The Expression of Chikungunya Virus Envelope 2 Glycoprotein Variants in *Nicotiana benthamiana* for the Development of a Diagnostic Reagent

By

Jason Christopher Delville Naude

Dissertation presented for the degree of Masters of Science

Department of Molecular and Cell Biology

Faculty of Science

University of Cape Town

February 2020

Supervisor: Dr Ann Meyers

Co-Supervisor: Dr Aleyo Chabeda

University of Cape Town

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

University of Cape Town

Acknowledgements

Firstly, to my supervisors Dr Ann Meyers and Dr Aleyo Chabeda, thank you not only for your support throughout my masters, but your undying patience and invaluable guidance. For your diligence and hands-on involvement with me as your student, I am eternally grateful.

To Associate Prof Inga Hitzeroth, thank you for filling in as acting-supervisor during the absence of my supervisors above. For your guidance and assistance on a daily basis, despite not being my supervisor, I, again, am eternally grateful.

And, to Prof Ed Rybicki, thank you for the opportunity to work in the Biopharming Research Unit. Your wealth of knowledge and willingness to educate has helped fashion me into the scientist I am today. Thank you, too, for the allowing me the opportunities to represent the Biopharming Research Unit in my many academic ventures outside of the lab.

To the other students, scientific officers and post-docs of the Biopharming Research Unit, old and new, thank you for three years of companionship, guidance, support and providing me with a healthy space to learn, fail, communicate, prosper, and be me. The relationships I have developed in the last three years have been heart-warming and invaluable to my day-to-day sanity.

To my funders, the Poliomyelitis Research Foundation and the University of Cape Town, thank you for providing me with the resources to pursue this degree. Your awarding me with financial support has provided me with great incentive and validation throughout my academic journey.

To my loving and supportive parents, Gregg Naude and Robyn Schmidt, thank you for your patience, your guidance and for the morals and values you have instilled within me throughout my life. No words can explain the love and appreciation I have for you for moulding me into the person I am today.

And last, but most certainly not least, to my dearest friends: your companionship has been the catalyst to my achievements and accolades. Again, words cannot describe the deep-felt love and appreciation I have for your years of support, love, laughs, advice, and the strength you have provided me. For someone as plagued by the unfortunate ailments of mental health as me, your friendships have rendered my demons near negligible.

"We can choose to be afraid of it, to stand there trembling not moving, assuming the worst that can happen or we can step forward into the unknown and assume it will be brilliant."

- Shonda Rhimes

Declaration

Name:	Jason Christopher Delville Naude
Student Number:	NDXJAS006
Course:	MCB5005W

I acknowledge that plagiarism is both wrong and unethical. Plagiarism is defined as using another's work and passing it of as those it is one's own.

I have used the Harvard convention for citation and referencing. Each contribution to, and quotation in, this Dissertation from the work(s) of other people has been attributed and has been cited and referenced.

This Dissertation presented for the degree of Master of Science in the Department of Molecular and Cell Biology is my own work.

I have not allowed, and will not allow, anyone to copy my work with the intention of passing it off as his or her own work.

Signature: Signed by candidate

Date: 10/02/2020

Abstract

Chikungunya fever is a non-fatal but highly debilitating disease that affects primates, birds and humans. The causative agent is the chikungunya virus (CHIKV), an arbovirus of the *Alphavirus* genus. CHIKV is responsible for the largest epidemic recorded for an *Alphavirus,* infecting an estimated 1.4 to 6 million patients worldwide. Furthermore, it has been recognised by the United States army as a potential biological weapon used for bioterrorism owing to the potential for infection via aerosol. CHIKV is primarily transmitted by infected *Aedes aegypti* mosquitos and is currently distributed in Africa, parts of Asia and South, Central and North America. As a result of the virus genetically adapting to infect the *Aedes albopictus* mosquito, its recent and rapid spread to non-endemic regions has occasioned increasing anxiety as well. Infection in humans presents as a sudden onset of fever, rash and severe arthralgia that persists for years. At present, there is no fast and effective diagnostic test to distinguish CHIKV from other similar viruses. This is a problem because viral infection displays the same symptoms as that of dengue, Zika, Ebola and yellow fevers while prognosis, patient care, and persistent symptoms of these viruses are very different. Usually, during the development of a diagnostic reagent, Biosafety Level 3 (BSL3) containment is required for purifying antigens from live viruses. These lab diagnostic tests are expensive to perform and, in regions facing a CHIKV epidemic, are inefficient due to their long waiting periods. This results in patients going undiagnosed or misdiagnosed and/or falling outside the window of prophylactic treatment. As such, a cheap and rapid diagnostic reagent to detect the presence of CHIKV antibodies would be most advantageous.

In this study, two recombinant variants of the CHIKV E2 glycoprotein were expressed in *Nicotiana benthamiana* plants to assess their viability for use in a diagnostic reagent for CHIKV infection. Two versions of a tobacco sp. codon-optimised, 6xhis-tagged CHIKV E2 envelope glycoprotein gene were synthesised and cloned into the plant expression vector, pTRAkc-ERH. The E2 glycoprotein is a desirable protein candidate used for a diagnostic reagent as it is a major target for neutralizing antibody production against CHIKV during early infection. One variant contained a ~52 kDa full length E2 glycoprotein (CHIKV E2-HIS) while the other contained a ~49 kDa truncated E2 glycoprotein lacking its transmembrane domain (CHIKV E2ΔTM-HIS). Following this, an expression time trial was performed whereby the recombinant proteins were expressed in *N. benthamiana* plants via *Agrobacterium*-mediated small-scale

syringe-infiltration at different optical densities, OD_{600} = 1.0 and 0.5. To improve expression, both genes were co-infiltrated and co-expressed with a human chaperone proteins calreticulin (CRT) or calnexin (CNX), or a plant silencing-suppressor protein NSs. Expression of the recombinant protein variants alone showed low to undetectable levels of expression in plant leaves across 7 days post infiltration (dpi) for both ODs tested. CHIKV E2ΔTM-HIS yielded the highest levels of all combinations tested at an OD_{600} = 1.0 when co-expressed with CRT and harvested at 3 dpi. These parameters were used for subsequent scaling up and production of E2 using vacuum infiltration.

Attempts at purifying CHIKV E2ΔTM-HIS proteins using Ni-NTA affinity chromatography and further investigation into the exposure of the 6xHis-tag on the native conformation of CHIKV E2ΔTM-HIS indicated that the 6xHis-tag was insufficiently exposed on E2 and thus inaccessible to facilitate purification by Ni-NTA affinity chromatography. Further attempts at purifying recombinant CHIKV E2ΔTM-HIS proteins by pH purification were also unsuccessful as large amounts of plant-protein contaminants present in all samples prevented adequate separation from CHIKV E2ΔTM-HIS. A different approach utilising ammonium sulphate precipitation facilitated separation of recombinant CHIKV E2ΔTM-HIS from some of the contaminating plant proteins in the 30 - 60% ammonium sulphate fraction; however, large amounts of recombinant CRT were co-purified with E2 in this fraction.

Although expression of a candidate diagnostic reagent in plants for detecting CHIKV antibodies in the form of E2 glycoprotein was achieved, further research needs to be done to optimise a purification strategy for CHIKV E2ΔTM-HIS proteins.

Abbreviations

6xhis (-HIS)	Six-histidine amino-acid tag
BCIP/NBT	5-bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium
bp	Base pairs
CaMV	Cauliflower mosaic virus
CHIKV	Chikungunya virus
CNX	Calnexin
CPMV	Chickpea mosaic virus
CRT	Calreticulin
DNA	Deoxyribose nucleic acid
Dpi	Days post-infiltration
DSP	Downstream processing
E1	Envelope 1 glycoprotein
E2	Envelope 2 glycoprotein
E2ΔTM	Envelope 2 glycoprotein lacking transmembrane domain
EDTA	Ethylenediaminetetra-acetic acid
ELISA	Enzyme-linked immunosorbent assay
ER	Endoplasmic reticulum
g	Gram(s)
hr(s)	Hour(s)
kb	Kilobase(s)
kDa	Kilodalton(s)
L	Litre(s)
LB	Luria-Bertani
LPH	Leader protein heavy
M	Molar
mA	Milliampere
mg	Milligram(s)
min	Minute(s)
mL	Millilitre(s)
mM	Millimolar
MES	2-morpholinoethanesulfonic acid

ng	Nanogram(s)
nm	Nanometre(s)
O/N	Overnight
OD	Optical density
PBS	Phosphate buffer saline
PCR	Polymerase chain reaction
RNA	Ribonucleic acid
rpm	Revolutions per minute
RT	Room temperature
s	Second(s)
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
T-DNA	Transfer-DNA
TBE	Tris-borate-EDTA buffer
TM	Transmembrane
TMD	Transmembrane domain
Tris	Tris(hydroxymethyl)amonimethane
TSP	Total soluble protein
UV	Ultraviolet
V	Volt(s)
VLP(s)	Virus-like particle(s)
µF	Microfarad
µg	Microgram(s)
µL	Microlitre(s)
µm	Micrometre(s)
%	Percentage
^{o}C	Degrees Celsius
Ω	Resistance

Table of Contents

Chapter 1: Literature Review 11
1.1 Introduction 11
1.2 The Chikungunya Virus 12
1.2.1 Virus Classification 12
1.2.2 Virus Genome and Structure 13
1.2.3 Virus Vectors and Transmission 14
1.2.4 Geographical Distribution 15
1.2.5 Clinical Symptoms 16
1.3 Envelope Glycoproteins 17
1.4 CHIKV Therapeutics, Diagnostics and Vaccines 19
1.4.1 Current Therapeutics 19
1.4.2 Current Diagnostic Assays 20
1.4.3 Current Vaccine Candidates 22
1.5 Plant-Based Protein Expression and Downstream Processing (DSP) 23
1.5.1 Expression in Plants 23
1.5.2 Downstream Processing of Recombinant Proteins 25
1.5.3 Transient Expression in Tobacco 25
1.6 Plant-Based Research for CHIKV 26
1.7 Project Aims and Objectives 27
Chapter 2: Expression of Recombinant CHIKV E2 Glycoprotein Variants in Plants 29
2.1 Introduction 29
2.2 Materials and Methods 33
2.2.1 Gene Design and Synthesis 33
2.2.2 Cloning and Plant Expression Vectors 34
2.2.3.1 General Cloning 35
2.2.3.2 Isolation of gene variants 37
2.3.3.3 pTRAkc-ERH cloning strategy 38
2.2.3.4 Original pEAQ-HT Cloning Strategy 39
2.2.3.5 New pEAQ-HT Cloning Strategy 39
2.2.3 Small-Scale Agroinfiltration Expression Time-Trials 41
2.2.4 Small-Scale Protein Extraction and Quantification of recombinant proteins 43

2.2.5 Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE) 43
2.2.6 Western Blotting 44
2.2.7 Large-Scale Vacuum Agroinfiltration 44
2.2.8 Large-Scale Protein Extraction 45
2.2.9 Stability Studies 45
2.3 Results 46
2.3.1 Confirmation of Plant Expression Constructs 46
2.3.2 Small-Scale Protein-Expression Time Trials 53
2.3.3 Large-Scale Vacuum Infiltration 55
2.3.4 Stability of CHIKV E2ΔTM-HIS 57
2.4 Discussion 58
Chapter 3: Purification of CHIKV E2ΔTM-HIS Glycoprotein Variant 62
3.1 Introduction 62
3.2 Materials and Methods 64
3.2.1 Nickel-Resin Purification 64
3.2.2 Denaturing vs. Non-Denaturing Polyacrylamide Gels 65
3.2.3 pH Precipitation 66
3.3.4 Ammonium Sulphate Precipitation 66
3.3.5 Assessment of CHIKV E2ΔTM-HIS Purification Yields 67
3.3 Results 67
3.3.1 Nickel-Resin Purification 67
3.3.2 Comparison of CHIKV E2ΔTM-HIS Resolved by Denaturing and Non-Denaturing Gel 70
3.3.3 pH Precipitation 71
3.3.4 Ammonium Sulphate Precipitation 72
3.4 Discussion 73
Chapter 4: Conclusion 77
References 81

Chapter 1: Literature Review

1.1 Introduction

Mosquito-borne diseases are responsible for major global health threats. These vectors are responsible for the spread of viruses that cause dengue fever, Japanese encephalitis, Rift Valley fever, yellow fever, West Nile and Zika, to name a few (Rudolph et al., 2014). One other such arbovirus is the chikungunya virus (CHIKV), the causative agent of chikungunya fever - a non-fatal but highly debilitating disease that affects primates, birds, and humans. The virus was first isolated during an outbreak in 1952 in Tanzania where its name originated from, in the native Makonde language, to mean "that which bends upward", describing the contorted posture of patients suffering from the severe characteristic joint pain (Schwartz and Albert, 2010). Now, CHIKV is responsible for the largest epidemic recorded for an *Alphavirus,* infecting an estimated 1.4 to 6 million patients worldwide (Metz et al., 2013). As an *Alphavirus*, CHIKV is a spherical, enveloped, positive-sense single-stranded RNA virus (Weaver and Lecuit, 2015). Furthermore, CHIKV is transmitted by infected female *Aedes* mosquitos and is currently distributed in Africa, parts of Asia and South, Central and North America. Previously, the virus was primarily spread by *Aedes aegypti* mosquitoes (Moulay et al., 2012) found in tropical and subtropical regions; however, evidence of a genetic adaptation of the virus to better infect *Aedes albopictus* mosquitos has been credited for the spread of the virus to previously non-endemic temperate regions (Tsetsarkin et al., 2007), in addition to the rapid and aggressive re-emergence of the virus in countries like India (Mavalankar et al., 2007). Moreover, CHIKV has been recognised by the United States army as a potential biological weapon used in bioterrorism owing to the potential for infection via aerosol (Hasan et al., 2018). Typical infection in humans presents as a sudden onset of fever, rash and, its most distinct feature, severe arthralgia that can persist for years (Akahata et al., 2010).

While antiviral treatments against CHIKV infection have been studied in animal models and *in vitro*, the effectiveness and safety of these treatments have yet to be determined in a clinical setting (Tharmarajah et al., 2017). At present, there is no fast and effective diagnostic assay to distinguish CHIKV from other similar co-circulating viruses. This is a problem because CHIKV infection displays the same symptoms as that of dengue (Pialoux

et al., 2007), Zika (Petersen et al., 2016), Ebola (Rewar and Mirdha, 2014) and yellow (Rudolph et al., 2014) fevers while prognosis, patient care, and persistent symptoms of these viruses are very different (Prat et al., 2014). While research has gone into optimising existing CHIKV diagnostics, they focus primarily on improving the sensitivity rather than the time aspect of these assays (Edwards et al., 2007). This is a problem in countries like India where approximately 1.4 million cases have been reported since the outbreak in 2005 and there exists only two laboratories that can serologically diagnose CHIKV (Mavalankar et al., 2007).

In addition, there is no licensed vaccine available to aid in the prevention and spread of CHIKV infection (Pialoux et al., 2007). While considerable efforts have gone into the development of a candidate vaccine against CHIKV – going back as far as the 1960s – these vaccine candidates have only been tested in mouse models and most remain in the preclinical phase (Burt et al., 2017).

As such, a rapid and cost-effective diagnostic reagent is desperately needed to detect the presence of CHIKV in regions that are challenged with identically-symptomatic viruses like dengue fever virus (DENV). Additionally, there is a need for a cost-effective, efficacious candidate vaccine against CHIKV that will permit the control of CHIKV infection and spread.

1.2 The Chikungunya Virus

1.2.1 Virus Classification

CHIKV is an *Alphavirus* belonging to the *Togaviridae* family. The *Alphavirus* genus comprises 29 distinct species – those that are threatening to humans typically cause severe diseases like encephalitis to milder febrile illnesses accompanied by polyarthritis, the characteristic week- to month-long syndrome of CHIKV infection. Furthermore, CHIKV belongs to the same genus as Sindbis, Semliki Forest, Ross River, Venezuelan equine encephalitis, Barmah Forest, o'nyong-nyong, and Mayaro viruses (Sun et al., 2019). CHIKV has been listed by the National Institute of Allergy and Infectious Diseases (NIAID) as a Category C Priority Pathogen due to its emerging threat (Thiboutot et al., 2010).

1.2.2 Virus Genome and Structure

CHIKV contains a plus-sense, single-stranded RNA genome approximately 12 kb in length (Figure 1.1). Two-thirds of the genome encode four non-structural proteins (nsPs) on the 5' end, translated from genomic RNA (Galán-Huerta et al., 2015). These nsPs are translated into a polyprotein cassette and subsequently processed into mature nsPs involved in viral infection, replication, gene regulation, defence, packaging and protease activity (Burt et al., 2017). The remaining third of the genome encodes three structural proteins on the 3' end, translated from subgenomic RNA (Galán-Huerta et al., 2015). Like the nsPs, the structural proteins are translated into a polyprotein cassette and subsequently processed into mature structural proteins that are essential for virion infection and assembly. These proteins include the capsid protein (C) and two surface envelope glycoproteins, E1 and E2. The capsid protein is responsible for nucleocapsid formation as well as entry to and exit from the nuclei of infected cells (Burt et al., 2017). The E1 envelope glycoprotein is responsible for viral entry and the E2 envelope glycoprotein is responsible for initiating cell attachment by binding to an unknown receptor on the surface of the host cell (Weaver and Lecuit, 2015).

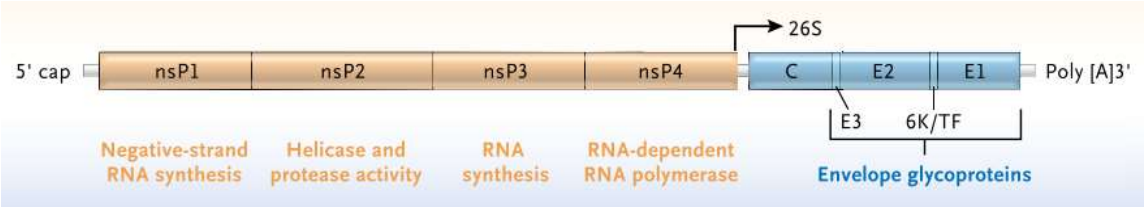

Figure 1.1: Organisation of the CHIKV genome. Illustration of the organisation and function of the four non-structural and three structural proteins of the CHIKV genome. 26S, 26S ribosomal RNA. Reproduced with permission from (Weaver and Lecuit, 2015), Copyright Massachusetts Medical Society.

CHIKV is an enveloped virus of ~70 nm with an icosahedral nucleocapsid structure of ~40 nm encapsidating a single molecule of viral RNA. The viral lipid-bilayer envelope is host-derived and displays the viral envelope glycoproteins which form 80 trimeric spikes, each composed of three E1-E2 heterodimers, on the surface of the virion (Metz et al., 2013) (Figure 1.2).

