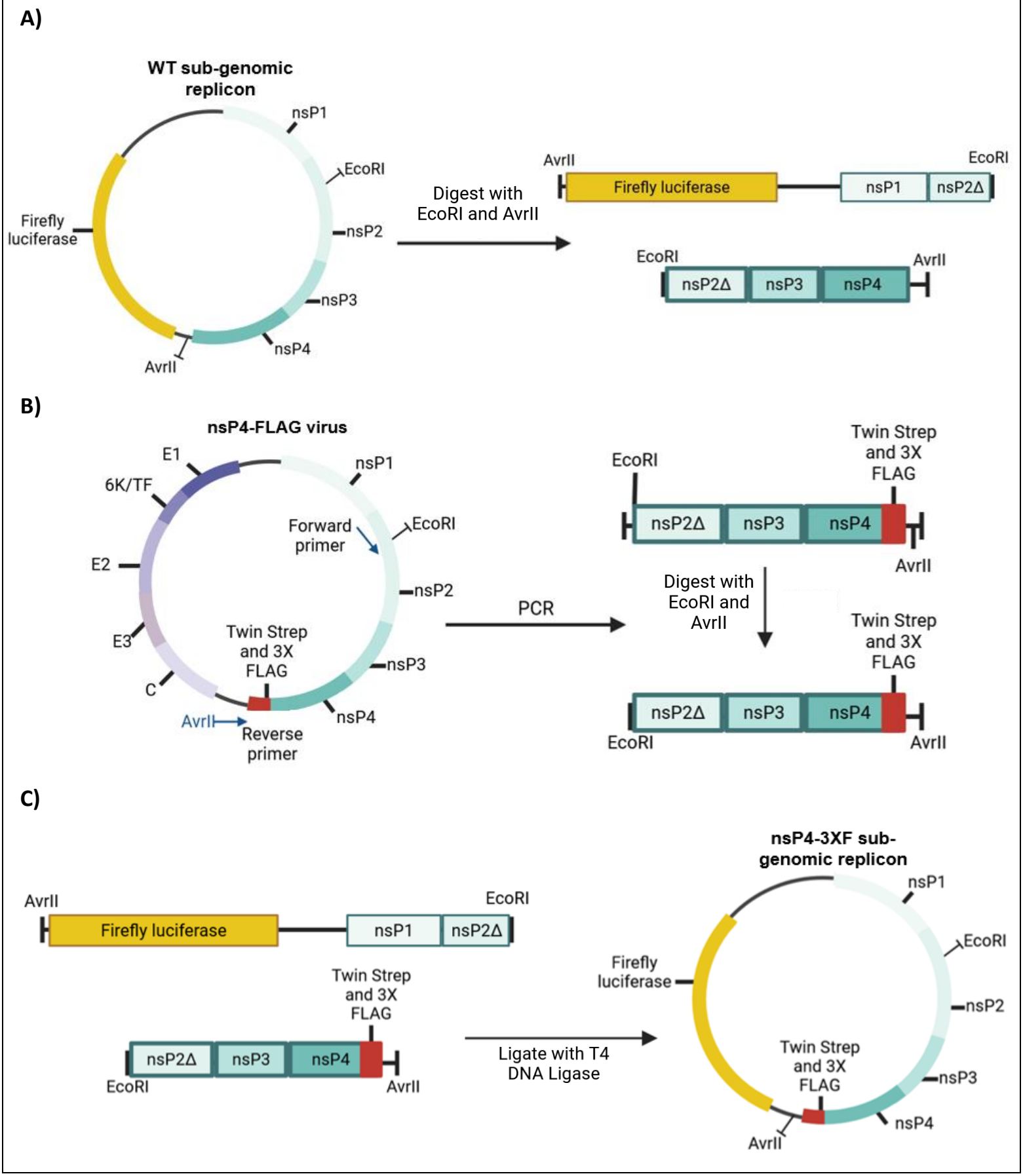

A)
WT sub-genomic replicon
nsP1
EcoRI
nsP2
nsP3
nsP4
AvrII
Firefly luciferase
Digest with EcoRI and AvrII
AvrII
EcoRI
Firefly luciferase
nsP1
nsP2Δ
EcoRI
AvrII
nsP2Δ
nsP3
nsP4
B)
nsP4-FLAG virus
E1
6K/TF
E2
E3
C
nsP1
EcoRI
nsP2
nsP3
nsP4
Forward primer
Twin Strep and 3X FLAG
AvrII
Reverse primer
PCR
EcoRI
Twin Strep and 3X FLAG
nsP2Δ
nsP3
nsP4
AvrII
Digest with EcoRI and AvrII
Twin Strep and 3X FLAG
nsP2Δ
nsP3
nsP4
EcoRI
AvrII
C)
AvrII
EcoRI
Firefly luciferase
nsP1
nsP2Δ
Twin Strep and 3X FLAG
nsP2Δ
nsP3
nsP4
EcoRI
AvrII
Ligate with T4 DNA Ligase
nsP4-3XF sub-genomic replicon
nsP1
EcoRI
nsP2
nsP3
nsP4
AvrII
Firefly luciferase
Twin Strep and 3X FLAG

Figure 16: Schematic representation of the process to engineer an SGR expressing recombinant nsP4 tagged with Strep II and FLAG. A) *Depicts the digestion of the WT SGR using EcoRI and AvrII to acquire the SGR backbone, which can be used as the vector in the ligation reaction* ***B)*** *Depicts the use of the nsP4-FLAG virus as a template for PCR to acquire a product that contains both the EcoRI site in nsP2 and an AvrII site downstream of the Strep II and FLAG tags, which can then digested with EcoRI and AvrII for use in cloning* ***C)*** *Depicts the WT SGR backbone (vector) and digested PCR product (insert) being ligated together to create a SGR expressing all four nsPs, including nsP4 tagged with Strep II and FLAG, and the luciferase reporter gene in place of the structural protein encoding genes in ORF-2.*

4.1.0 Engineering a novel system for detecting nsP4

4.1.1 Results

4.1.1.1 Touchdown PCR produced an insert for cloning containing nsP4 tagged with Strep II and three FLAG tags

To generate a CHIKV SGR expressing nsP4 tagged with a Strep II and three FLAG tags, the backbone of the WT SGR was first isolated for use in cloning. Two unique restriction sites were identified in the WT SGR sequence, EcoRI - which was located within the sequence for nsP2 - and AvrII, located in the junction sequence between nsP4 and the luciferase reporter gene **(Figure 9C)**.

The WT SGR was digested at these two restriction sites, which was confirmed by native agarose gel electrophoresis **(Figure 17)**. The backbone of the WT SGR (AvrII>EcoRI, 7327 bp) was extracted from the gel and CIP treated to dephosphorylate the linearised backbone and prevent religation.

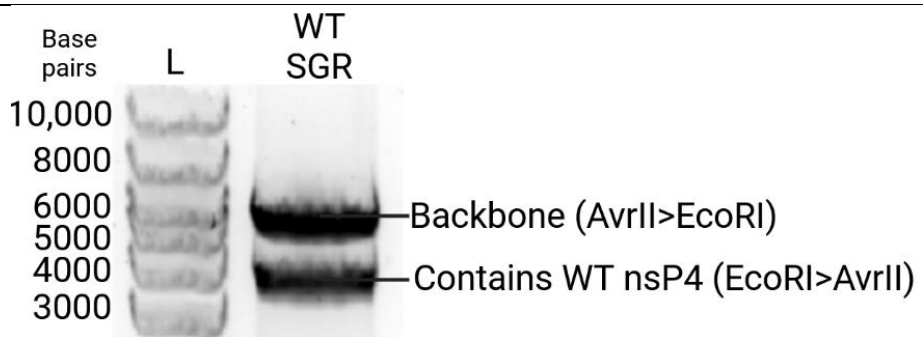


Figure 17: WT SGR digested with EcoRI and AvrII and separated by gel electrophoresis. *Digested WT SGR will produce two products: the backbone that will be used in the ligation reaction (7327 bp), and a second segment containing nsP4 (5397 bp).*

The insert containing nsP4 for use in cloning was generated by PCR amplification from the CHIKV nsP4-FLAG template and was engineered to contain an AvrII site downstream of the sequence encoding nsP4 in the final PCR product.

After several unsuccessful PCR experiments, touchdown PCR, in which the annealing temperature step is lowered gradually with each cycle, was performed. Eight different reactions were set up **(see Table 2 for details)**, using CHIKV nsP4-FLAG cDNA plasmid as a template in both linear and the uncut form. Additionally, two DNA polymerases were tested, Q5 and Phusion, as well as two different buffer set ups for each enzyme. Seven of the eight reactions successfully amplified a product of the expected size (5618 bp) and when analysed by native agarose gel electrophoresis indicated a band at the desired product size **(Reactions I-IV and 1-3, Figure 18).** However, as Reaction I had only a very faint band it was not extracted from the gel **(Figure 18)**. The other six successful reactions were extracted from the gel and digested with EcoRI and AvrII. Once the products were purified, they could then be used in cloning and be ligated into the WT SGR backbone between the same EcoRI and AvrII restriction sites.

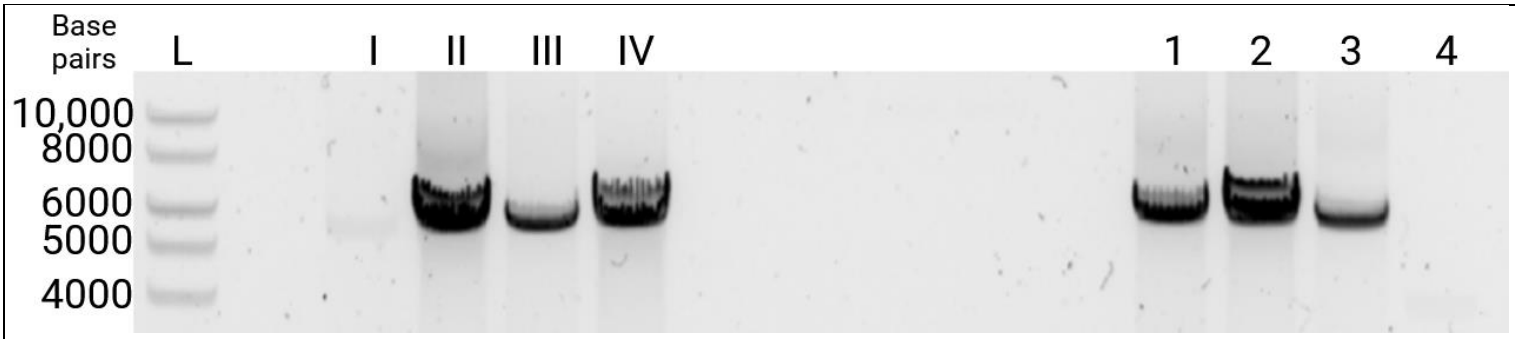


Figure 18: Products of touchdown PCR. *PCR was performed to amplify the segment of nsP4-FLAG virus containing the nsP4 tagged with Strep II and FLAG and add an AvrII site to the 5' so that the section contained the correct restriction sites for use in cloning. PCR was performed using two different enzymes: Q5 (I-IV) and Phusion (1-4). I-IV reactions used Q5 buffer, with II and IV containing the addition of GC enhancer buffer. Reactions 1 and 3 used Hifi buffer, while 2 and 4 used GC buffer. I, II,1 and 2 were amplified from a linear plasmid template, while III, IV, 3 and 4 were amplified from the uncut virus plasmid. Product had expected size of 5618 bp, bands were detected in lanes II-IV and 1-3.*

4.1.1.2 Ligation successfully produced a CHIKV SGR containing nsP4 tagged with Strep II and three FLAG tags

Two ligation reactions were performed using the PCR products of Reaction 1 and Reaction II **(Figure 18)**. The ligations were set up at a vector:insert ratio of 1:3 and incubated at 16°C overnight. The reactions were then transformed into competent *E. coli* cells, plated on agar plates supplemented with ampicillin, and incubated overnight at 37°C. Plate 1, which used the product from Reaction 1 as an insert, had 12 colonies, while Plate 2, which used the product from Reaction II, grew 15 colonies. From each plate, 9 colonies were picked and digested with EcoRI and AvrII and analysed by native agarose gel electrophoresis to determine the presence of the vector and insert **(Figure 19)**. The WT SGR was used as a positive control and allowed the comparison between the nsP4 containing regions (EcoRI-AvrII) of the

WT and new nsP4-3XF SGRs, which showed a size shift that indicated the presence of the desired tags in nsP4.