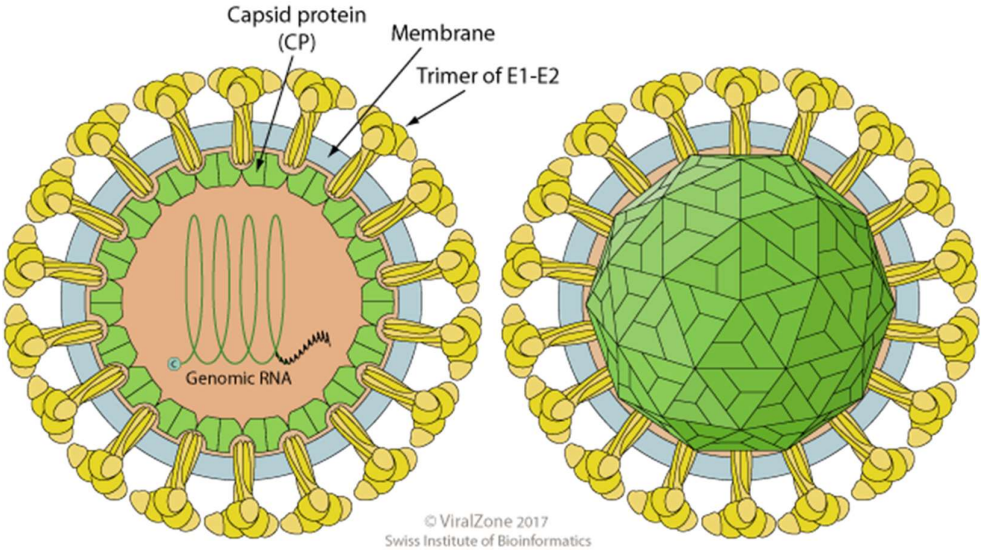

Figure 1.2: The typical structure of an *Alphavirus* virion. The organisation of the structural proteins (E1, E2 and CP) and the lipid-bilayer membrane are shown to form heteromeric spikes on the surface of the virion. The encapsidated RNA genome is approximately 11.6 kb. Image taken from ViralZone: www.expasy.org/viralzone *SIB Swiss Institute of Bioinformatics*.

1.2.3 Virus Vectors and Transmission

CHIKV is transmitted by infected female *Aedes* mosquitos. Usually, the virus circulates in a sylvatic cycle between mammalian reservoir or non-human primate hosts; however, it has been identified that during urban epidemics CHIKV is transmitted via a human-mosquito-human transmission cycle .

The primary carrier of CHIKV is *A. aegypti*, also called the yellow fever mosquito owing to its contribution to the spread of yellow fever virus. With the increase in geographical distribution of the mosquito from rural to urban settings, the incidence of CHIKV infection has increased. Additionally, increases in travel, insecticide- and drug-resistance development, and human activity have contributed to the re-emergence of CHIKV. Furthermore, human activity has facilitated the carriage of mosquito eggs over long distances. It is thus people, not mosquitoes, that are responsible for the rapid spread of *A. aegypti* as a hatched mosquito does not have a perimeter greater than 200 m (Moulay et al., 2012).

The spread of the virus to previously non-endemic regions has provided evidence that CHIKV has genetically adapted to better infect a new mosquito host, *Aedes albopictus*, an inhabitant of temperate regions with a wider geographical distribution (Tsetsarkin et

al., 2007). *A. albopictus* is an aggressive, diurnal and silent vector, rendering prevention strategies like bed nets ineffective. It is a resilient species that can thrive in both rural and urban environments. Consequently, it has a remarkable capacity to adapt to human activity and urbanisation allowing its survival in almost all human-habituated environments in dual zoophilic and anthropophilic preferences. Coupled with its relatively longer-lived lifespan (4 -8 weeks) and larger flight radius of 400-600m, it has superceded *A. aegypti* in many aspects (Knudsen, 1995).

1.2.4 Geographical Distribution

CHIKV was first isolated in Tanzania in 1952 and is believed to have originated in Africa. Before 2000, CHIKV infection was reported in Angola, Benin, Burundi, Cameroon, the Central African Republic, Democratic Republic of the Congo, Gabon, Guinea, Kenya, Liberia, Madagascar, Malawi, Nigeria, Uganda, Senegal, Sierra Leone, southern Africa, and Sudan. Its subsequent spread to Asia is believed to have occurred via shipping where the earliest reports of disease by CHIKV were confirmed in the Philippines in 1954. Subsequent outbreaks have been documented in Asia in Bangladesh, Bhutan, Cambodia, China, India, Indonesia, Laos, Malaysia, Maldives, Burma/Myanmar, Pakistan, Saudi Arabia, Singapore, Sri Lanka, Taiwan, Thailand, Timor, Vietnam, and Yemen (Burt et al., 2017). After disappearing for more than a decade, the virus re-emerged through dissemination by *A. albopictus* causing outbreaks of higher severity in the same areas and areas previously not affected by CHIKV. For example, several outbreaks were reported after the re-emergence of CHIKV in new geographical areas such as Europe (UK, Belgium, Germany, Czech Republic, Norway, Spain and France), Hong Kong, Canada, Taiwan, Sri Lanka, and the USA (Powers and Logue, 2007). Figure 1.3 represents the spatial distribution of the respective *Aedes* mosquito species and CHIKV strains.

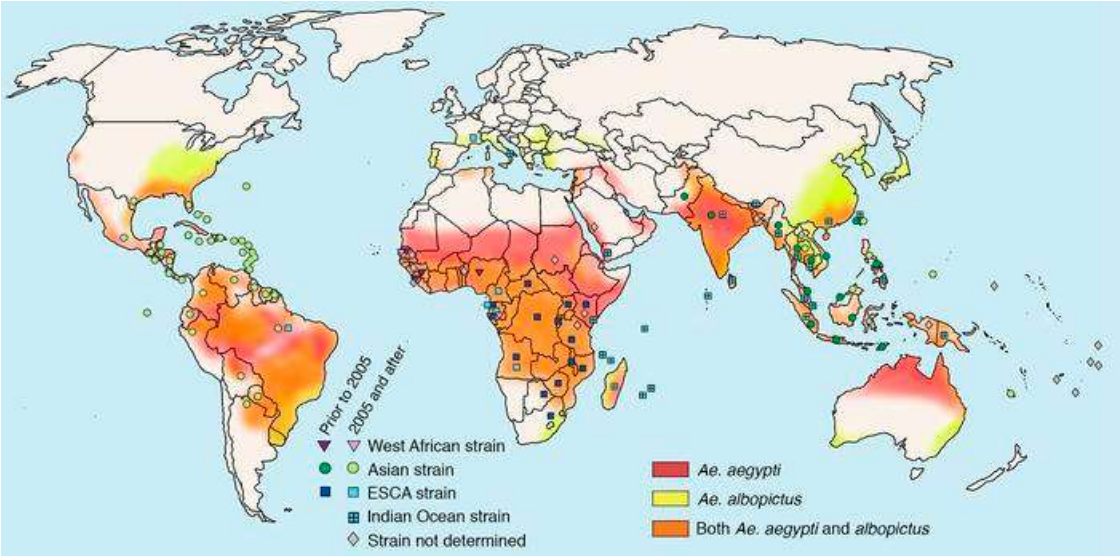

Figure 1.3: The geographical distribution of the *Aedes* mosquito and respective CHIKV strains. Coloured symbols indicate reported autochthonous CHIKV cases with a colour overlay indicating the geographical range of the two primary vectors responsible for mosquito-human-mosquito transmission of CHIKV. ESCA = East/South/Central African. Reproduced with permission from (Silva and Dermody, 2017), Copyright American Society for Clinical Investigation.

Most dramatically, in 2005, the Indian government reported more than 1.4 million cases of CHIKV infection in 12 states in the country. This was an underestimate because no surveillance system existed for CHIKV until 2006 (Mavalankar et al., 2007). It is estimated that nearly 40% of the world's population is at risk of CHIKV infection as outbreaks continue to occur globally (Wimalasiri-Yapa et al., 2019).

1.2.5 Clinical Symptoms

Infection by CHIKV is accompanied by a range of signs and symptoms of variable intensities in patients, with only 5 – 15% of infections being asymptomatic (Appassakij et al., 2013). Acute infection is most commonly characterised by the sudden onset of severe joint and muscle pain, a maculopapular rash ranging from a mild, localised rash to one more widespread (some covering 90% of the body), a fever higher than 39 ^{o}C (Miner et al., 2015), headache, lethargy (Weaver and Lecuit, 2015) and back pain. These symptoms usually resolve within two weeks; however, arthralgia may persist for months or even years (Galán-Huerta et al., 2015). While death from CHIKV infection is rare (1-in-1000 patients), cases of encephalitis, encephalopathy, myocarditis, hepatitis and multi-organ

failure have been reported in severe infections. Additionally, infected neonates, people above the age of 65 years and immunosuppressed patients are at an increased risk of severe infection (Burt et al., 2017).

Throughout history, these febrile symptoms were often mistaken for infection with o'nyong-nyong virus, Sindbis virus or DENV, the fastest spreading mosquito-borne virus in the world. However, the incapacitating arthralgia is what discerns CHIKV from DENV specifically (Carey, 1971). While CHIKV and DENV are two of the most difficult infections to distinguish between, symptoms of CHIKV infection are more abruptly onset. Rash and arthralgia are also more frequent in CHIKV than DENV infection (Pialoux et al., 2007).

1.3 Envelope Glycoproteins

The CHIKV envelope glycoproteins, E1 and E2, form part of the structural polyprotein cassette translated from a 26S subgenomic mRNA and are further proteolytically processed into the viral envelope glycoproteins. The 80 trimeric spikes formed by multiple E1-E2 heterodimers are essential for the budding of new virus particles. Cell entry via pH-dependent endocytosis is facilitated by E1, while recognition and attachment of unknown host receptors is facilitated by E2. The envelope polyprotein cassette (E3-E2-6K-E1) (Figure 1.1) is processed in the endoplasmic reticulum and the resulting E1 and membrane-bound precursor E2 (E3-E2) form heterodimers that trimerise to form the trimeric spikes. Finally, precursor E2 undergoes furin-dependent cleavage, releasing E3 to form mature trimeric spikes which are displayed on the surface of the virion (Metz et al., 2013).

CHIKV E1, a class II fusion glycoprotein, mediates viral membrane fusion with a cell membrane, as a trimer, via a low pH-dependent insertion and subsequent refolding to form a hairpin-like structure. The recently emerged CHIKV mutant (bearing an E1-A226V substitution), has been found to have allowed the virus to adapt to the broader-distributed host, *A. albopictus*. It is estimated that the E1 glycoprotein may have a lower chance of exposure to a host's immune system due to the envelope protein being more buried within the structure of the virions (Kuo, Chen et al. 2012). Therefore, CHIKV E1 is not an ideal candidate for use in the development of a diagnostic assay.

CHIKV E2, a type I transmembrane glycoprotein, mediates receptor binding during the *Alphavirus* lifecycle (Cho et al., 2008). A study by Kuo et al. (2012) showed that the E2 glycoprotein in alphaviruses both facilitates CHIKV E1 glycoprotein folding and regulates E1 fusogenicity. Like E1, there is evidence to suggest that the E2 glycoprotein underwent a mutation event that facilitated the adaptation of CHIKV to the more successful *A. albopictus* vector. It is suggested that the mutation enhanced CHIKV infectivity in the mosquito midgut possibly via the fusion step in the endosome (Matusali et al., 2019). In addition, the CHIKV E2 glycoprotein is a major target for neutralizing antibodies against CHIKV (Fox et al., 2015) during early infection owing to an epitope located on the N-terminus of the glycoprotein. As a result, the E2 glycoprotein is an appropriate protein candidate for use in the development of a diagnostic assay for early detection of CHIKV infection in sufferers (Kam et al., 2012).

CHIKV E1 and E2 are predicted to be glycosylated at N141, and N263 and N273, respectively (Metz et al., 2011). Asparagine (N)-linked glycosylation is an important co- and post-translational modification that these envelope proteins must undergo when entering the secretory pathway. In addition, N-glycans are crucial for protein folding, stability and modulating protein-protein interactions. Many human therapeutics derived from recombinant biopharmaceuticals are N-glycosylated; as such, N-glycosylation is important for recombinant glycoprotein production as it influences their heterogeneity and affects the biological function of these products. Furthermore, incomplete N-glycosylation may give rise to recombinant proteins with unwanted characteristics. Recombinant glycoprotein glycosylation in mammalian cells is typically efficient, whereas glycosylation in plants, insect cells and other expression systems, such as *Pichia pastoris*, display differences in glycosylation compared to mammalian cells (e.g. plants harbour a β(1,2)-xylose glycan core whereas mammals do not and the α(1,6)-fucose in mammals is substituted by an α(1,3)-fucose in plants) – this may alter the above mentioned protein folding, stability and modulation of protein-protein interactions with unknown effects thus affecting the efficiency of glycosylation in these alternative, non-natural host systems. To remedy this, in a *Nicotiana benthamiana* expression system Castilho et al. (2018) co-expressed immunoglobulins and other recombinant human glycoproteins (erythropoietin, interferon-ɤ and the Fc antibody region) with a recombinant

oligosaccharyltransferase from *Leishmania major*. The results showed a significant increase in N-glycosylation efficiency for some proteins and a promising tool for producing recombinant glycoprotein therapeutics in *N. benthamiana*.

1.4 CHIKV Therapeutics, Diagnostics and Vaccines

While the mortality rate of CHIKV infection is relatively low, the debilitating nature of the disease and its corresponding economic burden is such that the development of therapeutics, diagnostic assays and candidate vaccines against CHIKV are important considerations for combating CHIKV spread and infection, effectively diagnosing possible sufferers early in high risk-areas and limiting the burden of the disease (Silva and Dermody, 2017).

1.4.1 Current Therapeutics

At present, there are no licensed anti-CHIKV therapeutics available; however, numerous therapeutic strategies have been investigated. Certain antivirals, for example, have been discovered or developed to directly target the CHIKV replication cycle thereby inhibiting viral entry, replication and/or viral release (Tharmarajah et al., 2017). Pre-existing antivirals have been tested for anti-CHIKV properties and it was found that ribavirin in combination with interferon-alpha (IFN-α), used in CHIKV-infected patients, showed positive anti-CHIKV effects *in vitro* – it is the only FDA-approved drug in humans that produced this response (Turner et al., 2014). Similarly, favipiravir (T-705), an anti-influenza virus agent, in combination with its defluorinated analogue T-1105 inhibited CHIKV replication *in vitro* (Delang et al., 2014). Likewise, harringtonine, a plant alkaloid compound used to treat chronic myeloid leukaemia, mycophenolic acid, an immunosuppressant used for organ transplantations (Kaur et al., 2013), and 6-azauridine, an anti-metabolite (Pohjala et al., 2011), displayed CHIKV replication impairment *in vitro*. Additional small-molecule antivirals such as nucleoside analogue β-D-N^4-hydroxycitidine (NHC) have been shown to inhibit CHIKV replication *in vitro* (Ehteshami et al., 2017) and proved to be more potent than ribavirin and favipiravir (Tharmarajah et al., 2017).

Monoclonal antibodies as therapeutics or prophylactics against CHIKV have also been investigated since a humoral response and neutralizing antibodies against CHIKV clades and other alphaviruses are often protective strategies. In mice, passive immunisation by transfer of antisera or isolated monoclonal antibodies has been shown to be protective against CHIKV (Selvarajah et al., 2013). In addition, monoclonal antibodies as a postexposure therapeutic, even when administered in the late stages of infection, have been shown to be protective in mice against CHIKV (Fric et al., 2013). These showed further promise for the use of immunotherapy as a therapeutic against CHIKV disease.

Although the discovery of these antiviral and therapeutic compounds has shown protective promise against CHIKV infection, further studies in animal and human trials are needed for the advancement of these drugs. The best strategy against CHIKV infection and spread is likely to be an effective vaccine (Silva and Dermody, 2017).

1.4.2 Current Diagnostic Assays

The overlapping symptoms, geographical distribution and, in some cases, vectors shared by CHIKV, DENV, o'nyong-nyong, Mayorovirus and other similarly-symptomatic viruses highlight the need for definitive diagnostics to unambiguously identify CHIKV infection. An effective diagnostic will permit early detection of infection and, consequently, an appropriate clinical and public health response (Pezzi et al., 2019).

To date, three common laboratory assays are used to diagnose CHIKV infection: virus isolation, serological testing and reverse transcriptase-polymerase chain reaction (RT-PCR) (Galán-Huerta et al., 2015). While virus isolation is the most definitive diagnostic strategy, it is not a strategy that fits the need for early diagnosis of CHIKV infection. Virus isolation is being replaced with faster, more sensitive and safer molecular techniques that do not require the laboratory infrastructure needed for virus isolation. However, virus isolation should not be discounted for the fact that it can accurately characterise circulating strains and can further assist with the development of future diagnostic assays (Pezzi et al., 2019). A limitation of this method is that the blood sample needs to be transported between 2 – 8°C and used within 48 hours.

For serological assays, sera from whole blood, obtained from acute and recovering sufferers, can be used in enzyme-linked immunosorbent assays (ELISA). Like virus isolation, the non-frozen specimen needs to be transported at 2-8°C. These assays rely on CHIKV-specific IgM levels and a four-fold rise in IgG titre for diagnosis (Galán-Huerta et al., 2015). Unfortunately, samples collected during the acute and convalescent stages of the disease are used for a reliable diagnosis of CHIKV infection. As such, this technique is not ideal for early reliable diagnosis of infection.

RT-PCR is the more frequently used technique for diagnosis and the use of serum from whole blood is favoured. Other clinical samples like urine, saliva, semen, cerebrospinal fluid, and breast milk have been used; however, the time window for some samples (i.e. breast milk and urine) remain unclear and are often contradictory in the literature (Grivard et al., 2007, Campos et al., 2017). Issues with this strategy of diagnosis arise when mismatches occur between different strains of the virus and in-house primers (published in the literature) are used. In addition, for commercial real-time PCR assays, very limited or no information is available regarding the adequacy of the lineages of the CHIKV strains detected and how these kits were evaluated (Pezzi et al., 2019).

While these tests are reliable and sensitive, they require regularly maintained and sophisticated infrastructure and medium-to-high level skilled technicians to operate. These requirements may not be available for certain laboratories, especially in rural areas and areas lacking sophisticated infrastructure where outbreaks typically occur. As such, a reliable capture antigen-based rapid diagnostic assay has been recommended for the control of outbreaks and discrimination between co-circulating febrile viruses (Burdino et al., 2016). A recently developed capture antigen-based diagnostic assay used a recombinant CHIKV E2 capture antigen produced in E. coli. While the developed assay was simple, specific and sensitive, it was produced in a relatively expensive expression system (Fumagalli et al., 2018). Recombinant CHIKV E1 and E2 envelope glycoproteins have also been expressed in Sf9 cells (insect cell line) using a baculovirus vector-based approach. Their seroreactivities were evaluated by indirect IgM-capture ELISA. Recombinant CHIKV E1 showed a sensitivity of 77.5% while recombinant CHIKV E2 showed a sensitivity of 90% (Cho et al., 2008). However, these recombinant antigens were also produced in a relatively expensive expression system.

1.4.3 Current Vaccine Candidates

While there is no commercialised vaccine against CHIKV, various vaccine strategies have been investigated, allowing for some candidates to advance to preclinical and clinical studies. Summarised here are the candidates that have advanced to human studies.

TSI-GSD-218 is the first reported live-attenuated candidate vaccine against CHIKV. Attenuation of CHIKV is credited to a two-point mutation (T12I and G82R) of the E2 glycoprotein. The candidate vaccine displayed immunogenicity in mice and non-human primates (NHPs) and conferred protection in NHPs when challenged with CHIKV (Levitt et al., 1986). TSI-GSD-218 showed no adverse effects in phase I clinical trials and was highly immunogenic in phase II clinical trials with a single immunisation; however, approximately 8% of the subjects displayed transient febrile symptoms. Furthermore, as residue 82 is the dominant attenuating factor and influences CHIKV-specific neutralizing antibody production, observing reversion at this residue in mice and humans raised concerns about attenuation stability and the resulting safety of the vaccine. Consequently, the candidate vaccine was ceased after phase II studies (Edelman et al., 2000).

Another vaccine candidate, CHIKV/IRES, contains an encephalomyocarditis virus internal ribosomal entry site (IRES) sequence in place of the subgenomic promoter which serves to limit the expression of viral structural proteins. This modification allows for the attenuation of viral replication in mammalian cells and prevents replication in insect cells, thereby inhibiting replication in mosquito cells. Furthermore, the CHIKV/IRES vaccine candidate is safe, it displayed high immunogenicity in mice and NHPs, and remained attenuated (Plante et al., 2015).

Other such promising strategies are those of virus-like particle (VLP)-based vaccine candidates. CHIKV VLPs are made by expressing the structural polyprotein cassette via a DNA-expression plasmid transformed or transfected into an expression platform. The resultant particles mimic intact virions, however, due to a lack of genomic viral RNA, they are replication-incompetent. VRC-CHKVLP059-00-VP, a VLP-based candidate, is derived from DNA-transfected human embryonic kidney (HEK) cells. While it has been shown to

be safe, highly immunogenic and cross-protective against multiple CHIKV strains in mice and NHPs, multiple doses are required, and manufacturing costs are high (Akahata et al., 2010, Plante et al., 2015). VRC-CHKVLP059-00-VP is in phase II trials. Another VLP-based candidate vaccine in phase II trials makes use of a recombinant, live-attenuated measles virus (MV) vector (MV-CHIKV). The vaccine candidate was well tolerated and protective against lethal challenge due to its cross-neutralising immune response; however, multiple doses were required for 100% seroconversion and 58% of the high-dose cohort displayed adverse effects to multiple vaccinations (Ramsauer et al., 2015).

Other approaches have included other live-attenuated strategies including viral gene deletions, glycosaminoglycan binding modifications, alphavirus chimaeras, and codon alterations, as well as strategies that have included inactivated particles, subunit, and DNA- and recombinant virus-vectored vaccines (Silva and Dermody, 2017).

Even with the promise of CHIKV treatments and diagnostics, the most effective eradication and control strategy for CHIKV infection is still prevention via an efficacious vaccine. Despite the potential of these vaccine candidates, there are still numerous obstacles that need to be overcome. The ideal vaccine candidate must illicit a robust immune response, be cost-effective and ideally be administered in a single dose. This introduces the challenge of balancing vaccine safety and immunogenicity while maintaining economic viability regarding vaccine production and licensing (Silva and Dermody, 2017).

1.5 Plant-Based Protein Expression and Downstream Processing (DSP)

1.5.1 Expression in Plants

Until recently, mammalian cell lines such as Chinese Hamster Ovary (CHO), yeast and bacterial cell lines were predominantly used for biopharmaceutical production. With the current rise in demand for biopharmaceuticals, alternatives to mainstream expression systems that are cheaper, safer and more convenient are being investigated. Plants are one such attractive alternative that have emerged as a result of their high scalability, low capital and production costs, minor differences in glycosylation compared to traditional expression systems, and safety owing to their lack of oncogenic DNA sequences and

prions (Tschofen et al., 2016, Rybicki, 2017). However, there are some factors to consider pertaining expression yields in plant expression technologies.

While plant expression systems are impressive alternatives, they were not without their issues in their early years and, in some respects, even now. Since, many technical advances have been achieved to relieve the technology of these issues and further shuttle plant-based expression into the affordable technology that it is now (Fischer et al., 2004). One key problem was/is low expression yields. While it is rare for plant-based based system to reach the protein yields of traditional expression system, protein yield in plants can be further optimised by controlling genetic and epigenetic factors, in addition to cultivation parameters (Schillberg et al., 2019). Factors that determine expression yield include: the host species, gene silencing, transcription and translation efficiency, promoter activity and specificity, codon usage, subcellular targeting and the corresponding stability of the produced recombinant protein, and the expression system/construct. These factors should all be considered during upstream study design. A range of plant species have been used in molecular farming: leafy crops, cereals, fruit and vegetable crops, and some miscellaneous species. Tobacco has the highest biomass of the crop species used, it is easily and highly scalable, fully optimised for gene transfer and expression and, owing to its non-food, non-crop status, it has a lower risk of food or feed chain contamination (Ma et al., 2003). Transcription efficiency can be optimised using an effective promoter, like the strong constitutive promoter of the Cauliflower mosaic virus (CaMV), 35S (Fischer et al., 2004), and a polyadenylation signal, like the *Agrobacterium tumefaciens nos* gene . Translation efficiency can be optimised by adapting, and thus optimising, codon usage to that of the host species (Liénard et al., 2007). Subcellular targeting of proteins also affects protein stability. This is because the compartment in which the protein accumulates influences its assembly, folding and post-translational modifications. The secretory pathway has been shown to result in better assembly and folding of glycoproteins, for example, than the cytosol. This must be considered during construct design by the addition of a compartment-specific C- or N-terminal signal-peptide sequence in the expression construct. Previous optimisation of these factors has allowed for cheaper upstream costs of plant-produced biologics in contrast to itself (in the early years of the technology) and other traditional expression

systems (mammalian/insect cell lines and microbial fermentation). Like upstream processing, there is much to be considered for downstream processing of recombinant plant-produced proteins (Nagels et al., 2012).

1.5.2 Downstream Processing of Recombinant Proteins

Downstream processing of plant-derived recombinant proteins accounts for the majority of the production costs in extraction, purification and characterisation (Nandi et al., 2016). Hence, reducing upstream costs is one of the fundamental advantages of plant-expression systems. It has become clear that different expression hosts come with their own set of advantages and disadvantages regarding DSP. As a result, DSP infrastructure availability and extraction/purification protocols should be considered.

Purification costs can be reduced using affinity tags (e.g. polyhistidine tag) on one or both termini of the protein. A disadvantage that may arise by the addition of a tag/tags is their interference on the activity and/or structure of the protein. Thus, the possibility of removing the tag(s) from purified recombinant proteins may be necessary for the restoration of the activity and/or structure of the protein (Fischer et al., 2004). The addition of an affinity tag, the polyhistidine tag specifically, allows for relatively inexpensive, large-scale one-step purification procedures through metal affinity chromatography. This method uses adsorbents that are resistant to chemical and/or biological degradation and boast a high protein-binding capacity. Another efficient method is ion-exchange chromatography. Both of these methods can be used on a variety of biomolecules. (Platis et al., 2008).

1.5.3 Transient Expression in Tobacco

The value of molecular farming, as an alternative to traditional production platforms like *E. coli* and mammalian cells, is such that large amounts of pharmaceutically important proteins can be produced in more affordable ways. Proteins can be expressed in transgenic or transient systems. Each method of expression has its respective advantages and disadvantages (Fischer and Emans, 2000).

With the emergence of transient expression, the process of expression has significantly sped up compared to transgenic expression as screening of the recombinant protein of interest and large-scale production are possible within days (Castilho et al., 2018). There are three major methods of transient-expression: delivering a gene into plant cells by particle bombardment, *Agrobacterium*-mediated gene transfer and the use of modified viral vectors. *Agrobacterium*-mediated gene transfer delivers DNA to more cells than particle bombardment that not only reaches fewer cells, but is also not ideal for large-scale expression and, the DNA also needs to reach the nucleus for DNA transcription (Christou, 1996). In addition, the T-DNA containing the gene of interest is actively delivered to the nucleus when using agroinfiltration. This method can produce suitable amounts of protein when scaled up which may be sufficient for pre-clinical trials thus bypassing the need for a stable-expression plant line. Viral vectors deliver the gene of interest to most cells and transcription of this gene is achieved by viral replication in the cytoplasm which transiently delivers many transcripts of the gene of interest (Fischer and Emans, 2000). Viral vectors offer the advantage of autonomous replication, and their characteristic high levels of multiplication and naturally occurring transient expression levels. Viral vectors operate through transient RNA and DNA replicons that are viral-based and offer the rapid advantage of high levels of expression; however, the lack of stability of the foreign gene in the viral genome may pose problems (Scholthof et al., 1996).

1.6 Plant-Based Research for CHIKV

Worldwide, developing countries are the most affected by CHIKV epidemics. To combat this, plant-based expression platforms have been proposed as the tool to address the lack of vaccine and rapid diagnostic kit development and availability. In addition to its scalability, safety, efficiency and product quality, it operates at a fraction of the cost (0.1 to 10%) of traditional expression platforms like bacteria and mammalian cell-culture (Cardona-Ospina et al., 2016).

Beforehand, protein candidates that may potentially be used for the development of a biopharmaceutical should be defined. For a diagnostic assay: these proteins should be highly conserved (Cardona-Ospina et al., 2016). As per previous research done on CHIKV,

CHIKV nsP1 cysteine protease has high genetic diversity and is unsuitable for use in a diagnostic assay. Likewise, the other non-structural proteins are not suitable for use in a diagnostic assay. (Bao et al., 2013). The CHIKV envelope glycoproteins E1 and E2 are each suitable for the development of a rapid diagnostic kit. Both glycoproteins are highly conserved; however, plant glycosylation of these recombinant proteins may modify their effectiveness (Khan et al., 2012, Weber et al., 2015). The CHIKV E2 protein is the most conserved protein of the structural protein cassette and is responsible for earliest elicitation of an antibody response (Kam et al., 2012).

In addition, the plant system for the expression of recombinant CHIKV proteins should be chosen based on availability of infrastructure, risk of contamination, the effects of secondary metabolites on DSP, and the previous success of transformation protocols on certain plant species e.g. *Agrobacterium*-mediated infiltration is particularly effective for transformation of *Nicotiana benthamiana, N. tabacum* and *Solanum tuberosum*. The transient method of agroinfiltration is preferred as it can provide suitable yields of protein short term (Cardona-Ospina et al., 2016).

1.7 Project Aims and Objectives

The development of a rapid diagnostic assay to screen for CHIKV antibodies in sera is currently expensive and hazardous owing to the BSL conditions required for the operation of live virus-infected tissue and its products. The aim of this study was to develop a CHIKV diagnostic reagent by expressing recombinant variants of the CHIKV E2 glycoprotein, using a tobacco plant-based transient-expression system. The following encompasses the approach used:

- Synthesis of two codon-optimised CHIKV E2 glycoprotein gene variants and cloning into two plant expression vectors, pEAQ-*HT* and pTRAkc-ERH, in *E. coli*.
- Transformation of appropriate *Agrobacterium tumefaciens* strains with recombinant plant expression constructs.
- *Agrobacterium*-mediated infiltration of *N. benthamiana* leaves with different plant expression constructs. This includes co-infiltration with recombinant human chaperone proteins and an RNA-silencing suppressor protein to determine which combination, if any, provides optimal CHIKV E2 glycoprotein expression.

- Screening of plant produced CHIKV E2 glycoprotein variants to determine which day post-infiltration the recombinant proteins are expressed the highest.
- Selection of plant expression construct and conditions for large-scale expression and subsequent development and optimisation of an extraction and purification protocol.

Chapter 2: Expression of Recombinant CHIKV E2 Glycoprotein Variants in Plants

2.1 Introduction

In this chapter, the small-scale transient expression of the two CHIKV E2 glycoprotein variants in tobacco plants is described and analysed to select the best-expressed variant and its conditions for analysis in large scale expression. The CHIKV envelope 2 (E2) glycoprotein has been shown to be potently immunogenic owing to a single linear epitope located on the N-terminus of the glycoprotein. This epitope protrudes from the virion and is the major target of detection by antibodies during the early phase of infection. This protein is therefore an appropriate target for use as a reagent in an assay for the diagnosis of relatively undiagnosed/misdiagnosed CHIKV (Kam et al., 2012). Two recombinant variants of the CHIKV E2 glycoprotein have previously been expressed in insect cells at high levels (Metz et al., 2011); however, the variant lacking the C-terminal transmembrane domain showed higher levels of expression than the full-length CHIKV E2 glycoprotein In addition, it has been seen that expression of general transmembrane-containing proteins utilising the plant-expression planform elicits necrotic symptoms on the leaves of tobacco-plant hosts (Peyret and Lomonossoff, 2013). These symptoms have been shown to alleviate upon expression of the same, transmembrane domain-deficient protein. This truncation also improved expression and purification of the protein. These more favourable aspects suggest that it is worth investigating the expression of recombinant transmembrane-containing and transmembrane-deficient CHIKV E2 glycoprotein variants in the more cost-effective, rapid and highly scalable plant-expression system.

During the design of a cloning strategy for recombinant protein expression in plants, there are many tools that can be implemented to assist with optimisation of gene expression, protein folding, protein detection and protein purification. Optimisation of these processes help with the time and cost aspects of the research and end product integrity. Codon optimisation of recombinant gene sequences is one important tool as codon usage, while seemingly random, varies between organisms (Stratford et al., 2000). The expression of an exogenous gene may be impeded as a result of the protein coding regions of the exogenous gene having a codon usage pattern different to that of the host organisms. Codon optimisation may improve the translation efficiency of exogenous

genes as it rhymes the protein coding usage of the coding regions (Fuglsang, 2003). Previous research on expressing a recombinant plant-produced human epidermal growth factor (hEGF), a powerful mitogen being recombinantly produced on a pharmaceutical scale that enhances the healing of a wide range of injuries, showed that optimising the recombinant gene's codons for plant expression increased hEGF yield by ~34% (Thomas and Walmsley, 2014). Similarly, the expression of a plant codon-optimised tetanus toxin's non-toxic fragment C (TetC), for the development of a subunit vaccine against tetanus in tobacco plants, accumulated to be 10% of the total soluble proteins (TSP) extracted from harvested leaf material (Tregoning et al., 2003). For the expression of a protein to become economically viable, it should preferably accumulate to at least 1% of TSP (Rybicki, 2009).

Furthermore, to assist with protein targeting to the endoplasmic reticulum (ER), where they will undergo the process of proper folding, a signal-peptide sequence can be fused to recombinant gene sequences. One such signal peptide is the Leader Peptide Heavy (LPH)-signal peptide derived from the murine monoclonal antibody mAb24 heavy-chain gene used to target recombinant proteins to the secretory pathway (Schillberg et al., 1999). Previous fusion of the LPH-signal peptide onto a recombinant *N. benthamiana* plant-produced Rift Valley fever virus (RVFV) Gn glycoprotein, chimeric RVFV Gn/influenza HA particle-forming fusion protein (Mbewana, 2017), and HIV Env gp140 protein (Margolin et al., 2019) has shown promise for the production of these proteins as a result of the signal peptide's targeting function.

An advantage of using tobacco plants as a transient expression system is that it allows for rapid production of proteins which can then be screened through antibody-related assays like western blotting and enzyme-linked immunosorbent assays (ELISA); however, detection of these plant-produced proteins is dependent on the availability and specificity of commercialised/in-house antibodies which are often lacking and not very specific. In order to bypass this problem, recombinant proteins can be fused to an assortment of affinity tags for which commercial antibodies are commonly available. This bypasses the requirement for protein-specific antibodies and allows for efficient downstream one-step purification of the recombinant tagged protein through the tag's ability to bind to certain resins (Einhauer and Jungbauer, 2001). One such affinity tag is

the calmodulin-binding protein (CBP). CBP is a 26 amino-acid peptide obtained from the muscle myosin light chain kinase (MLCK) which binds to calmodulin with relatively high affinity (Zheng et al., 1997); however, the affinity resin is relatively expensive (Waugh, 2005). Another useful affinity tag is glutathione S-transferase (GST), a 218 amino-acid peptide that offers the advantage of potential protein solubility (Sehr et al., 2001) and inexpensive GSH-Sepharose affinity resin; however, large protein tags often burden the expression host metabolically (Waugh, 2005).

The smaller and highly-specific 8-amino acid FLAG™ marker peptide is very effective but scale-up procedures are costly as a result of the tag's low binding capacity and expensive resin (Einhauer and Jungbauer, 2001). One of the most popular affinity tags is the polyhistidine tag (6xHis-tag), which comprises six neighbouring histidine residues fused to either the N- and/or C-terminus of the recombinant protein. Additionally, the tag can be removed by engineering a cleavage site between the N- and/or C-terminus of the protein of interest and the fused 6xHis-tag (Zentgraf et al., 1995) and primary antibodies are commercially available for immunoassay detection. Thus, the fusion of a 6xHis-tag onto recombinant gene sequences during cloning is often beneficial to the costs and efficiency involved with downstream protein detection and purification. One significant disadvantage of using the 6xHis-tag is that the detection antibody can be promiscuous. This has the potential to result in background detection of proteins that are not the protein of interest (Kimple et al., 2013).

Although some proteins are expressed at high levels in plants, glycoproteins in particular seem to be expressed at very low levels. Evidence suggests this can be credited to the inherent differences in endogenous chaperone machinery between the plant and natural mammalian cell hosts - where viral glycoproteins are naturally produced - in addition to a potential bottleneck in the plant folding machinery to produce exogenous viral glycoproteins (Margolin et al., 2018). Furthermore, it has been shown that the expression of viral glyco-antigens produces a pathology in the plants indicative of plant ER-stress (Pera et al., 2015, Phoolcharoen et al., 2011). Chaperone proteins are responsible for the correct folding, during or after synthesis, of proteins. Two chaperones that have previously been shown to improve expression of recombinant proteins in plants, when co-expressed with the protein of interest, are human chaperones calreticulin (CRT) and

calnexin (CNX). Membrane-bound CNX and soluble CRT are both quality-control proteins that recognise glycoproteins, facilitate their folding in the ER and prevent incompletely or misfolded proteins from exiting the ER (Nagels et al., 2012). Previous work on improving recombinant protein production by co-infiltration of *N. benthamiana* plants with chaperones showed an increase in yields of *E. coli* type-3 secretion system proteins when co-expressed with the chloroplast-targeted chaperone CesT (MacDonald et al., 2017).

Another helper protein that may improve recombinant protein expression yields in plants is the RNA-silencing suppressor NSs protein from the tomato spotted wilt virus (TSWV). Plants harbour a post-transcriptional RNA silencing defence system that recognises and degrades transgenes from invading pathogens. The NSs protein encoded by the TSWV supresses this innate RNA silencing activity of plants (Takeda et al., 2002); as such, it can be used as a tool to suppress the deleterious effects of RNA silencing of exogenous expression of recombinant proteins and improve their expression.

In this study, two codon-optimised gene variants of a recombinant CHIKV E2 glycoprotein were designed to be expressed in a *N. benthamiana* transient-expression system. The 6xHis-tag was fused to the C-terminus of each gene variant for efficient downstream detection and purification purposes. The glycoprotein gene variants were cloned into transient plant expression vectors, pTRAkc-ERH (Meyers et al., 2008) and pEAQ-*HT* (Sainsbury et al., 2009). pTRAkc-ERH contains an LPH signal-peptide sequence on the 5' of the multiple cloning site (MCS) which is consequently fused onto the 5' terminus of the gene inserts during cloning; the pEAQ-*HT* vector does not contain this signal-peptide sequence. These constructs were transformed into *Agrobacterium tumefaciens* and infiltrated into *N. benthamiana* leaves using *Agrobacterium*-mediated infiltration. They were infiltrated either individually or co-infiltrated together with *Agrobacterium* strains harbouring the human chaperones calreticulin or calnexin or the RNA-silencing suppressor NSs. Samples were harvested at various days post infiltration to determine on which day post-infiltration (dpi) the highest yields are achieved and under which parameters.