Seven colonies (C1, C3, C7 from Plate 1 and C2, C3, C7, and C8 from Plate 2) possessed the correct size products and were therefore taken forward for further validation.

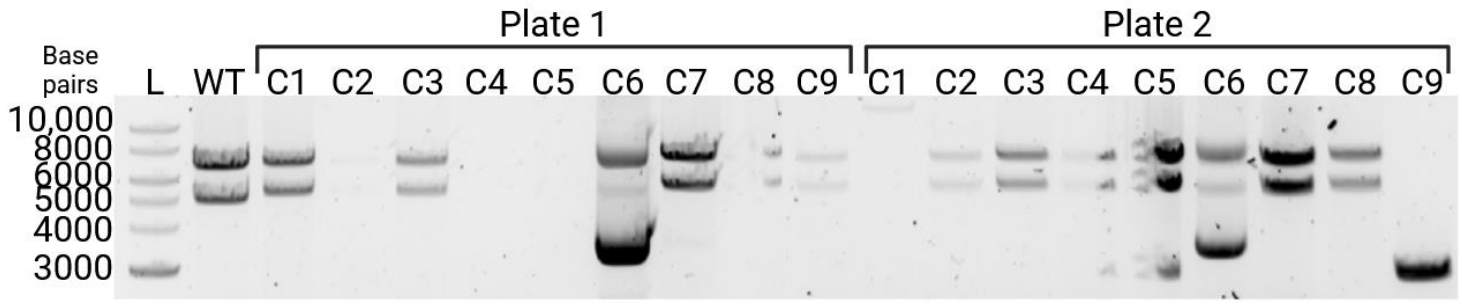


Figure 19: Diagnostic digest of colonies transformed with ligated plasmids. *Plasmids were digested with EcoRI and AvrII. nsP4-3XF SGR should produce two products: the backbone (7327 bp) and the insert containing the region encoding nsP4 tagged with Strep II and FLAG (5611 bp). WT SGR was used as a positive control for the digest, producing products of 7327 bp and 5397 bp.*

4.2.0 Validation of the novel system for detecting nsP4

4.2.1 Results

4.2.1.1 Sequencing of nsP4-3XF SGR confirmed the engineering of an SGR encoding nsP4 tagged with Strep II and FLAG

Colonies that were successfully digested with EcoRI and AvrII were sent for Sanger sequencing to confirm the desired plasmid had been engineered by insertion of the digested PCR product containing nsP4 tagged with Strep II and three tandem FLAG tags into the WT SGR backbone. All colonies sent for sequencing confirmed the

presence of the Strep II and FLAG tags in the N-terminus of the nsP4 sequence. From the eight plasmids sent for sequencing, C1 from Plate 1 (henceforth referred to as the nsP4-3XF SGR) was chosen to be taken forward for further validation, as it was the plasmid with the fewest sequencing errors.

Whole plasmid sequencing of the nsP4-3XF SGR was then performed. This again confirmed the successful engineering of the SGR expressing the FLAG tagged nsP4. When aligned with the sequence of the nsP4-3XF SGR, all 12,492 bases were aligned with 0 mismatches and 1 gap, where four adenine bases were missing from the original sequence of the 3' UTR.

4.2.1.2 Expression of nsP4-3XF decreases the level of SGR replication

To determine if the Strep II and FLAG tags ligated into nsP4 had an inhibitory effect on SGR replication, luciferase assays determining the level of luciferase expressed from cells transfected with either the WT SGR or nsP4-3XF SGR were performed and compared. As the WT SGR is known to replicate to the highest level within BHK-21 cells **(Roberts et al., 2017),** SGR replication was first compared in these cells **(Figure 20)**. These assays demonstrated that the insertion of the Strep II and FLAG tags into nsP4 had a detrimental effect on the level of SGR replication, as there was a significant decrease in the level of luciferase expressed in the nsP4-3XF SGR transfected cells compared to the WT SGR at every timepoint tested (4, 24 and 48 hpt). The level of SGR replication peaked at the same timepoint in both WT and nsP4-3XF transfected cells, with the highest level of luciferase signal detected at 24 hpt. Though there was a significant difference between the level of SGR replication, the nsP4-3XF SGR still replicated to a high level, with the level of luciferase detected from these lysates reaching approximately $9x10^5$ relative light units in BHK-21 cells.

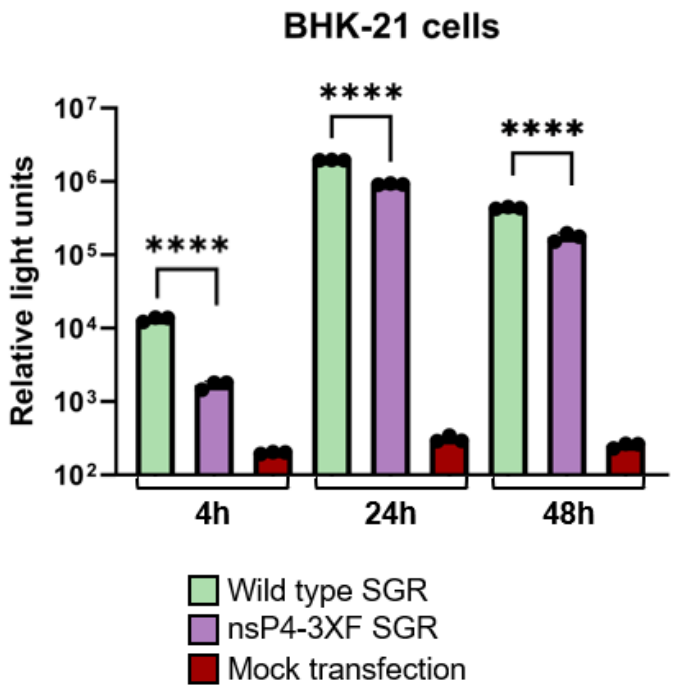


Figure 20: Insertion of Strep II and FLAG tags to nsP4 N-terminus significantly decreases level of luciferase expression in BHK-21 cells. *Comparison of firefly luciferase signal (in relative light units) in lysates of BHK-21 cells transfected with the WT SGR, nsP4-3XF SGR or DEPC water (mock transfection). Cells were assayed at three timepoints (4 hpt, 24 hpt and 48 hpt). Expression of nsP4-3XF significantly decreased the level of SGR replication at all timepoints. N=3 biological repeats. Assay gain was normalised to the lysate with the strongest reading. Error bars signify standard deviation from the mean and significance was measured by two-tailed T-test (****= P< 0.0001).*

Comparisons of SGR replication within biologically relevant model cells was then investigated, comparing the rate of replication within human RD and Huh7 cells and the *Ae. albopictus* C6/36 cells **(Figure 21)**. As in BHK-21 cells, SGR replication was significantly decreased in cells transfected with the nsP4-3XF SGR in comparison to those transfected with the WT SGR at 24 hpt and 48 hpt in both RD and Huh7 cells **(Figures 21A and 21B)**. The level of SGR replication peaked at the same time in all mammalian cells, with the highest expression of luciferase detected at 24 hpt

(Figures 20 and 21). However, unlike the BHK-21 cells transfected with nsP4-3XF SGR, there was no significant difference between the level of SGR replication between the nsP4-3XF SGR and WT SGR in RD and Huh7 cells at 4 hpt. Additionally, the difference in the level of SGR replication between the WT and nsP4-3XF SGRs was reduced in RD and Huh7 cells ($P<0.001$ and $P<0.01$, respectively) in comparison to the decrease observed in BHK-21 cells ($P<0.0001$) **(Figures 20 and 21).**

In contrast to the mammalian cell lines tested, there was no significant difference in the level of SGR replication between C6/36 cells transfected with the nsP4-3XF SGR or WT SGR at 4 hpt and 24 hpt **(Figure 21C)**. Additionally, the level of SGR replication did not peak at 24 hpt before decreasing at 48 hpt in C6/26 cells, as was observed in the mammalian cell lines **(Figures 20 and 21)**. Instead, the highest level of SGR replication was detected in C6/36 cells lysed at 48 hpt **(Figure 21C)**. This is likely due to the lytic effect that CHIKV has in mammalian cells, resulting in a decrease in CHIKV SGR after 24 hpt due to cell death and a shutdown of protein translation. However, as CHIKV persistently infects mosquito cells, the level of SGR replication continues to increase after 24 hpt.

While nsP4-3XF SGR did show a slightly reduced growth profile when compared to the WT SGR in the cells assayed, expression of nsP4-3XF does not prevent the nsP4-3XF SGR from replicating to high levels. Consequently, the nsP4-3XF SGR was taken forward for further validation to determine if it could potentially be used in the identification of novel interactions of nsP4 and host cell proteins.

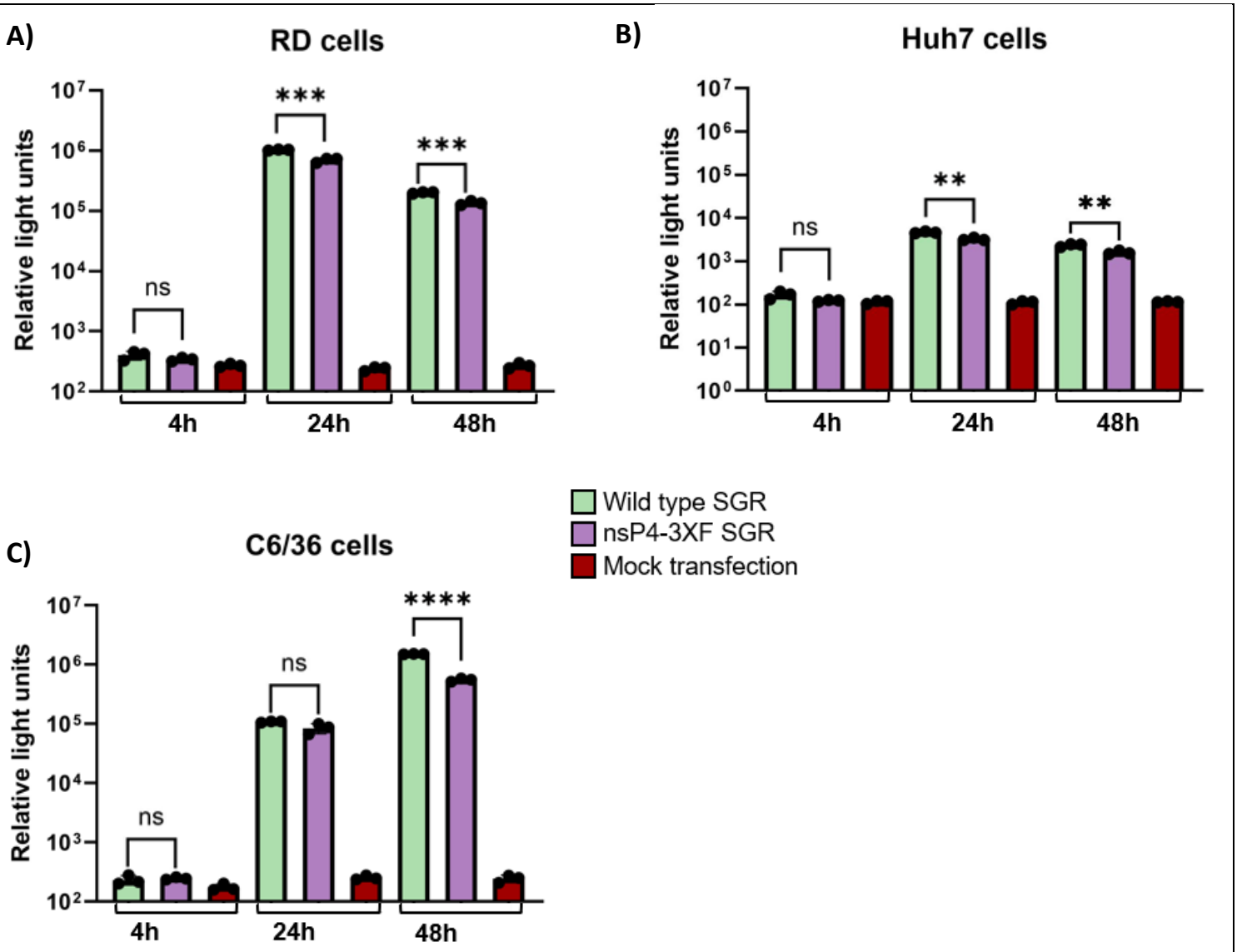


Figure 21: Insertion of Strep II and FLAG tags to nsP4 N-terminus significantly decreases level of replication in biologically relevant cells. *Comparison of firefly luciferase signal (in relative light units) in lysates of RD, Huh7 and C6/36 cells transfected with the WT SGR, nsP4-3XF SGR or DEPC water (mock transfection). Cells were assayed at three timepoints (4 hpt, 24 hpt and 48 hpt). Expression of nsP4-3XF significantly decreased the level of SGR replication at all timepoints. N=3 biological repeats. Assay gain was normalised to the lysate with the strongest reading. Error bars signify standard deviation from the mean and significance was measured by two-tailed T-test (ns = no significance, ** = P < 0.01, *** = P < 0.001, **** = P < 0.0001).*

4.2.1.3 nsP4-3XF was only detected by anti-FLAG antibodies from BHK-21 cell lysates

To determine if nsP4-3XF could be detected with anti-FLAG antibodies, the lysates of cells transfected with the nsP4-3XF SGR were analysed by western blotting. BHK-21, RD and Huh7 cells were transfected on a 6 well plate and harvested at 24 hpt, when SGR replication reached its peak in mammalian cells **(Figures 20 and 21)**. C6/36 cells were harvested at 48 hpt, when SGR replication reached its highest level in mosquito cells **(Figure 21C).**

WT SGR transfected cells and untransfected cells were used as negative controls for nsP4 detection by the anti-FLAG antibody. The lysate of BHK-21 cells transfected with human protein YAP encoding a FLAG tag were also included as a positive control for anti-FLAG antibody detection of FLAG tags. nsP4-3XF was only detected in BHK-21 cells **(Figure 22A).** From the BHK-21 sample transfected with the nsP4-3XF SGR, it was observed that the SGR may produce a second band, similar to the two bands observed when western blotting nsP4-FLAG lysate **(Figure 13-15)**. However, as the possible second band was close to the larger band it is difficult to tell if it is in fact a second smaller protein, or simply part of a single larger band. Additionally, these bands demonstrated similar levels of signal intensity when detected by the anti-FLAG antibody, which differs from nsP4-FLAG, as the second smaller band detected from the virus was less intense than the larger band **(Appendix 1)**.

While the anti-FLAG antibody detected YAP-FLAG on all four western blots, no signal for nsP4-3XF was detected in any biologically relevant cell line **(Figure 22B-D).** As the lysates were lysed with IP lysis buffer, for the potential use in Co-IP experiments, replication by luciferase assay could not be confirmed. The lack of

signal could therefore be a result of SGR replication not taking place in these cells, as nsP4 could be detected previously in RD and Huh7 cells transfected with the WT SGR **(Figure 13 and 15).** Due to time constraints, these experiments could not be repeated using PLB to confirm SGR replication of the nsP4-3XF SGR before testing for nsP4 detection using anti-FLAG antibodies, or tested with alternative antibodies, such as the Merits anti-nsP4 antibody to determine if nsP4 was present but not detected by the anti-FLAG antibody in RD, Huh7 and C6/36 cells. As nsP4-3XF was able to be detected in BHK-21 cells, they were taken forward to assess the potential use of anti-FLAG agarose beads to isolate nsP4-3XF and its interactors by Co-IP.

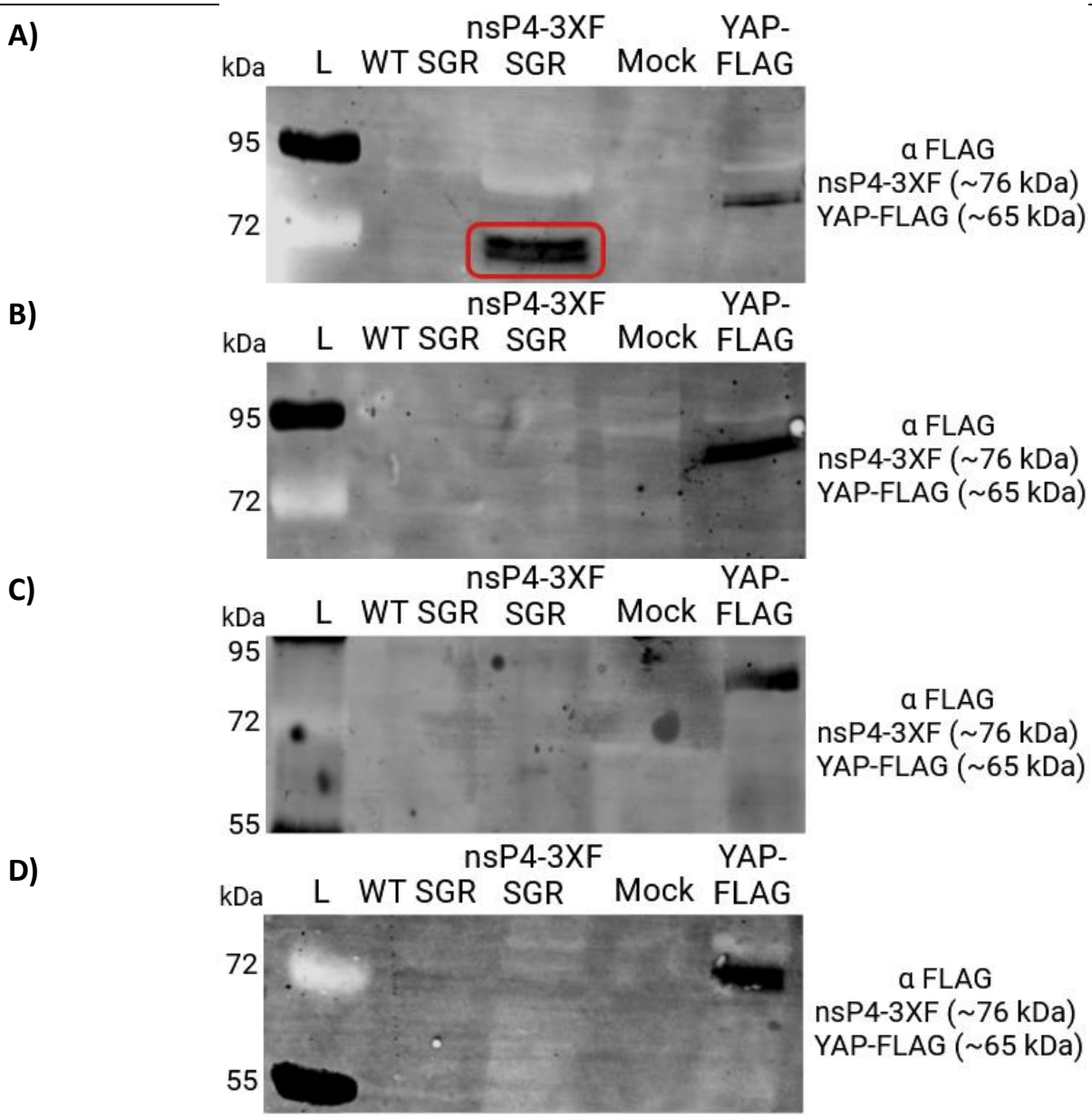


Figure 22: Western blots analysing the ability of anti-FLAG antibody to detect nsP4-3XF. *Western blots analysing the ability of FLAG agarose beads to perform Co-IPs to isolate nsP4-3XF (~76 kDa) and its interactors. BHK-21, RD, Huh7 and C6/36 cells were transfected with nsP4-3XF SGR and harvested using IP lysis buffer. Mammalian cells were harvested at 24 hpt and C6/36 cells at 48 hpt. WT SGR lysates were transfected with the WT SGR and mock lysates were transfected with DEPC water in place of SGR RNA as a negative controls for nsP4 detection. YAP-FLAG lysates were BHK-21 cells transfected with a plasmid expressing the human protein YAP with FLAG (~65 kDa) and used as a positive control for FLAG detection.* ***A)*** *BHK-21 cells, nsP4-3XF was detected* ***B)*** *RD cells, no nsP4 detected* ***C)*** *Huh7 cells, no nsP4 detected* ***D)*** *C6/36 cells, no nsP4 detected.*

4.2.1.4 Anti-FLAG magnetic agarose beads can be used for Co-IPs to isolate nsP4-3XF protein complexes from BHK-21 cells

As nsP4 was only detected by anti-FLAG antibodies in BHK-21 cells, BHK-21 cells transfected with nsP4-3XF were used to test the ability of anti-FLAG magnetic agarose beads to perform Co-IPs of nsP4-3XF. Cells were harvested at 24 hpt as this was the timepoint with the highest level of SGR replication demonstrated by luciferase assay **(Figure 20)**, and therefore the highest level of nsP4 expression. Cells were lysed with IP lysis buffer, as this buffer was designed to be used for immunoprecipitations and would therefore likely be used in Co-IP experiments to be sent for TMT-MS.

While only a faint band of approximate size to nsP4-3XF (~76 kDa) was detected in the input sample, a band of much higher intensity was detected at this size in the Co-IP sample **(Figure 23A)**. Additionally, no band for nsP4-3XF was observed in the flowthrough sample, suggesting that the majority of nsP4-3XF was bound to the anti-FLAG beads during this step. This clearly demonstrated the ability of anti-FLAG beads to bind to and isolate nsP4-3XF **(Figure 23A).** The higher band intensity observed in the Co-IP sample is likely due to the increased concentration of nsP4-3XF in the Co-IP compared to the input sample. The anti-FLAG antibody did not detect any signal from the mock lysates, demonstrating its specific nature in comparison to the Merits anti-nsP4 antibody, which resulted in multiple unspecific background bands being detected **(Figures 13-15)**.

Detection of β-actin indicates that there was a difference in concentration of β-actin between the input, flowthrough, and Co-IP samples of the transfected cell lysate **(Figure 23B)**. Previously, it has been reported that actin can be isolated alongside

the alphavirus RC during immunoprecipitation, but as the majority of β-actin was removed from the sample by the Co-IP sample stage, it is logical to assume that the presence of actin is indicative of non-specific binding rather than an interaction with nsP4 **(Barton et al., 1991).** This figure also indicates that less β-actin was present in the mock samples, and therefore there was a lower level of total protein in the mock samples than the samples from transfected cells **(Figure 23B)**. Due to the difference in the overall level of protein in the nsP4-3XF and mock samples, the results of the Co-IP cannot validly be compared. While this experiment would need to be repeated with a normalised level of input protein, it still provides valuable insight as an initial pilot study for the ability of anti-FLAG beads to isolate nsP4-3XF for future proteomic study of the nsP4 host interactome.

The current data does suggest that the nsP4 interactors HSP90 and TMEM45B can be immunoprecipitated alongside nsP4-3XF using anti-FLAG beads **(Figures 23C and 23D)**. A band of higher intensity the size of HSP90 was found in the nsP4-3XF SGR input in comparison to the mock input sample **(Figure 23C)**. Although this difference may be the result of the difference in protein concentration between the transfected and untransfected lysates, HSP90 is known to be upregulated in response to CHIKV infection **(Rathore et al., 2014).** Additionally, a band of higher intensity for HSP90 was observed in the Co-IP sample of the nsP4-3XF SGR expressing BHK-21 cells than the in the untransfected mock sample **(Figure 23C)**. Together, this suggests that anti-FLAG beads are capable of isolating nsP4-3XF alongside its host interactors by targeting the three tandem FLAG tags in the N-terminus of nsP4-3XF.

Additionally, a band of high intensity indicating the presence of TMEM45B was only observed in the input and Co-IP samples of the transfected cells, and with no bands

detected in untransfected flowthrough and Co-IP samples **(Figure 23D)**. As with the band observed for nsP4-3XF **(Figure 23A),** a fainter band for TMEM45B was detected in the flowthrough sample than in the input or Co-IP samples, suggesting that most of the protein was bound to the anti-FLAG beads at this stage **(Figure 23D)**. Additionally, as with nsP4-3XF, a band of higher intensity was observed in the nsP4-3XF SGR Co-IP sample than the input, indicating that TMEM45B was present in a higher concentration in the Co-IP sample than the input **(Figure 23D)**.