2.2 Materials and Methods

2.2.1 Gene Design and Synthesis

A DNA consensus sequence (Appendix, Fig. A1) of three South African CHIKV isolates (GenBank: HM045792.1, Vereeniging, Gauteng province; HM045795.1, SAH2123 strain; HM045805.1, CHIKV AR 18211 1976) was used to design two gene variants of the CHIKV E2 glycoprotein (Ann Meyers, Biopharming Research Unit). Information on the South African isolates were gathered from Volk et al. (2010). One gene variant contained the full length E2 glycoprotein sequence including the transmembrane domain (TMD), and the other gene variant lacked the TMD (ΔTM). These were *Nicotiana* sp. codon-optimised using the OptimumGene™ algorithm, synthesised and supplied in a pUC57 vector using the EcoRV blunt-end restriction enzyme by GenScript® (USA) in an *E. coli* TOP10-strain glycerol stock. To facilitate cloning, both gene variants contained 5' BspHI and 3' XhoI restriction enzyme cut sites on the proximal ends of the gene. For downstream protein detection and purification purposes, a 6xHis-tag was fused to the 3' termini of the gene variants separated by a linker sequence (Fig. 2.1) – 5' GAG GCT GCT GCA AAG GCT 3'. The supplied constructs are referred to as pUC57-CHIKV E2-HIS and pUC57-CHIKV E2ΔTM-HIS.

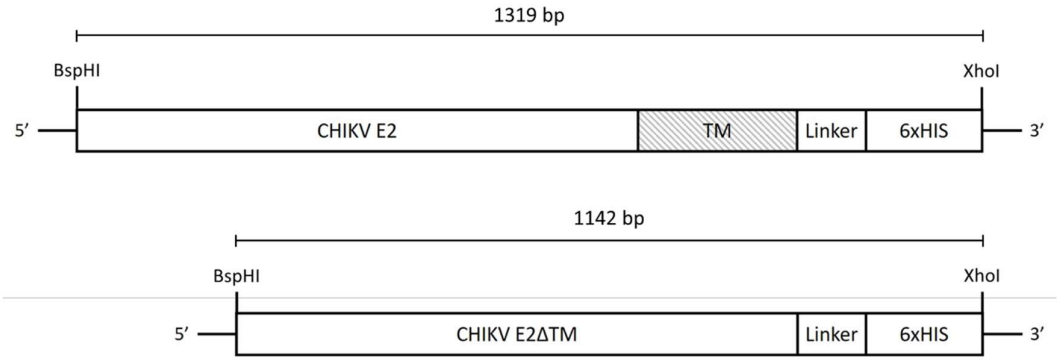

Figure 2.1: Synthesised CHIKV E2 glycoprotein gene variant constructs. Each variant contains a 5' BspHI and 3' XhoI restriction enzyme site, and a 3' polyhistidine tag (6xHIS) separated from the gene by a linker. The CHIKV E2ΔTM gene variant lacks the transmembrane domain (TM). The synthesised CHIKV E2-HIS construct is 1319 bp in length and the CHIKV E2ΔTM-HIS construct is 1142 bp in length

2.2.2 Cloning and Plant Expression Vectors

The pUC57 (containing CHIKV E2 gene variants synthesised and cloned by GenScript®, USA as pUC57-CHIKV E2-HIS and pUC57-CHIKV E2ΔTM-HIS), pUC18-LPH (Biopharming Research Unit, University of Cape Town; see 2.2.3.4) and pJET1.2/blunt (CloneJET PCR Cloning Kit, Thermo Fisher Scientific; see 2.2.3.5) cloning vectors were used for cloning purposes.

The vectors pEAQ-*HT* (Prof. G Lomonossoff, John Innes Centre, Norwich) and pTRAkc-ERH (Dr Rainer Fischer, Fraunhofer Institute for Molecular Biology and Applied Ecology, IME, Germany) (Fig. 2.2) were used for subcloning of the two gene variants and subsequent transient plant expression. The vectors and their antibiotic resistance markers are summarised in Table 2.1.

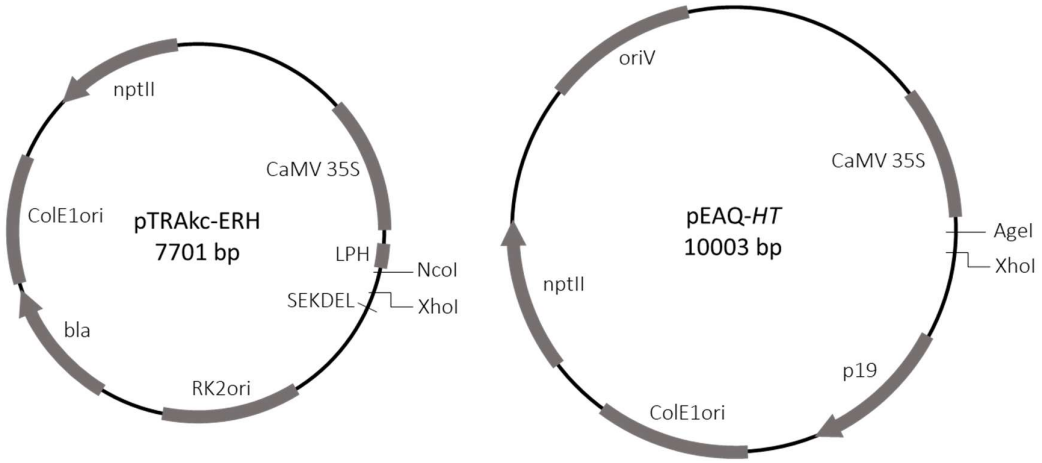

Figure 2.2: Plant expression vectors pTRAkc-ERH and pEAQ-*HT* used in this study. CaMV 35S, Cauliflower mosaic virus 35S promoter; ColE1ori, origin of replication for *E. coli*; nptII, kanamycin-resistance gene; LPH, signal-peptide sequence; SEKDEL, ER-retention signal sequence; RK2ori, origin of replication for *A. tumefaciens*; bla, ampicillin- and carbenicillin-resistance gene; p19, tomato bushy stunt-virus RNA-silencing suppressor; NcoI/AgeI/XhoI, restriction-enzyme sites for gene insertion.

Table 2.1: A summary of vectors used in this study.

VECTOR	*BACTERIAL STRAIN*	*ANTIBIOTIC, WORKING CONCENTRATION*
pUC57	*E. coli* DH5α	Ampicillin, 100 µg/mL

pUC18-LPH	*E. coli* DH5α	Ampicillin, 100 µg/mL
pJET1.2/blunt	*E. coli* DH5α	Ampicillin, 100 µg/mL
pEAQ-HT	*E. coli* DH5α	Kanamycin, 50 µg/mL
	A. tumefaciens LBA4404	Rifampicin, 50 µg/mL Kanamycin, 30 µg/mL
pTRAkc-ERH	*E. coli* DH5α	Ampicillin, 100 µg/mL
	A. tumefaciens GV3101:pMP90RK	Rifampicin, 50 µg/mL Carbenicillin, 50 µg/mL* Kanamycin, 30 µg/mL

* The GV3101:pMP90RK strain is a rifampicin- and kanamycin-resistant strain; however, upon generation of the recombinant clone, the clone gains additional carbenicillin resistance.

2.2.3 Cloning Strategy

2.2.3.1 General Cloning

CHIKV E2-HIS and CHIKV E2ΔTM-HIS were isolated from their pUC57 vectors for subsequent ligation into vectors. Vectors were linearised with appropriate restriction enzymes, resolved by 1% agarose gel electrophoresis and gel purified. Ligations of gene variants into vectors were facilitated by T4 DNA ligase (New England Biolabs) as per the manufacturer's instructions and incubated O/N at 4 ^{o}C. Thereafter, O/N ligation reactions were transformed into *E. coli* DH5α (E. cloni® 10G cells, Lucigen). Transformed bacterial suspensions were plated on LB agar plates with appropriate antibiotics and incubated at 37 ^{o}C O/N. Bacterial colonies were screened to confirm successful ligation and transformation by colony polymerase chain reaction (PCR) screening using appropriate primers. Colony PCR was performed using Ampliqon Taq DNA Polymerase (Ampliqon, Denmark) as per the manufacturer's instructions and resolved and analysed by 1% agarose gel electrophoresis. Primer names and sequences used in this study are summarised in Table 2.2. Plasmids were isolated from selected positive colonies and digested with appropriate restriction enzymes to confirm positive colony PCR results (single digest-control reactions were performed in the same buffer to that of double

digest reactions to ensure correct functionality of the individual restriction enzymes in the double digest reactions). If gene variants were cloned into an intermediary vector (pUC18-LPH or pJET1.2/blunt), the general cloning process was repeated for final cloning into the plant expression vector pEAQ-*HT*; however, for repetition of this process, like isolation of gene variants from pUC57 they were isolated from their intermediary vector instead in preparation for subsequent ligations.

Table 2.2: Summary of primers used in this study.

PRIMER	DIRECTION	NUCLEOTIDE SEQUENCE	PURPOSE
PTRA-F	Forward	5'-CAT TTC ATT TGG AGA GGA CAC G-3'	Colony PCR; Sequencing
PTRA-R	Reverse	5'-GAA CTA CTC ACA CAT TAT TCT GG-3'	Colony PCR; Sequencing; Gene amplification
AGEI-LPH-F	Forward	5'-ATA CCG GTC ATT TCA TTT GGA GAG GAC ACG-3'	Colony PCR; Gene amplification
PEAQ-F	Forward	5'-TTC TTC TTC TTG CTG ATT GG-3'	Colony PCR
PEAQ-R	Reverse	5'-CAC AGA AAA CCG CTC ACC-3'	Colony PCR
PJET1.2-F	Forward	5'-CGA CTC ACT ATA GGG AGA GCG GC-3'	Sequencing
PJET1.2-R	Reverse	5'-AAG AAC ATC GAT TTT CCA TGG CAG-3'	Sequencing
M13-F	Forward	5'-CGC CAG GGT TTT CCC AGT CAC GAC-3'	Colony PCR
M13-R	Reverse	5'-GAG CGG ATA ACA ATT TCA CAC AGG-3'	Colony PCR

Underlined sequence indicates added AgeI cut-site sequence for cloning purposes.

2.2.3.2 Isolation of gene variants

pUC57 CHIKV E2-HIS and pUC57 CHIKV E2ΔTM-HIS glycerol stocks were streaked onto Luria-Bertani (LB) agar plates supplemented with appropriate antibiotics (Table 2.1) and incubated at 37 ^{o}C overnight (O/N). A single colony containing each pUC57 construct was cultured in LB broth with appropriate antibiotics and incubated at 37 ^{o}C O/N, with agitation. pUC57-plasmid DNA was isolated using a QIAprep® Spin Miniprip Kit (Qiagen, Germany) as per the manufacturer's instructions. Thereafter, the gene variants were excised from their respective pUC57 plasmids by triple restriction enzyme digest with restriction enzymes BspHI and XhoI (Thermo Fisher Scientific), the two synthesised enzyme restriction sites framing the gene variants, and AflIII (Thermo Fisher Scientific). A double digest with BspHI and XhoI would have resulted in two identically sized fragments (one being the gene variants and one being a fragment from the vector backbone). Therefore, AflIII was used in these digests to cut the pUC57-vector backbone to ensure all fragments were different sizes for downstream gel purification of just the gene variant DNA fragments. The digested DNA was resolved by 1% agarose gel electrophoresis using a Tris-Borate-EDTA (TBE) buffer supplemented with 2 mg/mL of ethidium bromide for visualisation. The O'Gene Ruler™ 1kb DNA Ladder (Thermo Fisher Scientific) was used for size comparison of DNA fragments. The resolved gene variants were subsequently gel purified using the QIAquick® Gel Extraction kit (Qiagen, Germany) as per the manufacturer's instructions.

Each of the purified gene variants were then ligated into their destination expression vectors in this order:

- pTRAkc-ERH cloning strategy (Fig. 2.3): pUC57 → pTRAkc-ERH
- Original pEAQ-*HT* cloning strategy (Fig. 2.4): pUC57 → pUC18-LPH → pEAQ-*HT* (pEAQ-*HT* lacks an LPH signal-peptide sequence, thus, a subcloning step into pUC18-LPH was performed to add the LPH signal-peptide sequence onto the 5' terminus of the gene variants)
- New pEAQ-*HT* cloning strategy (Fig. 2.5): pTRAkc-ERH LPH CHIKV E2-HIS and LPH-CHIKV E2ΔTM → pJET1.2/blunt → pEAQ-*HT*

2.3.3.3 pTRAkc-ERH cloning strategy

The isolated gene variants were ligated into the pTRAkc-ERH vector (insert:vector ratio of 5:1), previously cut with NcoI (compatible with BspHI) and XhoI, adding an LPH ER signal-peptide sequence onto the 5' termini of the gene variants. After colony PCR screening, a positive clone was further confirmed by sequencing (Macrogen, Inc., South Korea). Both colony PCR screening and sequencing used pTRA-F and pTRA-R primers (Table 2.2). A summary of this cloning process can be seen in Fig. 2.3.

Confirmed pTRAkc-ERH LPH-CHIKV E2-HIS and pTRAkc-ERH LPH-CHIKV E2ΔTM-HIS constructs were transformed into electrocompetent *A. tumefaciens* GV3101:pMP90RK cells. Three hundred nanograms of DNA was added to 100 µL of electrocompetent cells in a chilled 0.1 cm electroporation cuvette (Molecular BioProducts, Inc.). The DNA and cells were mixed slowly and chilled on ice for 5 minutes. Cells were electroporated using parameters of 1.8 kV, 25 µF and 200 Ω; thereafter 900 µL of LB was transferred to the cells which were incubated at 27 °C for 2 hours. One hundred microlitres of transformed bacterial suspension was plated on LB agar plates with appropriate antibiotics. All LB agar plates were incubated at 27 °C for 3 days. Recombinant A*grobacterium* colonies were screened by colony PCR screening using pTRA-F and pTRA-R primers (Table 2.2).

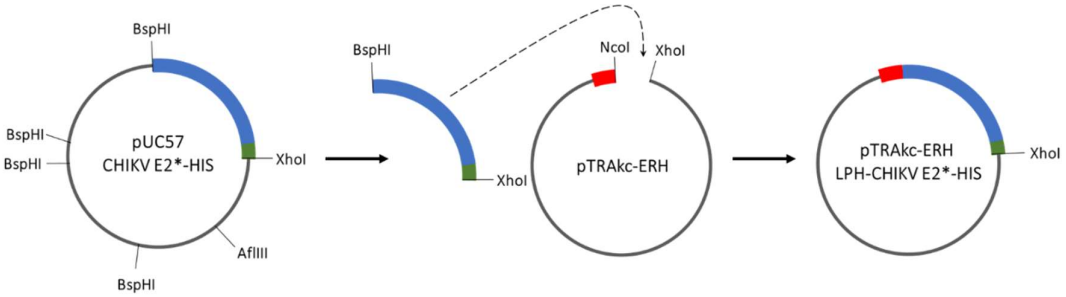

Figure 2.3: Summary of the cloning strategy to generate pTRAkc-ERH plant expression vectors harbouring the CHIKV E2 gene variants. The CHIKV E2 gene variants were excised from the pUC57 vector using restriction enzymes BspHI and XhoI, purified and subsequently ligated into pTRAkc-ERH previously cut with NcoI and XhoI. *, E2 or E2ΔTM.

2.2.3.4 Original pEAQ-HT Cloning Strategy

pEAQ-*HT* lacks an LPH signal-peptide sequence, thus, a subcloning step into pUC18-LPH (insert:vector ratio of 5:1) cut with NcoI (compatible with BspHI) and XhoI was necessary to fuse the LPH signal-peptide sequence onto the 5' terminus of the gene variants. In this strategy, the isolated gene variants were ligated into the pUC18-LPH vector and screened by colony PCR in order to isolate a clone containing LPH-fused gene variants. However, generation of pUC18 LPH-CHIKV E2-HIS and pUC18 LPH-CHIKV E2ΔTM-HIS was unsuccessful, despite multiple attempts (Fig. 2.12). A summary of the proposed cloning process can be seen in Fig. 2.4. An alternative strategy was developed (see 2.2.3.5).

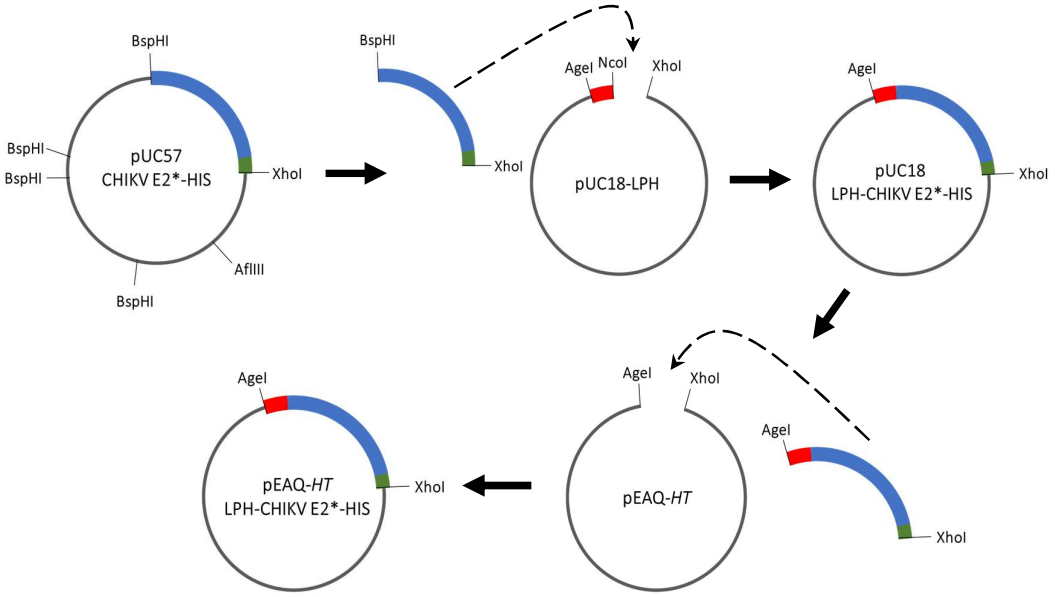

Figure 2.4: Summary of the cloning strategy to generate pEAQ-*HT* plant expression vectors harbouring the CHIKV E2 gene variants. The CHIKV E2 gene variants were excised from the pUC57 vector using restriction enzymes BspHI and XhoI, purified and subsequently ligated into pUC18-LPH previously cut with NcoI) and XhoI. The LPH-fused gene variants were then excised from the pUC18-LPH plasmid using restriction enzymes AgeI and XhoI, purified and subsequently ligated into pEAQ-*HT* previously cut with AgeI and XhoI. *, E2 or E2ΔTM.