Due to time limitations these experiments could not be repeated with a normalised level of protein in the input samples, but these results do indicate a potential for nsP4-3XF and its interactors to be isolated by this method.

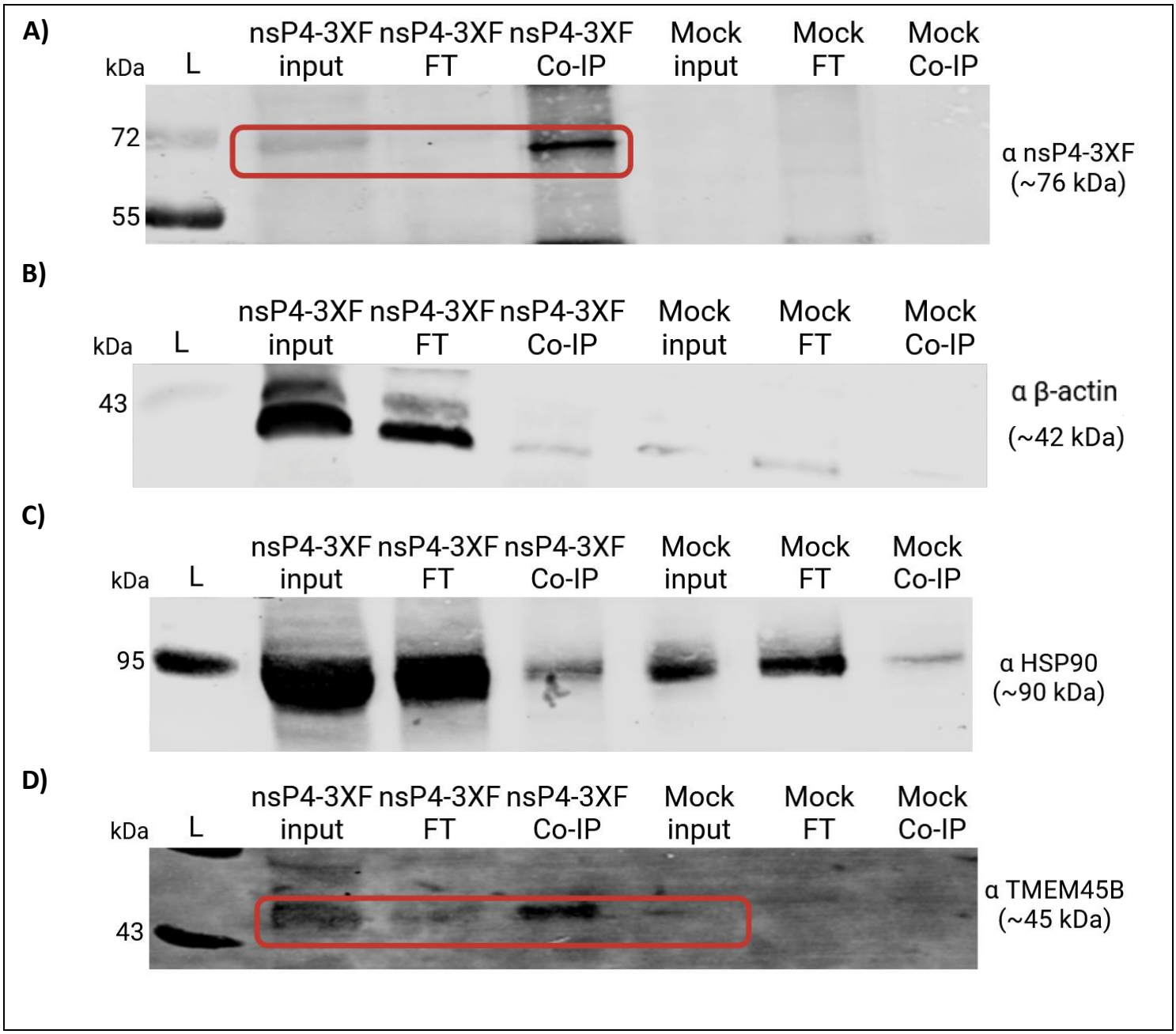

A)
kDa
L
nsP4-3XF input
nsP4-3XF FT
nsP4-3XF Co-IP
Mock input
Mock FT
Mock Co-IP
72
55
α nsP4-3XF
(~76 kDa)
B)
kDa
L
nsP4-3XF input
nsP4-3XF FT
nsP4-3XF Co-IP
Mock input
Mock FT
Mock Co-IP
43
α β-actin
(~42 kDa)
C)
kDa
L
nsP4-3XF input
nsP4-3XF FT
nsP4-3XF Co-IP
Mock input
Mock FT
Mock Co-IP
95
α HSP90
(~90 kDa)
D)
kDa
L
nsP4-3XF input
nsP4-3XF FT
nsP4-3XF Co-IP
Mock input
Mock FT
Mock Co-IP
43
α TMEM45B
(~45 kDa)

Figure 23: Western blots analysing anti-FLAG magnetic beads as a method for immunoprecipitating nsP4-3XF from cell lysates. *Western blots analysing the efficiency of anti-FLAG beads for isolating nsP4-3XF (~76 kDa) from BHK-21 cell lysates. Cells were transfected with the nsP4-3XF SGR and harvested at 24 hpt with IP lysis buffer. Mock lysates were transfected with DEPC water in place of SGR RNA were included as negative controls for immunoprecipitation of proteins using FLAG beads. Input (cell lysate), flow through (proteins from lysate unbound to beads) and Co-IP samples (proteins isolated from the cell lysate by Co-IP using the anti-FLAG beads) were tested for nsP4-3XF by FLAG antibody detection. Detection of proteins that should be isolated alongside nsP4 as interactors were also tested (HSP90 and TMEM45B, ~90 kDa and ~45 kDa respectively). β-actin (~42 kDa) was used as a loading control.* ***A)*** *Blots incubated in anti-FLAG antibodies to detect nsP4-3XF. A faint band in nsP4-3XF input sample and strong band in SGR Co-IP were detected at the size of the nsP4-3XF.* ***B)*** *β-actin levels in input, flowthrough and Co-IP samples in SGR and mock lysates* ***C)*** *HSP90 is present in nsP4-3XF and mock lysates but has stronger bands in nsP4-3XF input, flowthrough and Co-IP.* ***D)*** *TMEM45B is detected in the nsP4-3XF input, with a faint band in the flowthrough sample and strong band in the Co-IP.*

4.3 Discussion

Due to the poor efficiency of anti-nsP4 specific antibodies, a new system to study nsP4 during early replication events was engineered and initial validation steps were undertaken. Using the nsP4-FLAG virus, which encoded nsP4 with a Strep II and three FLAG tags inserted into the N-terminal domain, a version of the SGR encoding the tagged nsP4 was developed. FLAG tagged alphavirus nsP4s had previously been reported and were suggested as having a potential use for the discovery of nsP4 host protein interactions **(Tan et al., 2022b)**. The construction of this FLAG tagged version of CHIKV nsP4 circumvents the need to use the poorly efficient and often unspecific anti-nsP4 antibodies currently available. Development of the nsP4-3XF SGR will allow the study of nsP4 and its protein interactors during genome

replication and translation, in isolation to virion entry, assembly and release - which would not be possible using the virus nsP4-FLAG. Though the initial aims of this project were to identify interactions of nsP4 in mammalian cells, the nsP4-3XF SGR was also validated in mosquito C6/36 cells to ensure that the new SGR could efficiently replicate in these cells, for the potential use in investigating the nsP4 host interactome in mosquitoes, as nsP4 interactions also remain very poorly explored in mosquito vectors.

The nsP4-3XF SGR expressed significantly lower levels of firefly luciferase than the WT SGR at 24 and 48 hpt in all mammalian cell lines, and also at 4 hpt in BHK-21 cells **(Figures 20 and 21)**. This may suggest that genome replication and translation of CHIKV SGRs occurs earlier in BHK-21 cells than in RD or Huh7, resulting in the higher levels of luciferase expression observed in BHK-21 cells. The level of SGR replication peaked at the same time in all mammalian cells in both the WT and nsP4-3XF SGRs, with the luciferase expression decreasing between 24 and 48 hpt **(Figures 20 and 21)**. However, luciferase expression continued to increase in C6/36 cells between 24 and 48 hpt, and no significant difference in the level of SGR replication was demonstrated in C6/36 cells until 48 hpt **(Figure 21C)**. The decrease in mammalian cells can be attributed to CHIKV-induced host cell shutoff, which inhibits the translation of proteins by the host translation machinery **(Carey et al., 2019)**. Additionally, as the cytopathic response to CHIKV infection is lytic in mammalian cells **(Li et al., 2013)**, the decrease in luciferase expression may be attributed to the death of mammalian cells before the 48-hour timepoint. Conversely, CHIKV persistently infects mosquito cells, and has a much more moderate cytopathic effect during infection **(Li et al., 2013)**, allowing SGR replication to continue past 24 hpt.

Though expression of nsP4-3XF decreased the level of genome replication and translation in nsP4-3XF SGR transfected cells in comparison to the cells transfected with the WT SGR, the level of nsP4-3XF SGR replication was still high enough to enable the study of nsP4 using the nsP4-3XF SGR. For example, though SGR replication in nsP4-3XF SGR expressing BHK-21 and RD cells were significantly lower than their WT transfected counterparts, replication and translation still occurred in these cells to a much higher level than in the WT SGR in Huh7 cells **(Figures 20 and 21)**. Moreover, many other cell types including Vero cells, dermal fibroblast cells, and HeLa cells which were previously used as models to study CHIKV replication, have lower levels of firefly luciferase expression when transfected with the WT SGR **(Roberts et al., 2017).** This shows that although the insertion of the Strep II and FLAG tags into nsP4 decreases the level of SGR replication in the nsP4-3XF SGR, it is not so detrimental to the function of nsP4 that it no longer reaches high levels of replication. Therefore, the nsP4-3XF SGR still has the potential to be used for the study of nsP4 and its host interactome during early CHIKV genome events.

Despite the nsP4-3XF SGR being shown to replicate and be translated to high levels in RD and C6/36 cells **(Figure 21),** nsP4-3XF was not able to be detected by western blotting using an anti-FLAG antibody **(Figure 22)**. Previously, difficulties detecting CHIKV nsPs from mosquito cell lysates by western blotting have been reported, with **Roberts et al. (2017)** reporting that nsP3 was only able to be detected in C6/36 cells by IFA analysis**.** nsP4-3XF was also not detected in Huh7 cells, and though the CHIKV SGRs have been shown to be replicated to lower levels in Huh7 cells than the other cell types tested, WT nsP4 was identified in Huh7 cells previously during the lysis buffer experiments **(Figure 15)**. However, nsP4 was only weakly detected from RD and Huh7 cells lysed with IP lysis buffer, which may relate

to the lack of nsP4-3XF detected from these cells **(Figure 22).** The lack of time and amount of sample available prevented further western blot analysis of these lysates using the anti-nsP3 antibody (which would demonstrate SGR replication by nsP3 expression), the Merits anti-nsP4 antibody or repeats of this experiment with different lysis buffers, such as PLB to allow luciferase assay. Repeats of these experiments to determine if nsP4-3XF can be detected by FLAG antibodies by western blotting in should be performed, to determine if the nsP4-3XF SGR can be detected in biologically relevant cells, which can then be used in downstream applications.

As nsP4-3XF was only observed in BHK-21 cells using the anti-FLAG antibody **(Figure 22A)**, these cells were taken forward for use in the preliminary Co-IP experiment using anti-FLAG beads. As nsP4 was tagged with FLAG, these beads can be used to isolate nsP4 and any proteins complexed with the RdRp by binding to the FLAG tags **(Figure 23A)**. After optimisation of these Co-IPs, the results could then be sent for TMT-MS, allowing the identification of novel nsP4-host interactions. However, it would be beneficial to perform the Co-IPs to be sent for TMT-MS analysis in biologically relevant cells, such as RD or C6/36 cells, as there may be proteins identified in BHK-21 cells that interact with nsP4, that are not present or relevant during infection in human or mosquito cells. Additionally, the hamster proteome is not as well categorised as the human proteome, making analysis of the results from TMT-MS more difficult. However, as it is still a mammalian cell line, many of the proteins that interact with CHIKV proteins during infection will likely have homologues in human cells. Therefore, BHK-21 cells would still be a useful cell model for nsP4-host interactions if Co-IPs could not efficiently be performed in the other cell lines.

Time limitations of the project prevented the repeat of Co-IPs with normalised levels of protein being used in the input samples for nsP4-3XF Co-IP. Though it cannot validly be concluded that HSP90 and TMEM45B were specifically precipitated alongside nsP4-3XF, the initial experiment testing the ability of the FLAG beads to isolate nsP4 protein complexes is promising **(Figure 23)**. Once optimised, nsP4-3XF SGR could allow the isolation of nsP4 protein complexes by Co-IP, using a method that would more specifically target nsP4 by isolating the protein by the FLAG tag and evade the use of unspecific, inefficient nsP4 antibodies during Co-IP. This would be beneficial when evaluating TMT-MS data, as there would be fewer non-specific proteins in the samples, making it easier to identify proteins that interact with nsP4.

Chapter 5: Conclusions and Future Directions

5.1 nsP4 antibodies are not efficient enough for use in discovery of novel nsP4 interactions by Co-IP and TMT-MS

Though nsP4 antibodies have become more available in recent years, with commercial antibodies now available, such as the Genetex antibody employed in this study, it was concluded that the antibodies currently available are not sufficient for use in Co-IPs. The Genetex antibody was unable to detect nsP4 from transfected RD cell lysates and was therefore discarded **(Figure 11)**. Though the Merits lab antibody had previously been used in studies of nsP4 interactions **(Rathore et al., 2013, Rathore et al., 2014),** it was found that the antibody could not reliably detect nsP4 when expressed by the SGR **(Figures 12-15)**. A variety of lysis buffers were used during lysis of transfected cells as it was thought that the association of the RC with the host cell membrane may prevent nsP4 from being released during lysis. However, nsP4 was not consistently and conclusively detected from any one cell type or from any lysate lysed with a particular lysis buffer. The level of SGR replication did also not reliably predict the intensity of nsP4 observed using the Merits anti-nsP4 antibody, as demonstrated by the difference in signal intensity between BHK-21 cells and Huh7 cells lysed with PLB. Despite lower levels of SGR replication occurring in Huh7 cells than in BHK-21 cells, a more intense band for nsP4 was observed in PLB lysed Huh7 cells than the BHK-21 cells **(Figures 14A and 15A)**.

In addition to an optimal lysis buffer and cell line not being able to be determined, the Merits anti-nsP4 antibody itself was demonstrated to be inefficient, unspecific, and unreliable. Many unspecific bands were observed when blots were incubated with the Merits anti-nsP4 antibody, including at the size of nsP4 in untransfected cells

lysed with IP and RIPA buffer **(Figures 12-15)**, which resulted in the detection of nsP4 in transfected cells lysed with these buffers being inconclusive. Though originally planned for use in Co-IP and TMT-MS analysis of undiscovered nsP4 interactions, the unspecific and inefficient nature of the antibody resulted in this antibody being discarded for use in discovery of novel interactions of nsP4. Instead, a novel system that could reliably and specifically detect nsP4 from cell lysates was to be engineered to facilitate future studies of nsP4 interactions.

5.2 Engineering a novel system to study the nsP4 host interactome during early replication events

5.2.1 nsP4-3XF SGR was successfully engineered

Construction of a new SGR containing the tagged nsP4 from the nsP4-FLAG virus was successfully achieved **(Figure 19)**. The insert, amplified from the nsP4-FLAG virus, containing the tagged nsP4 and the added AvrII site **(Figure 18)**, was successfully cloned into the backbone of the WT SGR. The engineering of this SGR expressing the recombinant nsP4-3XF allows for the study of nsP4 and its protein interactors during replication and translation, without influence of the other viral lifecycle stages, by targeting the peptide tags ligated into nsP4. Successfully cloned plasmids were sent for sequencing, and the plasmid with the fewest sequencing errors was taken forward for validation.

5.2.2 Insertion of the Strep II and FLAG tags into nsP4 results in a decrease in the level of SGR replication and translation

The level of SGR replication was investigated in BHK-21 cells, in which the WT SGR replicates to the highest level, and in a range of biologically relevant cells (RD, Huh7 and C6/36 cells). Though the level of replication of nsP4-3XF SGR in the four cell

lines tested showed a reduced growth profile in comparison the WT SGR, expression of nsP4-3XF did not inhibit SGR replication to a detrimental level **(Figures 20 and 21)**. nsP4-3XF was still able to replicate to high levels, as demonstrated by the high levels of firefly luciferase expressed by nsP4-3XF SGR expressing cells. Indeed, the cell lines tested replicated nsP4-3XF SGR to higher levels than other biologically relevant cells (like dermal fibroblast cells) when transfected with the WT SGR. As the level of SGR replication was demonstrated to occur to a high level in the cells tested, nsP4-3XF SGR was taken forward for further validation, as it could be used for the study of nsP4 and host protein-protein interactions during early replication events.

5.2.3 nsP4-3XF can be used in the discovery of the nsP4 host interactome by western blotting and Co-IP

Despite the nsP4-3XF SGR replicating to high levels in the four cell lines tested, expression of nsP4-3XF was only observed from the cell lysates of BHK-21 cells and not in any of the biologically relevant cell lines tested **(Figure 22)**. Time constraints prevented the repetition of this experiment to validate that nsP4-3XF could be detected from biologically relevant cells by anti-FLAG antibodies, but as a positive result was observed from BHK-21 cells they were taken forward for use in a preliminary Co-IP experiment.

As no anti-nsP4 antibody could be optimised for use in Co-IP, the engineering of the SGR expressing nsP4-3XF was devised in order to perform Co-IPs by targeting the FLAG tags to circumvent the need for the inefficient anti-nsP4 antibodies. The Co-IPs were undertaken using anti-FLAG magnetic agarose beads, which bind to the three tandem FLAG tags in the N-terminus of nsP4-3XF. BHK-21 cells expressing

the nsP4-3XF SGR were lysed for use in the Co-IP. Co-IP using the anti-FLAG beads was also performed on untransfected cells as a negative control. The level of signal intensity for the bands representing nsP4-3XF increased between the initial transfected cell input and Co-IP samples, indicating that nsP4-3XF was isolated by the anti-FLAG beads resulting in an increase in nsP4-3XF concentration in the Co-IP sample **(Figure 23A)**. Additionally, interactors of nsP4 that have previously been reported, including HSP90 and TMEM45B, also appeared to have been isolated alongside nsP4-3XF **(Figures 23C and 23D)**. While the results of this Co-IP experiment cannot be validly compared to the negative control due to the difference in total protein level between the initial input samples **(Figure 23B)**, the Co-IP data is very promising and suggests that nsP4 protein complexes will be able to be isolated using the nsP4-3XF SGR and anti-FLAG beads.

5.3 Future directions

Confirming the ability of nsP4-3XF to be detected efficiently and reliably from transfected cell lysates would be the first and most important step when continuing the validation of the nsP4-3XF SGR. Though nsP4-3XF was detected from the lysate of BHK-21 cells by anti-FLAG antibodies **(Figure 22A),** it must be shown that detection of this nsP4-3XF is replicable and can be efficiently detected from BHK-21 cell lysates, unlike detection of WT nsP4 using the Merits anti-nsP4 antibody. Additionally, western blot analysis of nsP4-3XF expressing cells must be performed to confirm that nsP4-3XF can be detected from biologically relevant cell lines using the anti-FLAG antibody **(Figure 22)**. Lack of controls, such as the detection of nsP3 using the anti-nsP3 antibody or the lysis of the cells by PLB to perform luciferase assaying, resulted in SGR replication not being confirmed in the biologically relevant cells analysed. As such, it cannot be determined if the lack of nsP4-3XF signal

detection in these samples was the result of insufficient replication of nsP4-3XF SGR. Once nsP4-3XF can be efficiently and reliably detected from biologically relevant cell lysates, the nsP4-3XF SGR could be utilised to discover the nsP4 host interactome during early CHIKV replication events.

Preliminary data using the nsP4-3XF SGR to perform Co-IPs of nsP4 protein complexes is promising **(Figure 23)**. Though BHK-21 cells are not biologically relevant cells for identifying host proteins that interact with CHIKV for future drug development, as they originate from a hamster cell and not a human cell, they are a good candidate for optimisation of nsP4-3XF Co-IPs as nsP4-3XF SGR is replicated and transcribed to the higher level in BHK-21 cells than RD or Huh7 cells **(Figure 20 and 21)**. The increased level of nsP4-3XF expressed in nsP4-3XF SGR expressing BHK-21 cells could decrease the difficulty of detecting and isolating nsP4, which is expressed to lower levels than the other nsPs, making it easier to optimise Co-IPs. Once optimised, Co-IPs could then be performed in biologically relevant cells like RD cells which also replicates nsP4-3XF to a high level **(Figure 21)**. Three biological repeats of the Co-IPs would be completed in triplicate, with a negative control Co-IPs using untransfected cells also being performed. These samples would then be sent for TMT-MS, allowing the identification and quantification of novel nsP4 host protein interactions. Host proteins that were identified in high abundance in the transfected cell samples would be highlighted as potentially interacting with nsP4, and the interaction would be experimentally validated. Comparison to the untransfected samples would allow proteins that were non-specifically bound to the anti-FLAG beads during Co-IP to be discarded, as proteins that were found in high abundance of both transfected and untransfected samples would suggest that the protein was not bound in complex with nsP4 during the Co-IP of the transfected lysates.

As Co-IPs may not capture transient interactions, due to their unstable nature, crosslinking of nsP4 protein complexes bound to the FLAG beads could potentially be performed. Crosslinking would increase the likelihood of transient interactions being isolated during Co-IP, by preventing the disruption of the interaction, but would also increase the number of non-specific proteins in these samples. Performing western blots or staining of SDS-PAGE gels by Coomassie or Silver staining, to compare samples that have or have not been crosslinked and compare the level of background found in crosslinked Co-IP samples, could help determine if crosslinking of the Co-IP samples to be sent for TMT-MS should be performed.

Confirmation of the potential novel interactions identified by TMT-MS analysis would then be validated by methods such as IFA, inhibition of the potential interaction, and reverse Co-IPs that target the potential interactor instead of nsP4. IFAs can validate the potential interaction of two proteins by determining if the proteins colocalise together within the cell, with overlapping localisation indicating a potential interaction between the two proteins tested. Multiple timepoints can be analysed by IFA, which can provide additional insight to the lifecycle stage in which interaction takes place. Reverse Co-IPs using the potential interacting protein as bait for Co-IP, can also validate the interaction proposed from TMT-MS analysis. Immunoprecipitation of nsP4 alongside the potential interactor targeted by Co-IP, would indicate that nsP4 was complexed with the protein, reinforcing the likelihood of an interaction between the two proteins. Finally, inhibition of the potential interaction can provide insight into the effect of nsP4 interacting with that host cell protein. Inhibition of the interaction can be achieved by the use of siRNAs, shRNA or small molecule inhibitors that target the host factor. Inhibition of the expression of this host protein, or the ability of the interaction to occur between the host protein and nsP4, would give insight into

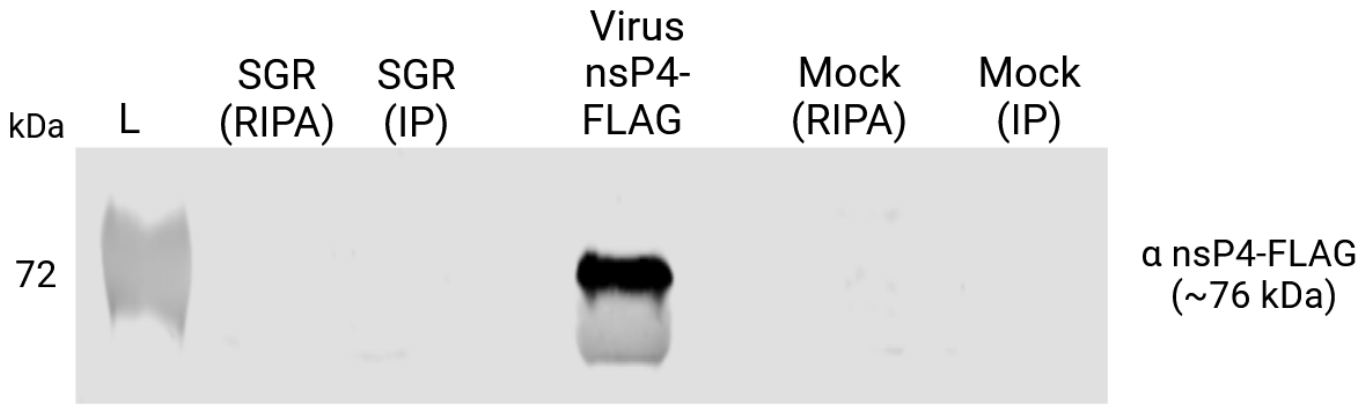
the interaction, and would determine if the effect of the interaction has a proviral or antiviral effect on CHIKV replication. The level of SGR replication in cells where the interactions is allowed to occur and cells in which the interaction has been inhibited can be compared by luciferase assay. If a significant increase in SGR replication was observed in the inhibited cells, this would indicate that the interaction between nsP4 and the host protein had an antiviral effect. On the other hand, a significant decrease in the level of SGR replication would demonstrate that the interaction had a proviral effect, with CHIKV exploiting the host factor to increase replication. Proteins that are critical for nsP4 function can be identified in this way, with an almost complete loss of SGR replication and luciferase expression in cells where the interaction had been inhibited demonstrating that CHIKV requires that interaction with the host factor to efficiently replicate.