2.2.3.5 New pEAQ-HT Cloning Strategy

Consequently, an alternative strategy was devised to clone CHIKV E2-HIS and CHIKV E2ΔTM-HIS into pEAQ-*HT*: the LPH-fused gene variants in the successfully-generated pTRAkc-ERH constructs (Fig. 2.10 and 2.11) were amplified by PCR using the pTRA-F primer with a 5' AgeI-sequence overhang (AgeI-LPH-F) and the pTRA-R reverse primer

(Table 2.2). An AgeI restriction cut site was added onto the 5' terminus of the gene variants, as a result, to facilitate downstream ligation of the amplified gene variants into pEAQ-*HT*. The CloneJET PCR Cloning Kit (Thermo Fisher Scientific) was used for amplification and subsequent subcloning of the amplified fragments into an intermediate pJET1.2/blunt vector according to the manufacturer's instructions to yield pJET1.2 AgeI-LPH CHIKV E2-HIS-pTRA-R and pJET1.2 AgeI-LPH CHIKV E2ΔTM-HIS-pTRA-R. Sequencing of pJET1.2 constructs was performed using pJET1.2-vector primers in Table 2.2. Following this, the isolated LPH-fused gene variants were excised out of the pJET1.2 constructs using AgeI and XhoI and inserted into pEAQ-*HT* (insert:vector ratio of 10:1) cut with AgeI and XhoI and dephosphorylated using rApid Alkaline Phosphatase (Roche) as per the manufacturer's instructions. For the new strategy, to rule out erroneous digestion and ligation steps as potential causes for the cloning issues seen in the original pEAQ-*HT* cloning strategy, a sample of the AgeI- and XhoI-digested pEAQ-*HT* vector was resolved by agarose gel electrophoresis (as described in 2.2.3.2) to ensure the vector was correctly digested in preparation for ligations of LPH-CHIKV E2-HIS gene variants. Likewise, samples of the subsequent ligation reactions of pJET1.2-isolated LPH-CHIKV E2-HIS gene variants were resolved by agarose gel electrophoresis to ensure correct ligations of gene variant isolates into the digested pEAQ-*HT* vector. The efficiency of these digests and ligations were screened before transformations. One microlitre of O/N ligation reactions were screened by PCR to investigate whether the recombinant pEAQ-*HT* LPH-CHIKV E2-HIS and pEAQ-*HT* LPH-CHIKV E2ΔTM-HIS constructs were generated successfully. Colonies were screened for positive clones by colony PCR. A summary of this cloning process can be seen in Fig. 2.5. Clones pJET1.2 LPH-CHIKV-E2-HIS and Clones pJET1.2 LPH-CHIKV-E2ΔTM-HIS were created.

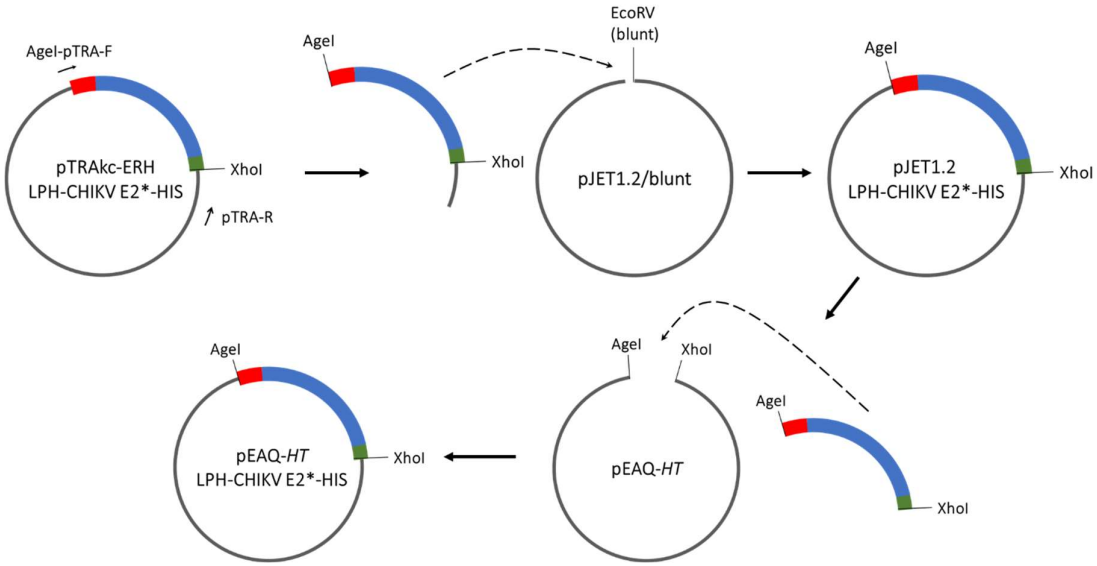

Figure 2.5: Summary of the cloning strategy to generate pEAQ-*HT* plant expression vectors harbouring the CHIKV E2 gene variants.The LPH-fused CHIKV E2 gene variants were amplified out of the pTRAkc-ERH LPH-CHIKV E2*-HIS constructs using the AgeI-LPH-F and pTRA-R primers, fusing a 5' AgeI sequence onto the gene variants, purified and subsequently ligated into pJET1.2/blunt vector previously blunt-end cut with EcoRV. The LPH-fused gene variants were then excised from the pJET1.2 plasmid using restriction enzymes AgeI and XhoI, purified and subsequently ligated into pEAQ-*HT* previously cut with AgeI and XhoI. *, E2 or E2ΔTM.

2.2.3 Small-Scale Agroinfiltration Expression Time-Trials

A. tumefaciens harbouring pTRAkc-ERH LPH-CHIKV E2 and pTRAkc-ERH LPH-CHIKV E2ΔTM were cultured in 10 mL LB media with appropriate antibiotics (Table 2.1) O/N at 27 ^{o}C with agitation (220 rpm). Thereafter, a 50 mL LB Broth base (2.5 g/L tryptone, 12.5 g/L yeast extract, 5 g/L NaCl, 1.95 g/L MES, pH 5.6) was inoculated with the O/N 10 mL LB culture with appropriate antibiotics (except rifampicin) and supplemented with 5 µL of 200 mM acetosyringone. The inoculated 50 mL culture was incubated as above. Subsequently, the concentration of the O/N culture was determined, using a 600 nm wavelength, and supplemented with resuspension solution (0.97 g/L MES, 2.03 g/L $MgCl_2$, pH 5.6) to achieve the desired optical density (OD_{600}) of 1.0 and 0.5, and 200 mM acetosyringone added to a final concentration of 20 µM. The resuspended cultures were left at room temperature (RT) for an hour prior to syringe infiltration to allow for induction of the *vir* genes by the acetosyringone. *A. tumefaciens* strains containing pTRAkc-ERH lacking any gene of interest (OD_{600} = 0.5) and pEAQ-*HT* Shuni-HIS (OD_{600} =

0.5)(Verbeek, 2019) were cultured, resuspended and infiltrated as above for use as negative and positive HIS controls, respectively.

Each resuspension was syringe infiltrated into the abaxial surface of three four to six-week old *N. benthamiana* plant leaves (grown under 16-h artificial light/ 8-h dark cycles, at 25 ^{o}C, 45% relative humidity).

A subsequent expression time trial was performed to investigate the effects that co-expressing human chaperone proteins, calreticulin (CRT) and calnexin (CNX), and plant RNA-silencing suppressor NSs have on CHIKV E2-HIS and CHIKV E2ΔTM-HIS expression levels. pTRAkc-ERH LPH-CHIKV E2 and pTRAkc-ERH LPH-CHIKV E2ΔTM were co-expressed with constructs encoding CRT, CNX or NSs. Constructs pEAQ-*HT* CRT (E. Margolin, Biopharming Research Unit, University of Cape Town), pEAQ-*HT* CNX (E. Margolin, Biopharming Research Unit, University of Cape Town) and pBIN-NSs (Takeda et al., 2002) were cultured and prepared for infiltration as above. Their respective ODs, *Agrobacterium* host strains and antibiotic resistance markers are summarized in Table 2.3.

Table 2.3: A summary of additional constructs used in this study to increase recombinant CHIKV E2 protein variant expression. Construct name, their optical densities in infiltration media, *Agrobacterium* strains and respective antibiotic resistance markers and antibiotic working concentrations.

CONSTRUCT	**OPTICAL DENSITY**	***Agrobacterium* STRAIN**	**ANTIBIOTIC, WORKING CONCENTRATION**
pEAQ-*HT* CRT	0.5	AGL-1	Kanamycin, 50 µg/mL Carbenicillin, 25 µg/mL
pEAQ-*HT* CNX	0.5	AGL-1	Kanamycin, 50 µg/mL Carbenicillin, 25 µg/mL
pBIN-NSs	0.5	LBA4404	Kanamycin, 30 µg/mL Rifampicin, 50 µg/mL

2.2.4 Small-Scale Protein Extraction and Quantification of recombinant proteins

Infiltrated plants were grown under the conditions described in 2.2.3 for 7 days post-infiltration (dpi). Leaf-disks were sampled on days 3, 5 and 7 dpi using the cap of a 2 mL Eppendorf reaction tube. Three leaf-disks were harvested (one clipping per leaf infiltrated) per construct per dpi per OD. Tubes containing leaf-disks were stored at -80 ^{o}C until further processed.

Once all samples had been harvested, leaf-disks were thawed at RT. Subsequently, 200 µL of 1xPBS supplemented with protease inhibitor (cOmplete™, EDTA-free Protease Inhibitor Cocktail, Roche) and a 60 mm diameter ceramic bead (Merck and Co, Kenilworth, USA) were added to each 2 mL Eppendorf tube containing the thawed leaf-disk samples. Tubes were subsequently added to a vortex mixer shaking at the highest setting at 4°C until plant material was sufficiently pulverised. Pulverised samples were then centrifuged for 15 000 x g for 10 mins to pellet insoluble material and supernatants were transferred to a new 1.5 mL Eppendorf reaction tube to be prepared for resolution by SDS-PAGE (see section 2.2.5). Denatured crude extract samples were then stored at -20 ^{o}C until analysis.

Prior to crude extract sample resolution by SDS-PAGE, samples were quantitated using a Bradford assay (BioRad). The standard curve, from which protein sample concentrations were calculated, was generated using a range of known bovine serum albumin (BSA) concentrations. All samples and standards were measured in triplicate.

2.2.5 Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Protein samples were prepared by adding 5 x sample application buffer (SAB) comprising of 94 mM Tris-Cl pH 7.5, 1.9 mM EDTA, 10% SDS, glycerol and bromophenol blue to a final dilution of 1:5. If samples had previously been stored at -80 ^{o}C, they were thawed at room temperature before the addition of 5 x SAB. Thereafter, the samples were denatured at 95 ^{o}C for 10 mins. Equal amounts of Denatured total soluble protein (TSP) were resolved by 10% SDS-PAGE at 120 V by means of BioRad Mini-Protean© Tetra Cell Systems (BioRad Laboratories, Hercules, USA) electrophoresis equipment.

For Coomassie staining, Coomassie brilliant blue stain (0.1% (m/v) Brilliant Blue R-250, 50% (v/v) glacial acetic acid and 10% (v/v) glacial acetic acid) was used to stain resolved samples at 37 °C for 1 hour prior to submerging the gel in destaining solution (45% (v/v) methanol, 10% (v/v) glacial acetic acid) O/N at room temperature with agitation.

2.2.6 Western Blotting

Following electrophoresis of equal TSP by SDS-PAGE, expression of recombinant plant-produced proteins was analysed by Western blot. The resolved proteins were transferred onto a nitrocellulose membrane (Amersham™ Protran™ Premium 0.45 µm NC) at 15 V for 90 mins using a Bio-Rad Trans-Blot Semi-dry transfer blotter (Bio-Rad Laboratories, Hercules, USA). Thereafter, the membrane was blocked using blocking buffer (5% (v/v) skim milk, 10% (v/v) 10xPBS, 0.1% (v/v) 10% Tween 20) for 30 mins at RT and then probed with mouse-produced primary anti-6xHis antibody (Thermo Fisher Scientific) diluted (1/2 000) in blocking buffer at 4 °C O/N with agitation. Following this, the membrane was washed with blocking buffer for 15 mins x 4. Diluted (1/10 000) goat-produced secondary alkaline phosphatase-conjugated antibody was used to probe the membrane for 1 hr at 37 °C with subsequent washes using blocking buffer without milk for 15 mins x 4. Probed proteins were detected using nitro-blue tetrazolium and 5-bromo-4-chloro-3'-indolylphosphate (NBT/BCIP; SeraCare Life Sciences Inc., Milford, USA). Rabbit-produced primary anti-CRT (abcam, Cambridge, United Kingdom), anti-RuBisCo (BIOCOM biotech, South Africa) and their goat-produced anti-rabbit secondary antibodies were diluted (1/5 000, 1/10 000, 1/10 000, respectively) and used as above.

2.2.7 Large-Scale Vacuum Agroinfiltration

Ten mL of LB media was inoculated with a glycerol stock of the recombinant *A. tumefaciens* harbouring the pTRAkc-ERH LPH-CHIKV E2ΔTM-HIS and pEAQ-*HT* CRT constructs with appropriate antibiotics. The cultures were incubated at 27 °C O/N with agitation. Thereafter, 50 mL LBB media was inoculated with the 10 mL O/N cultures with appropriate antibiotics and incubated O/N at 27 °C with agitation. The following day, 500 mL LBB media was inoculated with the 50 mL O/N cultures. The final inoculated culture

was supplemented with appropriate antibiotics, except rifampicin, and 200 mM acetosyringone to a final concentration of 20 µM. The cultures were incubated as above. On the day of infiltration, the OD_{600} of the cultures were determined and subsequently supplemented with resuspension media, to the desired respective OD_{600}, and 200 mM acetosyringone to a final concentration of 200 µM. Leaves were harvested on 5 dpi and stored at -80 ^{o}C for large-scale protein extraction (see section 2.2.8)

2.2.8 Large-Scale Protein Extraction

Harvested leaves from -80 ^{o}C storage were homogenised, without thawing, in 1 x PBS (2 mL/g leaf material) supplemented with protease inhibitor, at first, using an Ultra-Turrax® homogeniser (IKA® Works Inc., NC, USA). Thereafter, crude extract was agitated at 4 oC for 60 mins and then filtered through Miracloth™ (22-24 µm pore, Millipore, USA). The filtrate was centrifuged at 10 000 rpm for 15 mins at 4 oC and the supernatant was either stored at -80 ^{o}C or denatured (see section 2.2.5 at 95 ^{o}C for 10 mins.

2.2.9 Stability Studies

Once the optimal conditions of recombinant CHIKV E2 protein variant expression were chosen, it was worth investigating the stability of the chosen protein variant at -80 ^{o}C over time for the utilisation of the protein for incorporation into a diagnostic kit. Additionally, the stability of the protein variant was investigated with and without protease inhibitor as prolonged stability of a protein in an extraction buffer without protease inhibitor lessens the DSP costs. Proteins were extracted, as described in 2.2.8, with and without protease inhibitor (PI), from harvested leaves after large-scale vacuum infiltration. Three 100 µL samples for each type of crude extract (with PI and without PI) were transferred to an Eppendorf reaction tube and stored at -80 ^{o}C. One sample from each crude extract type was stored for 30, 60, and 90 days and thawed and denatured, as described in 3.2.3, on their allocated day. Once all samples were denatured and stored at -20 ^{o}C, they were analysed by western blotting (see section 2.2.6).

2.3 Results

2.3.1 Confirmation of Plant Expression Constructs

2.3.1.1 pTRAkc-ERH expression constructs

The CHIKV E2-HIS and CHIKV E2ΔTM-HIS gene variants were subcloned into pTRAkc-ERH. Confirmation of successful cloning and transformation was confirmed by colony PCR screening (Fig. 2.9), restriction digest analysis (Fig. 2.10) and sequencing (data not shown).

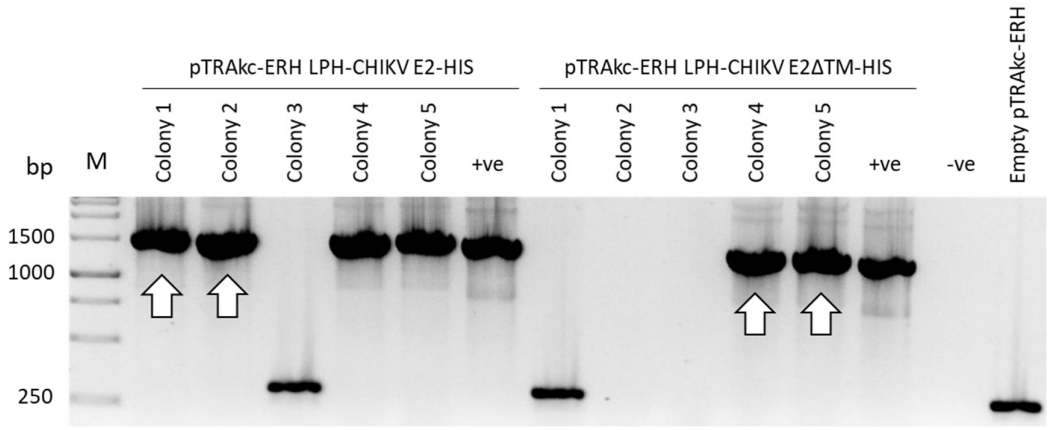

Figure 2.9: Colony PCR screening of recombinant pTRAkc-ERH *E. coli* transformants. Five colonies per construct were screened using pTRA-vector primers (Table 2.2). Expected PCR product sizes for pTRAkc-ERH LPH-CHIKV E2-HIS colonies were 1543 bp and for pTRAkc-ERH LPH-CHIKV E2ΔTM-HIS colonies 1366 bp. +ve, pUC57-CHIKV E2-HIS and pUC57-CHIKV E2ΔTM-HIS colony-PCR positive controls; -ve, no-template negative control; M, O'GeneRuler™ 1kb DNA Ladder. White arrows, constructs selected for restriction enzyme digest analysis.

Of the colonies screened for pTRAkc-ERH LPH-CHIKV E2-HIS, 4 out of 5 showed the expected fragment of 1543 bp. For pTRAkc-ERH LPH-CHIKV E2ΔTM-HIS, 2 out of 5 colonies showed the expected fragment of 1366 bp. DNA was extracted from two colonies per construct (white arrows, Fig. 2.9) for restriction enzyme digest analysis to confirm colony PCR screening (Fig. 2.10).

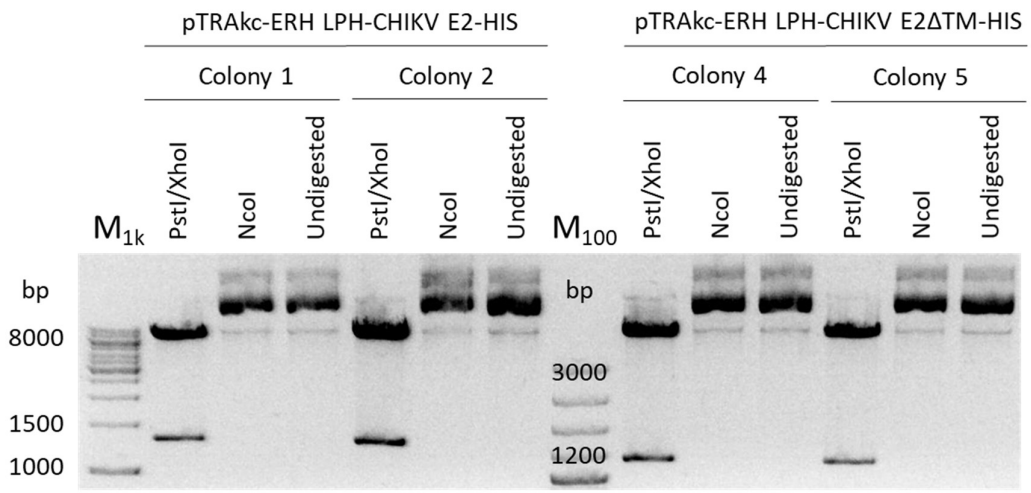

Figure 2.10: Restriction enzyme digest of recombinant pTRAkc-ERH constructs. DNA from two colonies per construct was double digested with PstI and XhoI or cut with NcoI. Undigested-vector reactions were used as a control. An expected band size for the double digest of pTRAkc-ERH LPH-CHIKV E2-HIS and pTRAkc-ERH LPH-CHIKV E2ΔTM-HIS are 1332 bp and 1155 bp, respectively, with a 7598 bp vector backbone. M_{1k}, O'GeneRuler™ 1kb DNA Ladder; M_{100}, O'GeneRuler™ 100 bp Plus DNA Ladder.

Constructs digested with PstI and XhoI (two restriction sites framing the inserted gene) resulted in DNA fragments of the expected sizes of 1332 bp and 1155 bp for pTRAkc-ERH LPH-CHIKV E2-HIS and pTRAkc-ERH LPH-CHIKV E2ΔTM-HIS, respectively. In addition, NcoI-digested and undigested constructs displayed identical banding patterns, indicating the abolishment of the vector's MCS NcoI site during the ligation of the pTRAkc-ERH NcoI site's compatible end with the gene variant's BspHI site's compatible end. Confirmed pTRAkc-ERH LPH-CHIKV E2-HIS and pTRAkc-ERH LPH-CHIKV E2ΔTM-HIS DNA were electroporated into *Agrobacterium* strain GV3101::pMP90RK and screened by colony PCR (Fig. 2.11).

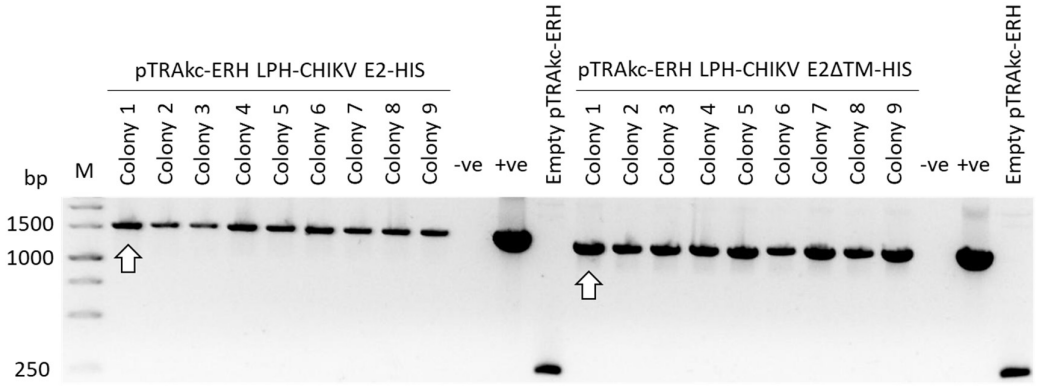

Figure 2.11: Colony PCR screening of recombinant pTRAkc-ERH *Agrobacterium* transformants. Nine colonies per electroporated construct were screened using pTRA-vector primers (Table 2.2). Expected PCR product sizes for pTRAkc-ERH LPH-CHIKV E2-HIS colonies are 1543 bp and for pTRAkc-ERH LPH-CHIKV E2ΔTM-HIS colonies are 1366 bp. +ve, positive transformed *E. coli* pTRAkc-ERH colonies; -ve, no-template negative control; M, O'GeneRuler™ 1kb DNA Ladder. White arrows, colonies sent for sequencing (data not shown) and used for downstream expression studies.

All recombinant pTRAkc-ERH LPH-CHIKV E2-HIS colonies yielded an amplified fragment of 1543 bp. Similarly, amplification of all recombinant pTRAkc-ERH LPH-CHIKV E2ΔTM-HIS colonies yielded an expected fragment of 1366 bp size.

The extracted pTRAkc-ERH LPH-CHIKV E2-HIS and pTRAkc-ERH LPH-CHIKV E2ΔTM-HIS constructs were sequenced and confirmed successful cloning of the CHIKV E2 gene variants in pTRAkc-ERH (data not shown). These constructs were used for downstream transient-expression time trials of CHIKV E2 protein variants.

2.3.1.2 pEAQ-HT Expression Constructs

2.3.1.2.1 Original Cloning Strategy Results

Both gene variants were ligated into pUC18-LPH, constructs were transformed into *E. coli* and clones were screened by colony PCR. Representative data is shown in Fig. 2.12. As can be seen, all but one of the screened colonies yielded bands identical in size to that of the pUC18-LPH vector lacking the insert indicating that ligation and/or transformation was unsuccessful. Colony 1 of the pUC18-LPH CHIKV E2ΔTM-HIS clones yielded a product

lower than the empty pUC18-LPH control. This may be due to non-specific binding of primers to supplied template DNA.

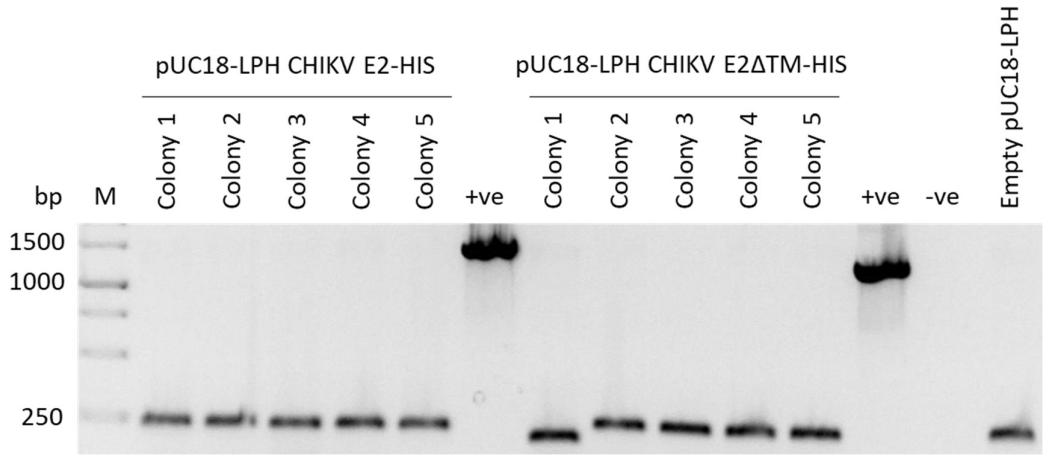

Figure 2.12: Colony PCR screening of recombinant pUC18-LPH *E. coli* transformants. Five colonies per transformed construct were screened using M13 primers (Table 2.2). Expected PCR product sizes for pUC18-LPH CHIKV E2-HIS colonies were 1512 bp and for pUC18-LPH CHIKV E2ΔTM-HIS colonies are 1335 bp. +ve, pUC57-CHIKV E2-HIS and pUC57-CHIKV E2ΔTM-HIS colony-PCR positive controls; -ve, no-template negative control; M, O'GeneRuler™ 1kb DNA Ladder.

2.3.1.2.1 New Cloning Strategy Results

The original pEAQ-*HT* strategy described in 2.2.3.4 was to generate the LPH-fused gene variants for subsequent subcloning into the plant expression vector pEAQ-*HT*; however, as this was unsuccessful, a new strategy was devised. LPH-CHIKV E2-HIS and LPH-CHIKV E2ΔTM-HIS were amplified and isolated, to yield correctly sized products indicative of successful isolation of LPH-fused CHIKV E2-HIS variants (data not shown), and cloned into an intermediary vector (pJET1.2/blunt) to generate pJET AgeI-LPH CHIKV E2-HIS-pTRA-R and pJET AgeI-LPH CHIKV E2ΔTM-HIS-pTRA-R

PCR screening of transformed *E. coli* colonies and sequencing (data not shown) of DNA prepared from the putative recombinants indicated successful transformation of pJET1.2 constructs. The 1473 bp and 1296 bp amplified products obtained from 4 out of 5 colonies screened represent amplified LPH-CHIKV E2-HIS and LPH-CHIKV E2ΔTM-HIS from the transformed pJET1.2 constructs using the AgeI-LPH-pTRA-F and pTRA-R primers (Fig. 2.13).

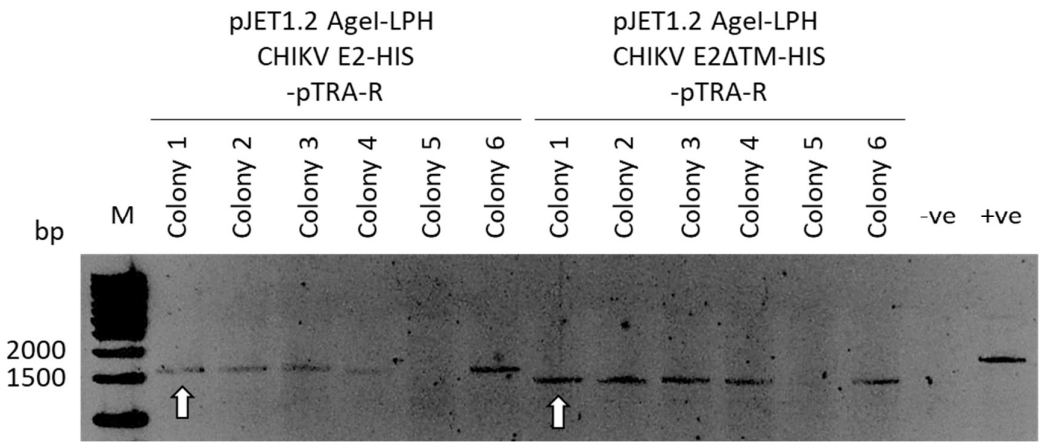

Figure 2.13: Colony PCR of pJET1.2 *E. coli* transformants. Six colonies per transformed construct were screened using AgeI-LPH-F and pTRA-R primers (Table 2.2). Expected PCR product sizes for pJET1.2 AgeI-LPH CHIKV E2-HIS-pTRA-R colonies are 1473 and for pJET1.2-AgeI-LPH CHIKV E2ΔTM-HIS-pTRA-R colonies are 1296 bp. -ve, no-template negative control; +ve, pJET1.2 West Nile Virus (WNV) AgeI-LPH-Envelope (Env)-pTRA-R positive control (~ 1.8 kb); M, O'GeneRuler™ 1kb DNA Ladder. White arrows, colonies from which DNA was used for sequencing.

As mentioned in section 2.2.3.5, a sample of the AgeI- and XhoI-digested pEAQ-*HT* vector was resolved by agarose gel electrophoresis to ensure apt digestion of the vector before performing ligations of LPH-CHIKV E2-HIS gene variants (Fig. 2.14A). Single digest control reactions of the pEAQ-*HT* vector showed that the respective restriction enzymes cut correctly. Therefore, since the single digests cut correctly in the same buffer used for the double digest, it can be inferred that the respective enzymes cut pEAQ-*HT* correctly during the double digest i.e. the digestion of pEAQ-*HT* with AgeI and XhoI in preparation for ligations was performed correctly.

Samples of the subsequent ligation reactions were also resolved by agarose gel electrophoresis to ensure LPH-CHIKV E2-HIS gene variants successfully ligated into phosphorylated and/or dephosphorylated variants of the digested pEAQ-*HT* vector (Fig. 2.14B) before performing transformations. The 1415 bp size bands in lanes 2 and 4 indicated that only the LPH-CHIKV E2ΔTM-HIS gene variant ligated into digested pEAQ-*HT* vectors (dephosphorylated and non-dephosphorylated). However, an unknown ~800 bp size band was amplified in these reactions. In addition, PCR screening predominantly amplified fragments identical to those seen in the empty vector control (E) indicating a

higher number of unsuccessful ligations present compared to the number of successful ligations present in the ligation reactions. Transformations of these ligation reactions were performed despite these results.

Colony PCR screening of *E. coli* transformed with pEAQ-*HT* constructs identified no positive clones (Fig. 2.14C). PCR screening of the clones amplified fragments that are identical to pEAQ-*HT* constructs lacking the ligated LPH-CHIKV E2-HIS gene variants of interest ("Empty pEAQ-*HT*"). Despite multiple attempts, the same result was seen (data not shown). Therefore, pEAQ-*HT* LPH-CHIKV E2-HIS and pEAQ-*HT* LPH-CHIKV E2ΔTM-HIS cloning was discontinued.

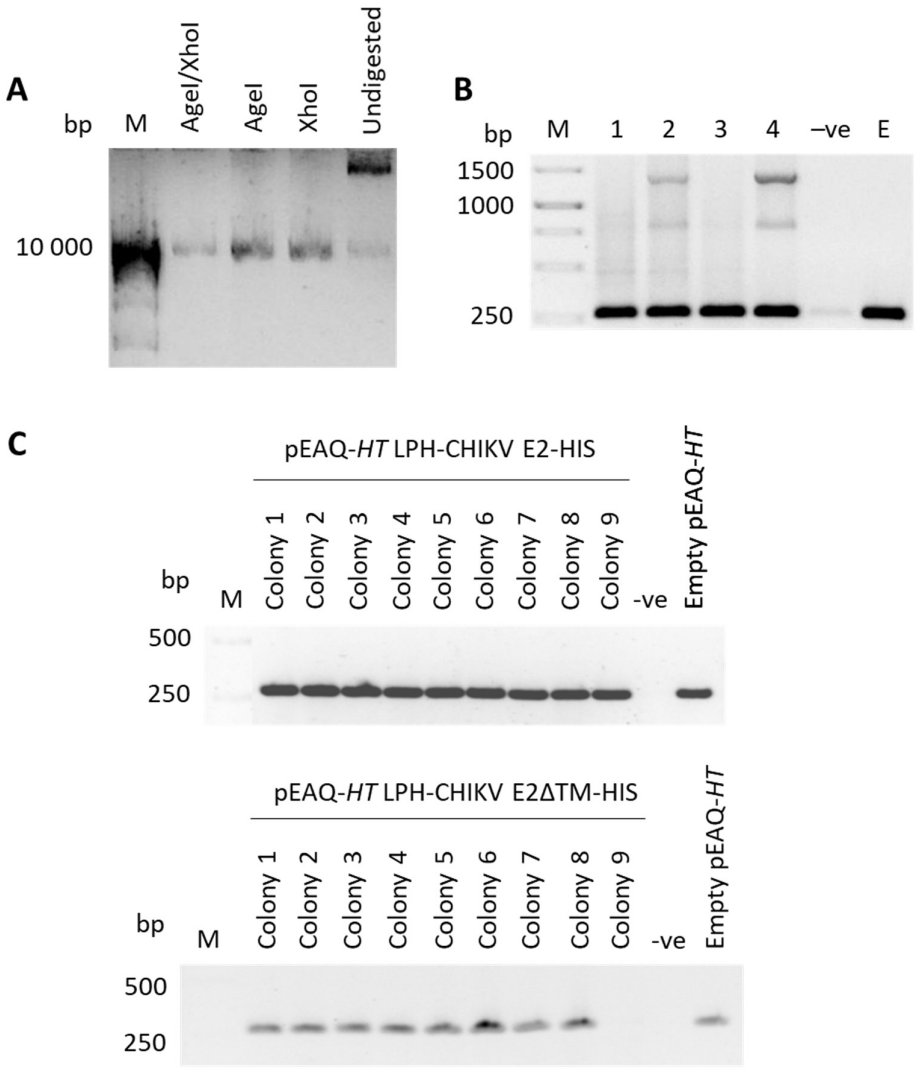

Figure 2.14: pEAQ-*HT* cloning analyses. (A) Restriction enzyme digest of empty pEAQ-*HT*. In preparation for ligations, pEAQ-*HT* vector digests were analysed to confirm correct double and single digests with AgeI and XhoI. Undigested pEAQ-*HT* was added as a control. **(B) PCR screening of ligation reactions.** Ligation efficiency was analysed for the insertion of gene variants into either dephosphorylated pEAQ-*HT* or non-dephosphorylated pEAQ-*HT* vector digested with AgeI and XhoI. 1 µL of ligation reaction mix was analysed using pEAQ-vector primers (Table 2.2). 1, LPH-CHIKV E2-HIS → non-dephosphorylated pEAQ-*HT* vector; 2, LPH-CHIKV E2ΔTM-HIS → non-dephosphorylated pEAQ-*HT* vector; 3, LPH-CHIKV E-HIS → dephosphorylated pEAQ-*HT* vector; 4, LPH-CHIKV E2ΔTM-HIS → dephosphorylated pEAQ-*HT* vector. E, empty pEAQ-*HT* control. **(C) Colony PCR screening of pEAQ-*HT* *E. coli* transformants.** Nine colonies per construct were screened using pEAQ-vector primers (Table 2.2). Expected PCR product sizes for pEAQ-*HT* LPH-CHIKV E2-HIS were 1592 bp and for pEAQ-*HT* LPH CHIKV E2ΔTM-HIS are 1415 bp. -ve, no-template negative control; M, O'GeneRuler™ 1kb DNA Ladder.

2.3.2 Small-Scale Protein-Expression Time Trials

A small-scale syringe-infiltration time trial was performed over 7 days (see section 2.2.3) to determine which of the two CHIKV E2 variants yielded the highest levels of protein, on which day post-infiltration (dpi) and at what optical density (OD_{600} = 0.5 or 1) tested. Recombinant cultures harbouring pTRAkc-ERH lacking any insert or pTRAkc-Shuni N-HIS (OD_{600} = 0.5) were used as negative and positive controls, respectively (Fig. 2.15).

Western blot analyses of equal TSP showed recombinant plant-produced CHIKV E2-HIS to be ~52 kDa in size and CHIKV E2ΔTM-HIS to be ~49 kDa in size. Low to no detectable levels of recombinant protein were seen for both CHIKV E2-HIS and CHIKV E2ΔTM-HIS on all dpi for both OD_{600} tested. For CHIKV E2-HIS, at OD_{600} = 1.0, expression could be seen (red arrow) at 3 dpi only; however, it was higher than expression levels observed when using an OD_{600} = 0.5 on the same dpi. Minimal expression was observed for OD = 0.5 at 5 dpi and no expression was detected on any other day. For CHIKV E2ΔTM-HIS, low levels of expression were seen on 3 dpi only for OD_{600} = 1.0; however, the level of expression was seemingly higher than that of CHIKV E2-HIS's for all dpi and both OD_{600}.