Viral RdRps are considered prime targets for drug development due to their critical role in viral replication. As the limited information available for CHIKV nsP4's structure inhibits rational drug design, interactions of the protein with host factors can instead be targeted to inhibit the virus by inhibiting the action of nsP4. Host proteins that enhance viral replication, such as the interaction with eIF2α which prevents activation of the PERK pathway in the early stages of the CHIKV lifecycle **(Rathore et al., 2013)**, can be targeted to inhibit efficient replication of the virus during infection. Identification of an interaction of nsP4 that is critical for efficient CHIKV replication would be a prime target for drug development. Alternatively, identification of interactions that have an antiviral effect, such as TMEM45B with nsP4 **(Yan et al., 2022),** can offer insight into the development of drugs that inhibit nsP4 function. The critical nature of nsP4 for CHIKV replication, as the viral RdRp, means that targeting this protein specifically during infection could provide relief for CHIKV symptoms,

prevent the transmission of the disease to new hosts by inhibition of replication and decrease the high viraemia typically associated with CHIKV infections. Moreover, chronic CHIKV symptoms are linked to the severity of acute infections **(Wada et al., 2017),** and so the development of specific antivirals that decrease the severity of disease could prevent the development of the long-term pain associated with Chikungunya fever.

Discovery of novel interactions of nsP4 should focus on mammalian cells, as they have the most potential to be exploited for development of specific drug treatments against CHIKV infection in humans. However, interactions of mosquito cell proteins with nsP4 also provide valuable insight and may help in the development of specific CHIKV antivirals. The identification of interactions of CHIKV protein in mosquito cells can provide insight into how viruses replicate within insect cells, possibly helping to elucidate the differences between the moderate cell response observed in mosquito cells compared to the cytopathic effect observed during mammalian cell infection **(Li et al., 2013)**. Additionally, host proteins that interact with CHIKV proteins in mosquito cells may have homologues in mammalian cells, such as the mosquito protein Rasputin which aid CHIKV replication in mosquito cells in the same fashion as mammalian G3PB proteins in human cells **(Fros et al., 2015a).** The study of the nsP4 host proteome in mammalian and mosquito cells alike can therefore further the overall understanding of the mechanisms behind CHIKV early replication events, while providing possible targets for future development.

Appendix



Appendix Figure 1: Western blot analysing nsP4-FLAG detection by an anti-FLAG antibody. Western blot comparing nsP4-FLAG lysate with Huh7 samples lysed with RIPA and IP lysis buffers was incubated with an anti-FLAG antibody to detect nsP4-FLAG (~76 kDa).

References

ABDELNABI, R., NEYTS, J. & DELANG, L. 2015. Towards antivirals against chikungunya virus. *Antiviral Res,* 121**,** 59-68.

ABRAHAM, R., MUDALIAR, P., JALEEL, A., SRIKANTH, J. & SREEKUMAR, E. 2015. High throughput proteomic analysis and a comparative review identify the nuclear chaperone, Nucleophosmin among the common set of proteins modulated in Chikungunya virus infection. *J Proteomics,* 120**,** 126-41.

AHOLA, T. & MERITS, A. 2016. Functions of Chikungunya Virus Nonstructural Proteins. *Chikungunya Virus*.

BAKOVIC, A., BHALLA, N., ALEM, F., CAMPBELL, C., ZHOU, W. & NARAYANAN, A. 2021. Inhibitors of Venezuelan Equine Encephalitis Virus Identified Based on Host Interaction Partners of Viral Non-Structural Protein 3. *Viruses,* 13.

BARTON, D. J., SAWICKI, S. G. & SAWICKI, D. L. 1991. Solubilization and immunoprecipitation of alphavirus replication complexes. *J Virol,* 65**,** 1496-506.

BASAVAPPA, M. G., FERRETTI, M., DITTMAR, M., STOUTE, J., SULLIVAN, M. C., WHIG, K., SHEN, H., LIU, K. F., SCHULTZ, D. C., BEITING, D. P., LYNCH, K. W., HENAO-MEJIA, J. & CHERRY, S. 2022. The lncRNA ALPHA specifically targets chikungunya virus to control infection. *Mol Cell,* 82**,** 3729-3744 e10.

BREHIN, A. C., CASADEMONT, I., FRENKIEL, M. P., JULIER, C., SAKUNTABHAI, A. & DESPRES, P. 2009. The large form of human 2',5'-Oligoadenylate Synthetase (OAS3) exerts antiviral effect against Chikungunya virus. *Virology,* 384**,** 216-22.

BUSTOS CARRILLO, F., COLLADO, D., SANCHEZ, N., OJEDA, S., LOPEZ MERCADO, B., BURGER-CALDERON, R., GRESH, L., GORDON, A., BALMASEDA, A., KUAN, G. & HARRIS, E. 2019. Epidemiological Evidence for Lineage-Specific Differences in the Risk of Inapparent Chikungunya Virus Infection. *J Virol,* 93.

CAREY, B. D., BAKOVIC, A., CALLAHAN, V., NARAYANAN, A. & KEHN-HALL, K. 2019. New World alphavirus protein interactomes from a therapeutic perspective. *Antiviral Res,* 163**,** 125-139.

CARVALHO, C. A. M., CASSEB, S. M. M., GONCALVES, R. B., SILVA, E. V. P., GOMES, A. M. O. & VASCONCELOS, P. F. C. 2017. Bovine lactoferrin activity against Chikungunya and Zika viruses. *J Gen Virol,* 98**,** 1749-1754.

CHEN, M. W., TAN, Y. B., ZHENG, J., ZHAO, Y., LIM, B. T., CORNVIK, T., LESCAR, J., NG, L. F. P. & LUO, D. 2017. Chikungunya virus nsP4 RNA-dependent RNA polymerase core domain displays detergent-sensitive primer extension and terminal adenylyltransferase activities. *Antiviral Res,* 143**,** 38-47.

CHEN, R., MUKHOPADHYAY, S., MERITS, A., BOLLING, B., NASAR, F., COFFEY, L. L., POWERS, A., WEAVER, S. C. & ICTV REPORT, C. 2018. ICTV Virus Taxonomy Profile: Togaviridae. *J Gen Virol,* 99**,** 761-762.

CRISTEA, I. M., CARROLL, J. W., ROUT, M. P., RICE, C. M., CHAIT, B. T. & MACDONALD, M. R. 2006. Tracking and elucidating alphavirus-host protein interactions. *J Biol Chem,* 281**,** 30269-78.

DAS, I., BASANTRAY, I., MAMIDI, P., NAYAK, T. K., B, M. P., CHATTOPADHYAY, S. & CHATTOPADHYAY, S. 2014. Heat shock protein 90 positively regulates Chikungunya virus replication by stabilizing viral non-structural protein nsP2 during infection. *PLoS One,* 9**,** e100531.

DE GROOT, R. J., RUMENAPF, T., KUHN, R. J., STRAUSS, E. G. & STRAUSS, J. H. 1991. Sindbis virus RNA polymerase is degraded by the N-end rule pathway. *Proc Natl Acad Sci U S A,* 88**,** 8967-71.

DELANG, L., SEGURA GUERRERO, N., TAS, A., QUERAT, G., PASTORINO, B., FROEYEN, M., DALLMEIER, K., JOCHMANS, D., HERDEWIJN, P., BELLO, F., SNIJDER, E. J., DE LAMBALLERIE, X., MARTINA, B., NEYTS, J., VAN HEMERT, M. J. & LEYSSEN, P. 2014. Mutations in the chikungunya virus non-structural proteins cause resistance to favipiravir (T-705), a broad-spectrum antiviral. *J Antimicrob Chemother,* 69**,** 2770-84.

DOMINGUEZ, F., SHILIAEV, N., LUKASH, T., AGBACK, P., PALCHEVSKA, O., GOULD, J. R., MESHRAM, C. D., PREVELIGE, P. E., GREEN, T. J., AGBACK, T., FROLOVA, E. I. & FROLOV, I. 2021. NAP1L1 and NAP1L4 Binding to Hypervariable Domain of Chikungunya Virus nsP3 Protein Is Bivalent and Requires Phosphorylation. *J Virol,* 95**,** e0083621.

ECHAVARRIA-CONSUEGRA, L., DINESH KUMAR, N., VAN DER LAAN, M., MAUTHE, M., VAN DE POL, D., REGGIORI, F. & SMIT, J. M. 2023. Mitochondrial protein BNIP3 regulates Chikungunya virus replication in the early stages of infection. *PLoS Negl Trop Dis,* 17**,** e0010751.

FOX, J. M. & DIAMOND, M. S. 2016. Immune-Mediated Protection and Pathogenesis of Chikungunya Virus. *J Immunol,* 197**,** 4210-4218.

FOY, N. J., AKHRYMUK, M., AKHRYMUK, I., ATASHEVA, S., BOPDA-WAFFO, A., FROLOV, I. & FROLOVA, E. I. 2013. Hypervariable domains of nsP3 proteins of New World and Old World alphaviruses mediate formation of distinct, virus-specific protein complexes. *J Virol,* 87**,** 1997-2010.

FREIRE, M., BASSO, L. G. M., MENDES, L. F. S., MESQUITA, N., MOTTIN, M., FERNANDES, R. S., POLICASTRO, L. R., GODOY, A. S., SANTOS, I. A., RUIZ, U. E. A., CARUSO, I. P., SOUSA, B. K. P., JARDIM, A. C. G., ALMEIDA, F. C. L., GIL, L., ANDRADE, C. H. & OLIVA, G. 2022. Characterization of the RNA-dependent RNA polymerase from Chikungunya virus and discovery of a novel ligand as a potential drug candidate. *Sci Rep,* 12**,** 10601.

FREPPEL, W., SILVA, L. A., STAPLEFORD, K. A. & HERRERO, L. J. 2024. Pathogenicity and virulence of chikungunya virus. *Virulence,* 15**,** 2396484.

FROS, J. J., GEERTSEMA, C., ZOUACHE, K., BAGGEN, J., DOMERADZKA, N., VAN LEEUWEN, D. M., FLIPSE, J., VLAK, J. M., FAILLOUX, A. B. & PIJLMAN, G. P. 2015a. Mosquito Rasputin interacts with chikungunya virus nsP3 and determines the infection rate in Aedes albopictus. *Parasit Vectors,* 8**,** 464.

FROS, J. J., MAJOR, L. D., SCHOLTE, F. E. M., GARDNER, J., VAN HEMERT, M. J., SUHRBIER, A. & PIJLMAN, G. P. 2015b. Chikungunya virus non-structural protein 2-mediated host shut-off disables the unfolded protein response. *J Gen Virol,* 96**,** 580-589.