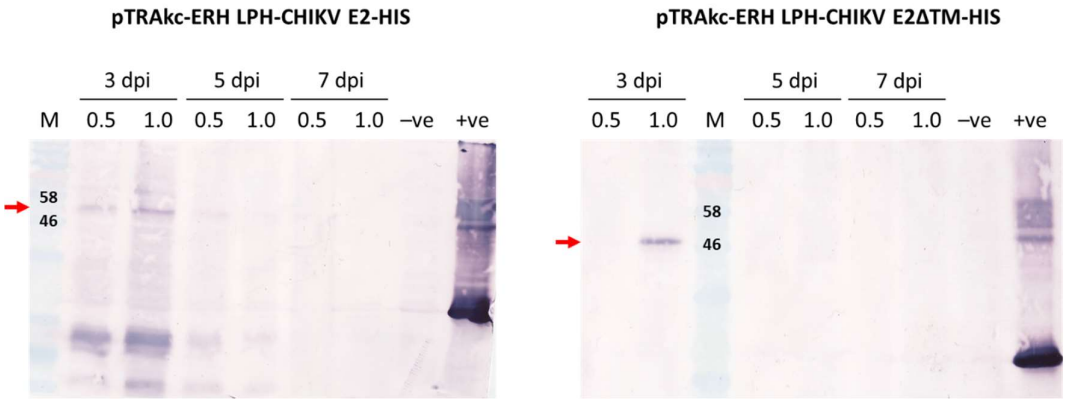

Figure 2.15: Equal TSP western blot screening of pTRAkc-ERH LPH-CHIKV E2-HIS (left) and pTRAkc-ERH LPH-CHIKV E2ΔTM-HIS (right) expression. Expression time trial of CHIKV E2 gene variants at OD_{600} 1.0 and 0.5 over a 7-day period. Nitrocellulose membrane was probed with anti-HIS antibodies. M, MW marker; red arrow, expected size of CHIKV E2ΔTM-HIS; -ve, empty pTRAkc-ERH negative control; +ve, Shuni N-HIS positive anti-HIS control.

Thereafter, the CHIKV E2 gene variants were co-expressed with either CRT (OD_{600} = 0.5), CNX (OD_{600} = 0.5), NSs (OD_{600} = 0.5) to determine whether this increased protein expression.

Co-expression of pTRAkc-ERH LPH-CHIKV E2-HIS with either CRT, NSs or CNX showed higher levels of expression on select dpi than that without CRT, NSs or CNX and for a longer period post-infiltration (Fig 2.16). CRT significantly improves expression for both ODs on 5 dpi with little to no detectable improvement of expression outside of this dpi for both ODs. Similarly, the other human chaperone, CNX, improves expression slightly on 5 dpi for OD_{600} 0.5 and on 7 dpi for OD_{600} 1.0 only. Furthermore, recombinant CHIKV E2-HIS expression was significantly better improved by NSs very slightly on 3dpi and higher on 5 and 7 dpi for both ODs. However, these improved levels of expression for CHIKV E2-HIS are still low.

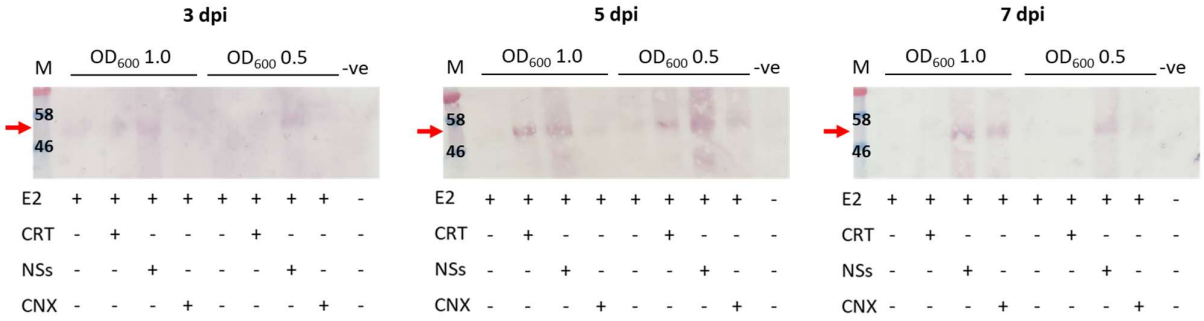

Figure 2.16: Equal TSP western blot screening of pTRAkc-ERH LPH-CHIKV E2-HIS. Co-expression time trial of CHIKV E2-HIS (OD_{600} = 1.0 and 0.5) with CRT/NSs/CNX (OD_{600} = 0.5) over a 7-day period. Nitrocellulose membrane was probed with anti-HIS antibodies. M, MW marker; red arrow, expected size of CHIKV E2ΔTM-HIS; -ve, empty pTRAkc-ERH negative control.

Co-expression of pTRAkc-ERH LPH-CHIKV E2ΔTM-HIS with either CRT, NSs or CNX showed higher levels of expression on select dpi than expression without CRT, NSs or CNX and for a longer period post-infiltration. Co-expression with CRT increased expression of E2ΔTM from being barely visible to a ~ 49 kDa band that is very visible across 3 to 7 dpi

using cultures of OD_{600} = 0.5 and 1. NSs showed lower expression improvement than CRT, but improvement nevertheless, on 3 dpi for both ODs tested, 5 and 7 dpi for OD_{600} = 0.5 as well as 7 dpi for OD_{600} = 1.0. Interestingly, improvement of expression by NSs was better and more prolonged for OD_{600} = 0.5, compared to OD_{600} = 1.0, as seen for all dpi. Furthermore, CNX improved expression on all dpi for OD_{600} 1.0 and on 3 and 5 dpi only for OD_{600} = 0.5 (Fig. 2.17); however, the improved expression levels were lowest compared to improvement by the other co-expressed proteins. Co-expression of pTRAkc-ERH LPH-CHIKV E2ΔTM-HIS (OD_{600} = 1.0) with CRT showed the highest levels of expression for any of the conditions investigated, compared also to the co-expression time trials of CHIKV E2-HIS. Consequently, these were the conditions chosen scaling up expression of the protein.

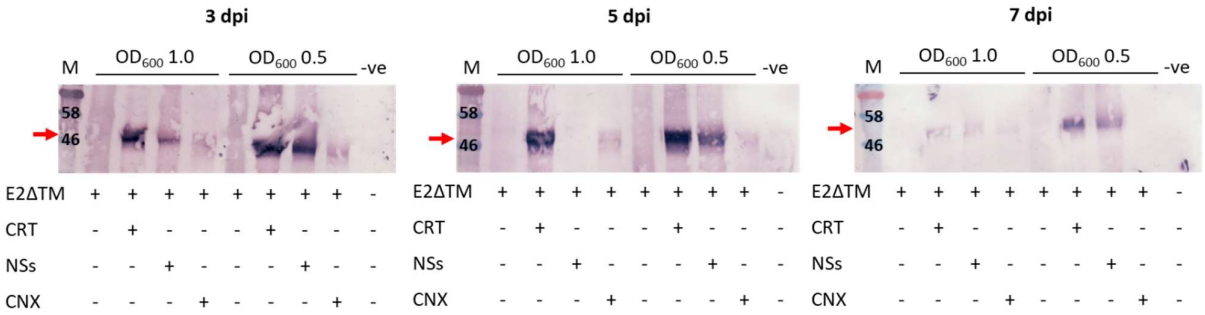

Figure 2.17: Equal TSP western blot screening of pTRAkc-ERH LPH-CHIKV E2ΔTM-HIS. Co-expression time trial of CHIKV E2ΔTM-HIS (OD_{600} = 1.0 and 0.5) with CRT/NSs/CNX (OD_{600} = 0.5) over a 7-day period. Nitrocellulose membrane was probed with anti-HIS antibodies. M, MW marker; red arrow, expected size of CHIKV E2ΔTM-HIS; -ve, empty pTRAkc-ERH negative control.

2.3.3 Large-Scale Vacuum Infiltration

A scaled-up co-infiltration time-trial repeat of pTRAkc-ERH LPH-CHIKV E2ΔTM-HIS (OD_{600} = 1.0) and pEAQ-*HT* CRT (OD_{600} = 0.5) was performed on 15 plants (5 plants per dpi). Western blot screening of crude extract equal TSP indicated the highest expression levels of CHIKV E2ΔTM-HIS on 3 dpi (Fig. 2.18 red arrow); however, a faint, similarly-sized signal was detected in the crude extract of plants infiltrated with CRT only (Fig.2.18, "CRT only" lane) indicating possible non-specific probing. Additionally, Coomassie staining of the same samples showed the resolution of a similarly-sized protein to that of CHIKV E2ΔTM-HIS (~53 kDa) in all samples ('CRT only' included) (Fig. 2.18 yellow arrows). This protein signal of similar size was worth investigating as it may indicate that the signal we saw

from probing CHIKV E2ΔTM-HIS might have been non-specific probing of the anti-6xHis antibodies to either the similar-sized large RuBisCo subunit (~53 kDa), CRT (~49 kDa) or both, or spill-over from the 7 dpi samples.

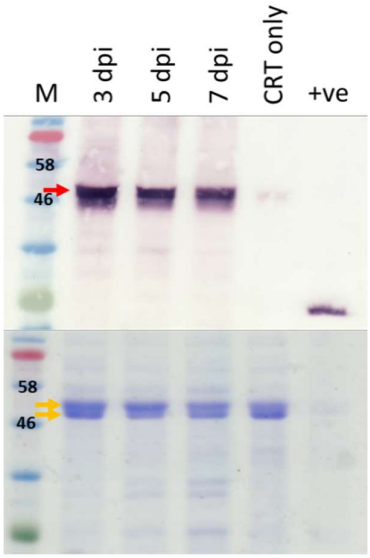

Figure 2.18: Equal TSP western blot screening (top) and Coomassie staining (bottom) of pTRAkc-ERH LPH-CHIKV E2ΔTM-HIS. Repeat of the co-expression time trial, on a large-scale, of CHIKV E2ΔTM-HIS (OD_{600} = 1.0) with CRT (OD_{600} = 0.5) over a 7-day period to select optimal dpi for large-scale harvesting. Nitrocellulose membrane was probed with anti-HIS antibodies. M, MW marker; red arrow, expected size of CHIKV E2ΔTM-HIS; yellow arrows, CHIKV E2ΔTM-HIS/CRT/large subunit of RuBisCo; +ve, Shuni N-HIS positive anti-HIS control.

Coomassie staining of 3 dpi crude extract samples from large-scale co-expression of CHIKV E2ΔTM-HIS and CRT expression only and empty pTRAkc-ERH show a set of dominantly expressed proteins resolving between the ¬50 kDa and ¬55 kDa marks (Fig. 2.19). Western blot screening of the same samples indicated no background signal as a result of non-specific binding of anti-HIS antibodies as seen in lane “2” where both CRT and the large subunit of RuBisCo is present (as seen by the αRuBisCo and αCRT blots). This may indicate that the signal seen in the “CRT only” lane in Fig. 2.18 may have been accidental spill-over from lane “7 dpi”. Interestingly, comparing the top-most prominent band, in the range observed above, between lane “1” and “2”, CRT levels seem to be

higher at the 3-dpi time point when CHIKV E2ΔTM-HIS is not present ("2"). This may illustrate the 3 dpi-expression levels of CRT when not actively facilitating the folding of CHIKV E2ΔTM-HIS.

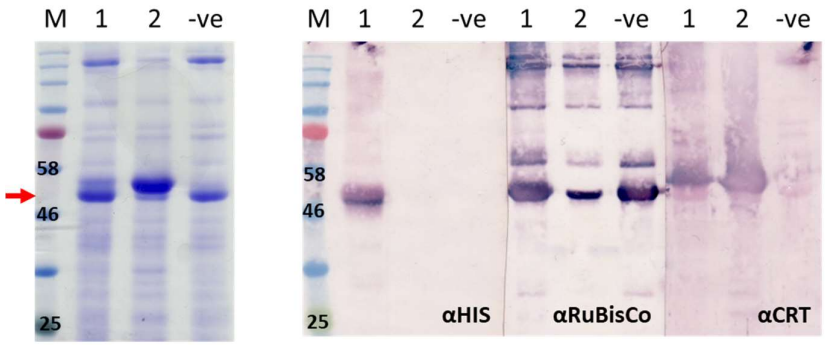

Figure 2.19: Coomassie staining (left) and equal TSP western blot screening (right) of CHIKV E2ΔTM-HIS, CRT and RuBisCo. Repeat of large-scale co-expression of CHIKV E2ΔTM-HIS (OD_{600} = 1.0, 3 dpi) with CRT (OD_{600} = 0.5) to screen for relative CHIKV E2ΔTM-HIS, CRT and RuBisCo expression levels and sizes. Identical SDS-PAGE gels were transferred to nitrocellulose membranes probed with anti-HIS, anti-CRT and anti-RuBisCo antibodies. M, MW marker; red arrow, expected size of CHIKV E2ΔTM-HIS; 1, pTRAkc-ERH CHIKV E2ΔTM-HIS and CRT co-expression; 2, CRT expression only; -ve, empty pTRAkc-ERH negative control; α-, anti-protein antibody.

2.3.4 Stability of CHIKV E2ΔTM-HIS

Since large-scale co-expression of CHIKV E2ΔTM-HIS and CRT was optimised (Fig. 2.15, Fig. 2.17 – 2.19), it was worth investigating the stability of the plant-produced CHIKV E2ΔTM-HIS at -80 °C, in the presence or absence of PI in the extraction buffer, over a 90 day post-extraction (dpe) period, to explore whether and for how long recombinant CHIKV E2ΔTM-HIS proteins would be stable at -80 °C for the utilization of these recombinant proteins in a diagnostic kit. The results indicated the presence of stable CHIKV E2ΔTM-HIS proteins on 30, 60 and 90 dpe in samples extracted with and without PI (Fig. 2.20); however, increasing levels of CHIKV E2ΔTM-HIS degradation was observed (black arrows) as time increased for samples extracted without PI. Similarly, small levels of degradation can be seen for CHIKV E2ΔTM-HIS extracted with PI starting after 60 dpe and before 90 dpe; however, these degradation levels are significantly lower than those seen after 60 dpe and before 90 dpe for samples extracted without PI.

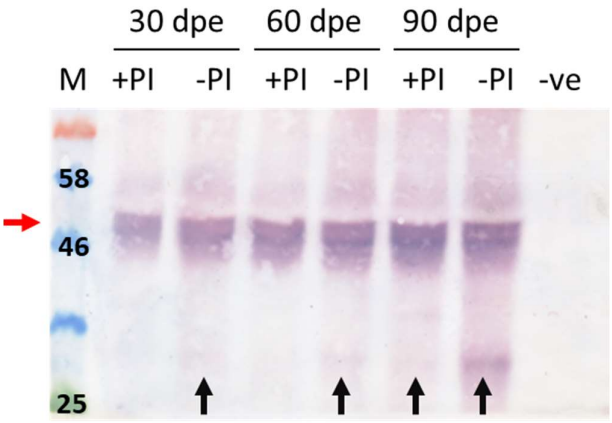

Figure 2.20: Western blot screening of pTRAkc-ERH CHIKV E2ΔTM-HIS. Stability study of large-scale extracted CHIKV E2ΔTM-HIS in the presence and absence of protease inhibitor over a 90-day period. Nitrocellulose membrane was probed with anti-HIS antibodies. M, MW marker; red arrow, expected size of CHIKV E2ΔTM-HIS; black arrow, potential degradation of CHIKV E2ΔTM-HIS; -ve, empty pTRAkc-ERH negative control.

2.4 Discussion

In this study, a consensus sequence of the CHIKV E2 gene generated from three South African CHIKV isolates was designed with and without its transmembrane domain (ΔTM) with each gene variant containing a 6xHis-tag on their 3' termini. The consensus sequence was generated by using the only three available South African isolates mentioned in section 2.2.1. The ORFs of the South African isolates and 77 other isolates found around the world are highly conserved; as such, the generated consensus sequence represents an appropriate sequence for the diagnostic assay. These constructs were codon optimised for optimal expression in the *Nicotiana benthamiana* plant-expression host.

The 6xHis-tag was chosen for its size, the availability of commercial anti-HIS antibodies, relative cost compared to other affinity tags, one-step downstream purification process, and its low transcriptional/translational burden on the tobacco plant-expression host due its small 6 amino-acid size. The plant host is likely to be burdened already as a result of transient expression of recombinant proteins of interest via agroinfiltration (Hitzeroth and van Zyl, 2016). In addition, at the time this study was being conducted, commercial anti-CHIKV E2 antibodies were not readily available. But, as seen by the efficient

detection of the 6xHis-tagged CHIKV E2 protein variants, the anti-HIS antibody substitute used in this study was a sufficient alternative. Previous studies have shown the efficiency of using Ni-NTA chromatography to purify recombinant 6xHIS-tagged CHIKV E2 glycoproteins in insect (Cho et al., 2008, Bagno et al., 2020), *E. coli* (Verma et al., 2016, Bagno et al., 2020) and mammalian cells (Bagno et al., 2020). These studies present further evidence for how the 6xHis-tag and Ni-NTA chromatography provide a universal purification system of recombinant antibodies: if recombinant proteins are insoluble, located in inclusion bodies and/or have a tertiary structure such that the 6xHis-tag is hidden within, proteins can be extracted under denaturing conditions and subsequently purified successfully (Crowe et al., 1994).

The difficulty experienced with the pEAQ-HT cloning may be due to the low amounts of correctly-ligated constructs available for transformation and/or the large size of the recombinant vector constructs containing the >1 kb gene variant constructs. It is known that transformation efficiency decreases with an increase in construct size (Sheng et al., 1995). The large sizes of genes to be ligated may also account for why insertion into the non-dephosphorylated pEAQ-*HT* vector was unsuccessful. Dephosphorylation of vector DNA reduces empty vector background; however, this was not the case as all ligation reactions contained similar amounts of empty vector DNA. A combination of these factors may be why transforming a construct of >11 kb size into competent cells not necessarily designed for this (unlike Maximum-Efficiency DH5α *E. coli* cells) is failing. Transformation of pEAQ-*HT* constructs into Maximum-Efficiency DH5α *E. coli* cells, or the like, was not performed due to time and cost limitations.

Cloning of CHIKV E2-HIS and CHIKV E2ΔTM-HIS into pTRAkc-ERH however, was successful in generating pTRAkc-ERH LPH-CHIKV E2-HIS and pTRAkc-ERH LPH-CHIKV E2ΔTM-HIS. Three small-scale syringe-infiltration time trials (data for one shown) of these constructs indicated that the highest yielding CHIKV E2 variant was CHIKV E2ΔTM-HIS (Fig. 2.15, 2.16). Moreover, western blot analysis of crude extracts from co-expression with either CRT, CNX or NSs indicated that the optimal conditions for CHIKV E2ΔTM-HIS expression were infiltration at OD_{600} = 1.0 and co-expression with CRT (Fig. 2.17). Using the optimal conditions for large-scale vacuum-infiltration time trials allowed for (a) replication of the small-scale infiltration in a large-scale setting and (b) determination of the best dpi to be

3 dpi (Fig. 2.18). This is expected as transient protein expression in tobacco plants often provides optimal expression within a week; hence sampling leaf clippings at 3, 5 and 7 dpi. Optimal expression does differ from protein to protein and should be repetitively determined on a case-by-case basis: rotavirus capsid protein VP6 (Pera et al., 2015), beak and feather disease virus (BFDV) coat protein (Duvenage et al., 2013) and HPV-16 L1 protein (Maclean et al., 2007) were optimally expressed at 3, 4 and 6 dpi, respectively. A similarly-sized signal was seen in the control lanes of Fig. 2.18 (CRT expression only) - this is troubling because: (a) as previously mentioned, anti-HIS antibodies are notoriously promiscuous and, in this case, are potentially binding non-specifically to a protein(s) of similar size; (b) CRT (~49 kDa), CHIKV E2ΔTM-HIS (~52 kDa) and the large subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCo) (~53 kDa) are similar in size and, therefore, will resolve similarly by electrophoresis. Therefore, whatever the anti-HIS antibodies are potentially non-specifically binding to may be contributing to the signal we see for CHIKV E2ΔTM-HIS and thus inflating the observed expression levels assumed to only be indicative of CHIKV E2ΔTM-HIS expression levels. Probing of 3 dpi crude extracts indicated this not to be the case (Fig. 2.19) and the signal seen is potentially spill over from the '7 dpi' lanes (Fig. 2.18).