FROS, J. J. & PIJLMAN, G. P. 2016. Alphavirus Infection: Host Cell Shut-Off and Inhibition of Antiviral Responses. *Viruses,* 8.

FU, J. Y. L., CHUA, C. L., VYTHILINGAM, I., SULAIMAN, W. Y. W., WONG, H. V., CHAN, Y. F. & SAM, I. C. 2019. An amino acid change in nsP4 of chikungunya virus confers fitness advantage in human cell lines rather than in Aedes albopictus. *J Gen Virol,* 100**,** 1541-1553.

GARMASHOVA, N., GORCHAKOV, R., VOLKOVA, E., PAESSLER, S., FROLOVA, E. & FROLOV, I. 2007. The Old World and New World alphaviruses use different virus-specific proteins for induction of transcriptional shutoff. *J Virol,* 81**,** 2472-84.

GHILDIYAL, R., GUPTA, S., GABRANI, R., JOSHI, G., GUPTA, A., CHAUDHARY, V. K. & GUPTA, V. 2019. In silico study of chikungunya polymerase, a potential target for inhibitors. *Virusdisease,* 30**,** 394-402.

GOERTZ, G. P., LINGEMANN, M., GEERTSEMA, C., ABMA-HENKENS, M. H. C., VOGELS, C. B. F., KOENRAADT, C. J. M., VAN OERS, M. M. & PIJLMAN, G. P. 2018. Conserved motifs in the hypervariable domain of chikungunya virus nsP3 required for transmission by Aedes aegypti mosquitoes. *PLoS Negl Trop Dis,* 12**,** e0006958.

HAPUARACHCHI, H. C., BANDARA, K. B., SUMANADASA, S. D., HAPUGODA, M. D., LAI, Y. L., LEE, K. S., TAN, L. K., LIN, R. T., NG, L. F., BUCHT, G., ABEYEWICKREME, W. & NG, L. C. 2010. Re-emergence of Chikungunya virus in South-east Asia: virological evidence from Sri Lanka and Singapore. *J Gen Virol,* 91**,** 1067-76.

HUCKE, F. I. L., BESTEHORN-WILLMANN, M. & BUGERT, J. J. 2021. Prophylactic strategies to control chikungunya virus infection. *Virus Genes,* 57**,** 133-150.

HYDE, J. L., CHEN, R., TROBAUGH, D. W., DIAMOND, M. S., WEAVER, S. C., KLIMSTRA, W. B. & WILUSZ, J. 2015. The 5' and 3' ends of alphavirus RNAs – Non-coding is not non-functional. *Virus Research,* 206**,** 99-107.

ISSAC, T. H., TAN, E. L. & CHU, J. J. 2014. Proteomic profiling of chikungunya virus-infected human muscle cells: reveal the role of cytoskeleton network in CHIKV replication. *J Proteomics,* 108**,** 445-64.

JAIN, J., MATHUR, K., SHRINET, J., BHATNAGAR, R. K. & SUNIL, S. 2016. Analysis of coevolution in nonstructural proteins of chikungunya virus. *Virol J,* 13**,** 86.

JIA, H. & GONG, P. 2019. A Structure-Function Diversity Survey of the RNA-Dependent RNA Polymerases From the Positive-Strand RNA Viruses. *Front Microbiol,* 10**,** 1945.

JOUBERT, P. E., WERNEKE, S. W., DE LA CALLE, C., GUIVEL-BENHASSINE, F., GIODINI, A., PEDUTO, L., LEVINE, B., SCHWARTZ, O., LENSCHOW, D. J. & ALBERT, M. L. 2012. Chikungunya virus-induced autophagy delays caspase-dependent cell death. *J Exp Med,* 209**,** 1029-47.

JUNGFLEISCH, J., BOTTCHER, R., TALLO-PARRA, M., PEREZ-VILARO, G., MERITS, A., NOVOA, E. M. & DIEZ, J. 2022. CHIKV infection reprograms codon optimality to favor viral RNA translation by altering the tRNA epitranscriptome. *Nat Commun,* 13**,** 4725.

KAUR, P., LELLO, L. S., UTT, A., DUTTA, S. K., MERITS, A. & CHU, J. J. H. 2020. Bortezomib inhibits chikungunya virus replication by interfering with viral protein synthesis. *PLoS Negl Trop Dis,* 14**,** e0008336.

KHAN, A. H., MORITA, K., PARQUET, M. D. C., HASEBE, F., MATHENGE, E. G. M. & IGARASHI, A. 2002. Complete nucleotide sequence of chikungunya virus and evidence for an internal polyadenylation site. *J Gen Virol,* 83**,** 3075-3084.

KREJBICH-TROTOT, P., GAY, B., LI-PAT-YUEN, G., HOARAU, J. J., JAFFAR-BANDJEE, M. C., BRIANT, L., GASQUE, P. & DENIZOT, M. 2011. Chikungunya triggers an autophagic process which promotes viral replication. *Virol J,* 8**,** 432.

KRIL, V., AIQUI-REBOUL-PAVIET, O., BRIANT, L. & AMARA, A. 2021. New Insights into Chikungunya Virus Infection and Pathogenesis. *Annu Rev Virol,* 8**,** 327-347.

KUMAR, S., MAMIDI, P., KUMAR, A., BASANTRAY, I., BRAMHA, U., DIXIT, A., MAITI, P. K., SINGH, S., SURYAWANSHI, A. R., CHATTOPADHYAY, S. & CHATTOPADHYAY, S. 2015. Development of novel antibodies against non-structural proteins nsP1, nsP3 and nsP4 of chikungunya virus: potential use in basic research. *Arch Virol,* 160**,** 2749-61.

LEE, R. C. & CHU, J. J. 2015. Proteomics profiling of chikungunya-infected Aedes albopictus C6/36 cells reveal important mosquito cell factors in virus replication. *PLoS Negl Trop Dis,* 9**,** e0003544.

LEE, R. C., HAPUARACHCHI, H. C., CHEN, K. C., HUSSAIN, K. M., CHEN, H., LOW, S. L., NG, L. C., LIN, R., NG, M. M. & CHU, J. J. 2013. Mosquito cellular factors and functions in mediating the infectious entry of chikungunya virus. *PLoS Negl Trop Dis,* 7**,** e2050.

LELLO, L. S., BARTHOLOMEEUSEN, K., WANG, S., COPPENS, S., FRAGKOUDIS, R., ALPHEY, L., ARIEN, K. K., MERITS, A. & UTT, A. 2021. nsP4 Is a Major Determinant of Alphavirus Replicase Activity and Template Selectivity. *J Virol,* 95**,** e0035521.

LEUNG, J. Y., NG, M. M. & CHU, J. J. 2011. Replication of alphaviruses: a review on the entry process of alphaviruses into cells. *Adv Virol,* 2011**,** 249640.

LI, R., SUN, K., TUPLIN, A. & HARRIS, M. 2023. A structural and functional analysis of opal stop codon translational readthrough during Chikungunya virus replication. *J Gen Virol,* 104.

LI, Y. G., SIRIPANYAPHINYO, U., TUMKOSIT, U., NORANATE, N., A, A. N., TAO, R., KUROSU, T., IKUTA, K., TAKEDA, N. & ANANTAPREECHA, S. 2013. Chikungunya virus induces a more moderate cytopathic effect in mosquito cells than in mammalian cells. *Intervirology,* 56**,** 6-12.

LULLA, A., LULLA, V. & MERITS, A. 2012. Macromolecular assembly-driven processing of the 2/3 cleavage site in the alphavirus replicase polyprotein. *J Virol,* 86**,** 553-65.

MARTIN, M. F., BONAVENTURE, B., MCCRAY, N. E., PEERSEN, O. B., ROZEN-GAGNON, K. & STAPLEFORD, K. A. 2024. Distinct chikungunya virus polymerase palm subdomains contribute to viral protein accumulation and virion production. *PLoS Pathog,* 20**,** e1011972.

MARTINS, D. O. S., SOUZA, R. A. C., FREIRE, M., DE MORAES ROSO MESQUITA, N. C., SANTOS, I. A., DE OLIVEIRA, D. M., JUNIOR, N. N., DE PAIVA, R. E. F., HARRIS, M., OLIVEIRA, C. G., OLIVA, G. & JARDIM, A. C. G. 2023. Insights into the role of the cobalt(III)-thiosemicarbazone complex as a potential inhibitor of the Chikungunya virus nsP4. *J Biol Inorg Chem,* 28**,** 101-115.

MATKOVIC, R., BERNARD, E., FONTANEL, S., ELDIN, P., CHAZAL, N., HASSAN HERSI, D., MERITS, A., PELOPONESE, J. M., JR. & BRIANT, L. 2019. The Host DHX9 DExH-Box Helicase Is Recruited to Chikungunya Virus Replication Complexes for Optimal Genomic RNA Translation. *J Virol,* 93.

METIBEMU, D. S., ADEYINKA, O. S., FALODE, J., CROWN, O. & OGUNGBE, I. V. 2024. Inhibitors of the Structural and Nonstructural Proteins of Alphaviruses. *ACS Infect Dis,* 10**,** 2507-2524.

MÜLLER, M., JONES, N., TODD, E., KHALID, H., MERITS, A., MANKOURI, J. & TUPLIN, A. 2019. Replication of the Chikungunya virus genome requires cellular chloride channels. *Access Microbiology,* 1.

NGUYEN, T. H., NGUYEN, H. L., NGUYEN, T. Y., VU, S. N., TRAN, N. D., LE, T. N., VIEN, Q. M., BUI, T. C., LE, H. T., KUTCHER, S., HURST, T. P., DUONG, T. T., JEFFERY, J. A.,

DARBRO, J. M., KAY, B. H., ITURBE-ORMAETXE, I., POPOVICI, J., MONTGOMERY, B. L., TURLEY, A. P., ZIGTERMAN, F., COOK, H., COOK, P. E., JOHNSON, P. H., RYAN, P. A., PATON, C. J., RITCHIE, S. A., SIMMONS, C. P., O'NEILL, S. L. & HOFFMANN, A. A. 2015. Field evaluation of the establishment potential of wMelPop Wolbachia in Australia and Vietnam for dengue control. *Parasit Vectors,* 8**,** 563.

PAINGANKAR, M. S. & ARANKALLE, V. A. 2014. Identification of chikungunya virus interacting proteins in mammalian cells. *J Biosci,* 39**,** 389-99.

PAREEK, A., KUMAR, R., MUDGAL, R., NEETU, N., SHARMA, M., KUMAR, P. & TOMAR, S. 2022. Alphavirus antivirals targeting RNA-dependent RNA polymerase domain of nsP4 divulged using surface plasmon resonance. *FEBS J,* 289**,** 4901-4924.

PEINADO, R. D. S., EBERLE, R. J., ARNI, R. K. & CORONADO, M. A. 2022. A Review of Omics Studies on Arboviruses: Alphavirus, Orthobunyavirus and Phlebovirus. *Viruses,* 14.

PIETILA, M. K., HELLSTROM, K. & AHOLA, T. 2017. Alphavirus polymerase and RNA replication. *Virus Res,* 234**,** 44-57.

POHJALA, L., UTT, A., VARJAK, M., LULLA, A., MERITS, A., AHOLA, T. & TAMMELA, P. 2011. Inhibitors of alphavirus entry and replication identified with a stable Chikungunya replicon cell line and virus-based assays. *PLoS One,* 6**,** e28923.

PROSSER, O., STONEHOUSE, N. J. & TUPLIN, A. 2023. Inhibition of Chikungunya virus genome replication by targeting essential RNA structures within the virus genome. *Antiviral Res,* 211**,** 105523.

RANA, J., RAJASEKHARAN, S., GULATI, S., DUDHA, N., GUPTA, A., CHAUDHARY, V. K. & GUPTA, S. 2014. Network mapping among the functional domains of Chikungunya virus nonstructural proteins. *Proteins,* 82**,** 2403-11.

RATHORE, A. P., HAYSTEAD, T., DAS, P. K., MERITS, A., NG, M. L. & VASUDEVAN, S. G. 2014. Chikungunya virus nsP3 & nsP4 interacts with HSP-90 to promote virus replication: HSP-90 inhibitors reduce CHIKV infection and inflammation in vivo. *Antiviral Res,* 103**,** 7-16.

RATHORE, A. P., NG, M. L. & VASUDEVAN, S. G. 2013. Differential unfolded protein response during Chikungunya and Sindbis virus infection: CHIKV nsP4 suppresses eIF2alpha phosphorylation. *Virol J,* 10**,** 36.

RAUSALU, K., UTT, A., QUIRIN, T., VARGHESE, F. S., ZUSINAITE, E., DAS, P. K., AHOLA, T. & MERITS, A. 2016. Chikungunya virus infectivity, RNA replication and non-structural polyprotein processing depend on the nsP2 protease's active site cysteine residue. *Sci Rep,* 6**,** 37124.

REIS, E. V. S., DAMAS, B. M., MENDONCA, D. C., ABRAHAO, J. S. & BONJARDIM, C. A. 2022. In-Depth Characterization of the Chikungunya Virus Replication Cycle. *J Virol,* 96**,** e0173221.

REYES-GASTELLOU, A., JIMENEZ-ALBERTO, A., CASTELAN-VEGA, J. A., APARICIO-OZORES, G. & RIBAS-APARICIO, R. M. 2021. Chikungunya nsP4 homology modeling reveals a common motif with Zika and Dengue RNA polymerases as a potential therapeutic target. *J Mol Model,* 27**,** 247.

REZZA, G. & WEAVER, S. C. 2019. Chikungunya as a paradigm for emerging viral diseases: Evaluating disease impact and hurdles to vaccine development. *PLoS Negl Trop Dis,* 13**,** e0006919.

ROBERTS, G. C., ZOTHNER, C., REMENYI, R., MERITS, A., STONEHOUSE, N. J. & HARRIS, M. 2017. Evaluation of a range of mammalian and mosquito cell lines for use in Chikungunya virus research. *Sci Rep,* 7**,** 14641.

RUBACH, J. K., WASIK, B. R., RUPP, J. C., KUHN, R. J., HARDY, R. W. & SMITH, J. L. 2009. Characterization of purified Sindbis virus nsP4 RNA-dependent RNA polymerase activity in vitro. *Virology,* 384**,** 201-8.

RUPP, J. C., JUNDT, N. & HARDY, R. W. 2011. Requirement for the amino-terminal domain of sindbis virus nsP4 during virus infection. *J Virol,* 85**,** 3449-60.

SANCHEZ-ALDANA-SANCHEZ, G. A., LIEDO, P., BOND, J. G. & DOR, A. 2023. Release of sterile Aedes aegypti mosquitoes: chilling effect on mass-reared males survival and escape ability and on irradiated males sexual competitiveness. *Sci Rep,* 13**,** 3797.

SAWICKI, D., BARKHIMER, D. B., SAWICKI, S. G., RICE, C. M. & SCHLESINGER, S. 1990. Temperature sensitive shut-off of alphavirus minus strand RNA synthesis maps to a nonstructural protein, nsP4. *Virology,* 174**,** 43-52.

SCHNEIDER, M., NARCISO-ABRAHAM, M., HADL, S., MCMAHON, R., TOEPFER, S., FUCHS, U., HOCHREITER, R., BITZER, A., KOSULIN, K., LARCHER-SENN, J., MADER, R., DUBISCHAR, K., ZOIHSL, O., JARAMILLO, J. C., EDER-LINGELBACH, S., BUERGER, V. & WRESSNIGG, N. 2023. Safety and immunogenicity of a single-shot live-attenuated chikungunya vaccine: a double-blind, multicentre, randomised, placebo-controlled, phase 3 trial. *Lancet,* 401**,** 2138-2147.

SCHOLTE, F. E., TAS, A., ALBULESCU, I. C., ZUSINAITE, E., MERITS, A., SNIJDER, E. J. & VAN HEMERT, M. J. 2015. Stress granule components G3BP1 and G3BP2 play a proviral role early in Chikungunya virus replication. *J Virol,* 89**,** 4457-69.

SCHUFFENECKER, I., ITEMAN, I., MICHAULT, A., MURRI, S., FRANGEUL, L., VANEY, M. C., LAVENIR, R., PARDIGON, N., REYNES, J. M., PETTINELLI, F., BISCORNET, L., DIANCOURT, L., MICHEL, S., DUQUERROY, S., GUIGON, G., FRENKIEL, M. P., BREHIN, A. C., CUBITO, N., DESPRES, P., KUNST, F., REY, F. A., ZELLER, H. & BRISSE, S. 2006. Genome microevolution of chikungunya viruses causing the Indian Ocean outbreak. *PLoS Med,* 3**,** e263.

SCHWARTZ, O. & ALBERT, M. L. 2010. Biology and pathogenesis of chikungunya virus. *Nat Rev Microbiol,* 8**,** 491-500.

SHIN, G., YOST, S. A., MILLER, M. T., ELROD, E. J., GRAKOUI, A. & MARCOTRIGIANO, J. 2012. Structural and functional insights into alphavirus polyprotein processing and pathogenesis. *Proc Natl Acad Sci U S A,* 109**,** 16534-9.

SHIRAKO, Y. & STRAUSS, J. H. 1998. Requirement for an aromatic amino acid or histidine at the N terminus of Sindbis virus RNA polymerase. *J Virol,* 72**,** 2310-5.

SIMON, F., JAVELLE, E., OLIVER, M., LEPARC-GOFFART, I. & MARIMOUTOU, C. 2011. Chikungunya virus infection. *Curr Infect Dis Rep,* 13**,** 218-28.

SINGH, A., KUMAR, A., UVERSKY, V. N. & GIRI, R. 2018. Understanding the interactability of chikungunya virus proteins via molecular recognition feature analysis. *RSC Adv,* 8**,** 27293-27303.

SINGH, S. K. & UNNI, S. K. 2011. Chikungunya virus: host pathogen interaction. *Rev Med Virol,* 21**,** 78-88.

SKIDMORE, A. M. & BRADFUTE, S. B. 2023. The life cycle of the alphaviruses: From an antiviral perspective. *Antiviral Res,* 209**,** 105476.

SOUSA, I. P., JR., CARVALHO, C. A. M. & GOMES, A. M. O. 2020. Current Understanding of the Role of Cholesterol in the Life Cycle of Alphaviruses. *Viruses,* 13.

SPUUL, P., BALISTRERI, G., HELLSTROM, K., GOLUBTSOV, A. V., JOKITALO, E. & AHOLA, T. 2011. Assembly of alphavirus replication complexes from RNA and protein components in a novel trans-replication system in mammalian cells. *J Virol,* 85**,** 4739-51.

SUN, K., APPADOO, F., LIU, Y., MÜLLER, M., MACFARLANE, C., HARRIS, M. & TUPLIN, A. 2024. A novel interaction between the 5' untranslated region of the Chikungunya virus genome and Musashi RNA binding protein is essential for efficient virus genome replication. *Nucleic Acids Research,* 52**,** 10654-10667.

TAN, Y. B., CHMIELEWSKI, D., LAW, M. C. Y., ZHANG, K., HE, Y., CHEN, M., JIN, J. & LUO, D. 2022a. Molecular architecture of the Chikungunya virus replication complex. *Sci Adv,* 8**,** eadd2536.

TAN, Y. B., LELLO, L. S., LIU, X., LAW, Y. S., KANG, C., LESCAR, J., ZHENG, J., MERITS, A. & LUO, D. 2022b. Crystal structures of alphavirus nonstructural protein 4 (nsP4) reveal an intrinsically dynamic RNA-dependent RNA polymerase fold. *Nucleic Acids Res,* 50**,** 1000-1016.

TEPPOR, M., ZUSINAITE, E. & MERITS, A. 2021. Phosphorylation Sites in the Hypervariable Domain in Chikungunya Virus nsP3 Are Crucial for Viral Replication. *J Virol,* 95.

THOKA, B., JAIMIPAK, T., ONNOME, S., YOKSAN, S., UBOL, S. & PULMANAUSAHAKUL, R. 2018. The synergistic effect of nsP2-L(618), nsP3-R(117), and E2-K(187) on the large plaque phenotype of chikungunya virus. *Virus Genes,* 54**,** 48-56.

TIOZZO, G., DE ROO, A. M., GURGEL DO AMARAL, G. S., HOFSTRA, H., VONDELING, G. T. & POSTMA, M. J. 2025. Assessing chikungunya's economic burden and impact on health-related quality of life: Two systematic literature reviews. *PLoS Negl Trop Dis,* 19**,** e0012990.

TOMAR, S., HARDY, R. W., SMITH, J. L. & KUHN, R. J. 2006. Catalytic core of alphavirus nonstructural protein nsP4 possesses terminal adenylyltransferase activity. *J Virol,* 80**,** 9962-9.

TORTOSA, P., COURTIOL, A., MOUTAILLER, S., FAILLOUX, A. B. & WEILL, M. 2008. Chikungunya-Wolbachia interplay in Aedes albopictus. *Insect Mol Biol,* 17**,** 677-84.

TSETSARKIN, K. A. & WEAVER, S. C. 2011. Sequential adaptive mutations enhance efficient vector switching by Chikungunya virus and its epidemic emergence. *PLoS Pathog,* 7**,** e1002412.

UTT, A., RAUSALU, K., JAKOBSON, M., MANNIK, A., ALPHEY, L., FRAGKOUDIS, R. & MERITS, A. 2019. Design and Use of Chikungunya Virus Replication Templates Utilizing Mammalian and Mosquito RNA Polymerase I-Mediated Transcription. *J Virol,* 93.

VAN HUIZEN, E. & MCINERNEY, G. M. 2020. Activation of the PI3K-AKT Pathway by Old World Alphaviruses. *Cells,* 9.

VERMA, S., NEWAR, J., MANOSWINI, M., DHAL, A. K. & GHATAK, A. 2023. Interactions of the nsP3 proteins of CHIKV with the human host protein NAP1 plays a significant role in viral pathogenesis – An in silico study. *Human Gene,* 36.

WADA, Y., ORBA, Y., SASAKI, M., KOBAYASHI, S., CARR, M. J., NOBORI, H., SATO, A., HALL, W. W. & SAWA, H. 2017. Discovery of a novel antiviral agent targeting the nonstructural protein 4 (nsP4) of chikungunya virus. *Virology,* 505**,** 102-112.

WEBER, W. C., STREBLOW, D. N. & COFFEY, L. L. 2024. Chikungunya Virus Vaccines: A Review of IXCHIQ and PXVX0317 from Pre-Clinical Evaluation to Licensure. *BioDrugs,* 38**,** 727-742.

WEISS, B., NITSCHKO, H., GHATTAS, I., WRIGHT, R. & SCHLESINGER, S. 1989. Evidence for specificity in the encapsidation of Sindbis virus RNAs. *J Virol,* 63**,** 5310-8.

WHO. 2020. *Chikungunya* [Online]. Available: https://www.who.int/health-topics/chikungunya/#tab=tab_1 [Accessed 26/11/24 2024].

YAN, F., YANG, W., WANG, X. & GAO, G. 2022. TMEM45B Interacts with Sindbis Virus Nsp1 and Nsp4 and Inhibits Viral Replication. *J Virol,* 96**,** e0091922.

YAN, Y., ZHANG, F., ZOU, M., CHEN, H., XU, J., LU, S. & LIU, H. 2024. Identification of RACK1 as a novel regulator of non-structural protein 4 of chikungunya virus. *Acta Biochim Biophys Sin (Shanghai),* 56**,** 1425-1436.

ZELLER, H., VAN BORTEL, W. & SUDRE, B. 2016. Chikungunya: Its History in Africa and Asia and Its Spread to New Regions in 2013-2014. *J Infect Dis,* 214**,** S436-S440.

ZHANG, L. & ELIAS, J. E. 2017. Relative Protein Quantification Using Tandem Mass Tag Mass Spectrometry. *Methods Mol Biol,* 1550**,** 185-198.

ZIMMERMAN, O., HOLMES, A. C., KAFAI, N. M., ADAMS, L. J. & DIAMOND, M. S. 2023. Entry receptors - the gateway to alphavirus infection. *J Clin Invest,* 133.