Furthermore, a stability study of the same samples as above over a 90-day period was performed. While the stability of a protein does not indicate stable biological function, it is worth investigating stability as unstable/degraded proteins will not exhibit any biological activity. When the biological activity of a protein is assayed, it is worth investigating the ideal time of stability to obtain accurate results. The stability study indicated stable recombinant CHIKV E2ΔTM-HIS proteins present for at least 90 dpe at -80 ^{o}C. However, increasing levels of degradation were seen for -PI samples over time (Fig. 2.20). Therefore, while CHIKV E2ΔTM-HIS can be extracted without PI, thus decreasing the cost of DSP, samples need to be processed within a month of extraction for optimal DSP yields.

The small-scale agroinfiltration time-trials and large-scale vacuum agroinfiltrations indicated that the transiently expressed CHIKV E2ΔTM-HIS glycoprotein variant, when co-infiltrated with human chaperone CRT, is the best expressed variant out of the two tested. While this may be so, investigations into the CHIKV E2ΔTM-HIS purification yield

must be done to assess and optimisation the most appropriate purification method to be used to achieve the highest yield of CHIKV E2ΔTM-HIS glycoproteins for use in the diagnostic assay.

Chapter 3: Purification of CHIKV E2ΔTM-HIS Glycoprotein Variant

3.1 Introduction

This chapter describes different techniques used to purify recombinant plant-produced CHIKV E2ΔTM-HIS proteins from plant leaf extracts. Downstream processing (DSP), of which purification is a component, is the process physical and/or chemical, used to extract recombinantly produced proteins from their starting materials. The cost of DSP often accounts for up to 80% of the overall biopharmaceutical productions costs (Tschofen et al., 2016). Purification of biopharmaceuticals is a critical part of the DSP process and must be reliable, robust, easily scaled-up and able to remove impurities specific to the protein expression system (Jozala et al., 2016). One such introduced process-related impurity is the superabundant plant protein RuBisCo – which may require early processing procedures to ensure removal (Tschofen et al., 2016). This may be achieved through a partial purification strategy that could potentially separate these process-related impurities from CHIKV E2ΔTM-HIS and minor plant proteins. These other minor proteins can then be separated from CHIKV E2ΔTM-HIS by a subsequent purification strategy that may no longer be hindered by the removed impurities.

The nickel-affinity purification system is an efficient tool for one-step purification of recombinantly produced proteins. The system's efficiency is credited to the tight-binding ability of the fused 6xHis-tag to the immobilised Ni-charged affinity resin (Ni-NTA affinity chromatography). As such, it allows for contaminants which are not fused to 6xHis-tags and therefore do not bind to Ni-NTA to be washed out under stringent conditions before eluting the protein of interest by competition with imidazole, or a slight altering of pH (Crowe et al., 1994). Nickel-affinity purification is a routinely-used system that also offers mild-elution conditions – both ideal features for large-scale purification - and the ability to purify proteins under denaturing conditions using 8 M urea (Gaberc-Porekar and Menart, 2001). One disadvantage that may arise during nickel-affinity purification is the lack of binding of a 6xHis-tagged recombinant protein with an inaccessible 6xHis-tag. This may lead to low or no binding of the tagged protein to the immobilised resin (Block et al., 2009).

Running a non-denaturing polyacrylamide gel is an informative experiment to perform to assess the accessibility of a 6xHis-tag on a recombinant protein. In contrast to

denaturing polyacrylamide gel, a non-denaturing gel does not contain the detergent SDS and preparation of samples are done without β-mercaptoethanol, SDS and heat-denaturing conditions (Arndt et al., 2012). This allows for the resolution of the protein of interest in its native conformation i.e. the conformation of the protein of interest when purified by Ni-NTA affinity chromatography. As such, the exposure of the 6xHis-tag fused to the native recombinant protein can be assessed by probing the transferred proteins with anti-6xHis antibodies on a nitrocellulose membrane. This should be done in parallel with a denatured sample of the same protein of interest to compare the probing signal of the denatured and non-denatured samples as a non-exposed 6xHis-tag, when probed with anti-6xHis antibodies, will not elicit an BCIP-NBT-facilitated colour reaction as intense as the denatured sample.

Failing efficient one-step purification using Ni-NTA affinity chromatography, pH purification and ammonium sulphate ($(NH_4)_2SO_4$) precipitation are useful protocols for the partial purification of soluble proteins. pH precipitation relies on the fact that each protein has their own isoelectric point (pI) and is least soluble at this pI i.e. protein-protein interactions become greater as a result of the electrostatic forces of the molecules being at their minimum and less water interacting with these proteins. These conditions are favourable for aggregation of proteins and possibly precipitation (Pelegrine and Gasparetto, 2005). Therefore, altering the pH conditions of crude extract from infiltrated and harvested leaf material may precipitate the protein of interest at one of the pH conditions tested. Contrastingly, altering the pH conditions of crude extract may precipitate out a majority of the other plant proteins at one of the pH conditions tested and leave the protein of interest in the supernatant.

The method of ammonium sulphate precipitation is a convenient first step for separation of recombinant proteins of interest from other proteins in crude homogenate due to its low cost and selectivity benefits compared to other separation methods like ion-exchange and gel filtration chromatographies (King, 1972). Ammonium sulphate precipitation makes use of the fact that increasing ionic strength reduces the solubility of soluble proteins in a process called 'salting-out'. Thus, based on a protein's solubility in the presence of varying salt ((NH4)2SO4) concentrations, the salting-out process can be used to separate proteins for partial purification (Duong-Ly and Gabelli, 2014).

(NH4)2SO4 is the salt of choice for initial separation of crude extracts since it yields precipitated proteins that are both denatured and concentrated in a small volume. However, contaminant proteins may be co-precipitated with the protein of interest. As such, if very pure preparations are needed, a further precipitation step, like ion-exchange chromatography, may be needed thereafter (Grodzki and Berenstein, 2010).

The purification strategies used and described in this chapter include Ni-NTA affinity chromatography, pH precipitation, and ammonium sulphate precipitation. In addition, the exposure of the 6xHis-tag fused to the recombinant protein of interest was investigated to determine if a Ni-resin purification strategy would be effective. Thereafter, partial purification of CHIKV E2ΔTM-HIS proteins was investigated using pH and ammonium sulphate purification strategies. In addition, the presence of plant-protein contaminants, like recombinant calreticulin (CRT) and RuBisCo, were investigated for in the ammonium sulphate fractions that precipitated CHIKV E2ΔTM-HIS to determine whether they co-purified with CHIKV E2ΔTM-HIS.

3.2 Materials and Methods

3.2.1 Nickel-Resin Purification

Fifteen tobacco plants were co-infiltrated, as described in section 2.2.7, and soluble protein was extracted from harvested leaves, as described in section 2.2.8. Variable buffer reagents, buffer pH, imidazole concentrations, nickel resin volumes added, and binding times were tested for subsequent purification attempts. Despite the variable conditions mentioned above, the small- and large-scale methodology remained the same. A summary of these variable conditions can be found in Table 3.1.

All steps were performed at room temperature and centrifugation done at 5000 x g for 30 seconds. Variable volumes of HIS-Select® Nickel Affinity Gel (Sigma-Aldrich) were added to a 1.5 mL Eppendorf reaction tube and centrifuged and the supernatant subsequently discarded. Two hundred microlitres of equilibration buffer (components mentioned in Table 3.1) was added, mixed with the Ni-NTA and subsequently centrifuged. The supernatant was discarded and 100 µL of crude extract from harvested leaves was added to the equilibrated resin, mixed and incubated to allow proteins to bind

for variable amounts of time. The mixture was centrifuged and the supernatant stored and labelled “S1”. Thereafter, the mixture was washed with 500 µL of wash buffer (components mentioned in Table 3.1), mixed for 10 seconds and centrifuged. The supernatant was stored and labelled “S2”. The wash was repeated once more to obtain “S3”. Following this, 50 µL of elution buffer (components mentioned in Table 3.1) was added to the mixture, mixed well and centrifuged and the supernatant stored and labelled “E1”. This was repeated twice more, and the supernatants of each step were collected and labelled “E2” and “E3”. All fractions were denatured and analysed by SDS-PAGE and Coomassie staining as described in section 2.2.5.

When performed under denaturing conditions, 1 x PBS supplemented with 8M urea was used to extract recombinant proteins from harvested leaves. Likewise, the equilibration, wash and elution buffers used were supplemented with 8M urea. The elution buffer also lacked imidazole.

3.2.2 Denaturing vs. Non-Denaturing Polyacrylamide Gels

To investigate whether the 6xHis-tag is sufficiently exposed on the native conformation of the CHIKV E2ΔTM-HIS protein, the recombinant proteins were resolved by non-denaturing PAGE and in parallel, for comparison, by denaturing PAGE.

In order to maintain the native conformation of the CHIKV E2ΔTM-HIS proteins, the non-denaturing polyacrylamide gel was made without 10% sodium dodecyl sulphate (SDS), to prevent irreversible linearization of the recombinant proteins. Additionally, the 5 x sample application buffer (SAB) used lacked β-mercaptoethanol, in order to retain disulphide bonds as well as SDS, and the samples were not heat denatured. In contrast, denaturing gels (SDS-PAGE) and treatment of CHIKV E2ΔTM-HIS proteins were carried out as described in section 2.2.5.

The non-denaturing and denaturing gels were transferred onto the same sheet of nitrocellulose membrane adjacent to each other and screened by Western blotting, as described in 2.2.6, by probing with anti-6xHis antibodies.

3.2.3 pH Precipitation

Crude extract was extracted from harvested leaves, as described in section 2.2.8, in a 1 x PBS, pH 7.4 buffer. The pH of the crude extract was subsequently decreased to reach pH levels of 6.0, 5.5, 5.0, 4.5, 4.0 and 3.5 and 1 mL of sample from each pH level was transferred to Eppendorf reaction tubes. The samples were centrifuged at 13 000 rpm for 5 mins and the supernatants transferred to new reaction tubes. The pellets were resuspended in 100 µL of 1 x PBS, pH 7.4 to resolubilise any proteins that pelleted at the respective pH levels. The respective supernatants and pellets were denatured and analysed by SDS-PAGE and Coomassie staining, as described in section 2.2.5. Since samples were taken from the same crude extract at different pH levels, equal total soluble protein was compared between samples.

3.3.4 Ammonium Sulphate Precipitation

Proteins previously extracted, as described in 2.2.8, were fractionated by ammonium sulphate precipitation by the addition of solid $(NH_4)_2SO_4$ to achieve 80% saturation. Each saturation step was done in 10% increments. The amount of $(NH_4)_2SO_4$ to be added to increase saturation by 10% was calculated using an online ammonium sulphate calculator (http://www.encorbio.com/protocols/AM-SO4.htm).

Initially, to achieve 10% saturation, solid $(NH_4)_2SO_4$ was slowly added to freshly extracted proteins and left to equilibrate for 1 hr. This was done with constant stirring at 4 ^{o}C. Once equilibrated, the fraction was centrifuged at 10 000 x g for 15 min at 4 ^{o}C. The pellet was resuspended in 1.5 mL of 1 x PBS, pH 7.4, and stored at 4 ^{o}C for later dialysis of all resuspended samples. The calculated amount of solid $(NH_4)_2SO_4$ was slowly added to the supernatant at 4 ^{o}C with constant stirring to perform the subsequent 20% saturation step in the precipitation and the above process repeated. This was repeated for the 30%, 40%, 50%, 60%, 70%, and 80% saturation intervals.

Each sample of resuspended pellet was dialysed to remove the $(NH_4)_2SO_4$ for subsequent analysis of samples. Samples were dialysed O/N using tubing with a 10 kDa molecular-weight cut-off (Thermo Fisher Scientific) in 1 x PBS, pH 7.4 at 4 ^{o}C with constant stirring. Dialysed pellet resuspensions were denatured and analysed by SDS-PAGE, as in section

2.2.5, and probed using Western blotting, as in section 2.2.6. Samples from the dialysed ammonium sulphate fractions in which CHIKV E2ΔTM-HIS were precipitated were also screened for the presence of CRT (as recombinant CRT was co-expressed with CHIKV E2ΔTM-HIS) to investigate whether CRT co-purified with CHIKV E2ΔTM-HIS by probing samples with anti-CRT antibodies (abcam, Cambridge, United Kingdom).

3.3.5 Assessment of CHIKV E2ΔTM-HIS Purification Yields

To quantify and compare the purification yield of CHIKV E2ΔTM-HIS in the respective samples obtained from the purification strategies investigated, the intensities of the ~49 kDa bands in Western blots and Coomassie stains were observed.

With regards to Table 3.1, the intensities of the bands observed in the wash (S) and elution (E) samples are represented by a '+', where more '+' indicate higher intensities of the ~49 kDa band that represents the purification yield of CHIKV E2ΔTM-HIS in that sample ('+' indicates the band is just visible and '+++++' indicates the band is at its highest intensity i.e. dark blue or dark purple in the Coomassies and Western blots, respectively). In contrast, a '-' indicates the lack of a ~ 49 kDa band i.e. an undetectable/no yield of CHIKV E2ΔTM-HIS obtained in the sample.

3.3 Results

3.3.1 Nickel-Resin Purification

A one-step purification strategy was attempted to purify recombinant CHIKV E2ΔTM-HIS proteins using Ni-NTA affinity chromatography, under different buffer reagent, buffer pH, imidazole concentration, nickel resin volume added, and binding time conditions, to optimise the purification strategy. Table 3.1 summarises the Coomassie staining results of the ~49 kDa band intensities of CHIKV E2ΔTM-HIS in the wash (S) and elution (E) fractions observed.

An in-house methodology using the protein extraction buffer (section 2.2.4 – 1 x PBS, pH 7.4) as the protocol's equilibration, wash and elution (supplemented with 250 mM imidazole) buffers was investigated. This methodology washed out the majority of the recombinant proteins in washes S1 - S3 with minimal protein elution (E1)(#1). The same

result was seen when the Ni-resin volume was increased (#2). Neither scaling up the methodology and increasing the salt and imidazole concentration, nor increasing the binding time, successfully eluted recombinant proteins (#3). When using the manufacturer-recommended buffers, the same result is seen (#4) even when increasing the binding time (#5). In addition, purification under denaturing conditions was unsuccessful (#6) including when the elution buffer pH was lowered from pH 5.5 to pH 4.5 (#7). Since almost all the recombinant proteins are washed out during the equilibration wash, there seems to be no binding of the CHIKV E2ΔTM-HIS to the Ni-resin suggesting the 6xHis-tag may not be sufficiently exposed.

Table 3.1: Ni-NTA affinity chromatography conditions and results. EqB, equilibration buffer; WB, wash buffer; ElB, elution buffer. S1, equilibration wash; S2 - S3, wash 2 - 3; E1 - E3, elution 1 – 3. +, CHIKV E2ΔTM-HIS protein present in sample. Multiple '+' indicate higher CHIKV E2ΔTM-HIS proteins present; -, protein absent in sample.

#	Buffers	Imidazole concentration	Buffer pH	Resin volume	Binding time	S1	S2	S3	E1	E2	E3	Modification
1	EqB: 1 X PBS WB: 1 X PBS ElB: 1 X PBS	0 M 0 M 250 mM	7,4	50 µL	1 min	+++++	+++	++	+	-	-	
2	EqB: 1 X PBS WB: 1 X PBS ElB: 1 X PBS	0 M 0 M 250 mM	7,4	200 µL	1 min	+++++	+++	++	+	-	-	Resin volume increased
3	EqB: 1 X high-salt PBS (0.5 M NaCl) WB: 1 X high-salt PBS (0.5 M NaCl) ElB: 1 X high-salt PBS (0.5 M NaCl)	0 mM 0 mM 0.5 M	7,4	10 mL	60 mins	++++	+++	+	+	-	-	Salt concentration, binding time and imidazole concentration increased
4	EqB: 50 mM sodium phosphate, 0.3M NaCl WB: 50 mM sodium phosphate, 0.3M NaCl ElB: 50 mM sodium phosphate, 0.3M NaCl	0 M 0 M 250 mM	8,0	200 µL	1 min	+++++	++	+	+	-	-	Manufacturer's buffers used

Table 3.1 continued: Ni-NTA affinity chromatography conditions and results. EqB, equilibration buffer; WB, wash buffer; ElB, elution buffer. S1, equilibration wash; S2 - S3, wash 2 - 3; E1 - E3, elution 1 – 3. +, CHIKV E2ΔTM-HIS protein present in sample. Multiple '+' indicate higher CHIKV E2ΔTM-HIS proteins present; -, protein absent in sample.

5	EqB: 50 mM sodium phosphate, 0.3M NaCl WB: 50 mM sodium phosphate, 0.3M NaCl ElB: 50 mM sodium phosphate, 0.3M NaCl	0 M 0 M 250 mM	8,0	200 µL	1 hr	++++	+	-	-	-	-	Binding time increased
6*	EqB: 0.1 M sodium phosphate, 8 M urea WB: 0.1 M sodium phosphate, 8 M urea EluB: 0.1 M sodium phosphate, 8 M urea	0 M	8,0 6,3 5,5	200 µL	1 hr	++++	-	-	-	-	-	Denaturing conditions
7*	EqB: 0.1 M sodium phosphate, 8 M urea WB: 0.1 M sodium phosphate, 8 M urea EluB: 0.1 M sodium phosphate, 8 M urea	0 M	8,0 6,3 4,5	200 µL	1 hr	++++	-	-	-	-	-	Denaturing conditions, elution buffer more acidic

* CHIKV E2ΔTM-HIS glycoproteins extracted under denaturing conditions (1 X PBS supplemented with 8M urea)

3.3.2 Comparison of CHIKV E2ΔTM-HIS Resolved by Denaturing and Non-Denaturing Gel

The binding of anti-6xHis antibodies to denatured and non-denatured CHIKV E2ΔTM-HIS proteins was investigated by probing anti-6xHis antibodies to CHIKV E2ΔTM-HIS proteins resolved by denaturing and non-denaturing polyacrylamide gels, respectively (Fig. 3.1)

A dark purple band (~ 49 kDa) corresponding to CHIKV E2ΔTM-HIS was visualised on the denaturing blot (Den) whereas a much fainter one was visualised on the non-denaturing blot (ND). This provided evidence that the 6xHis-tag on CHIKV E2ΔTM-HIS was not exposed and therefore explained the lack of CHIKV E2ΔTM-HIS purification using Ni-NTA affinity chromatography. Consequently, other purification protocols were investigated.

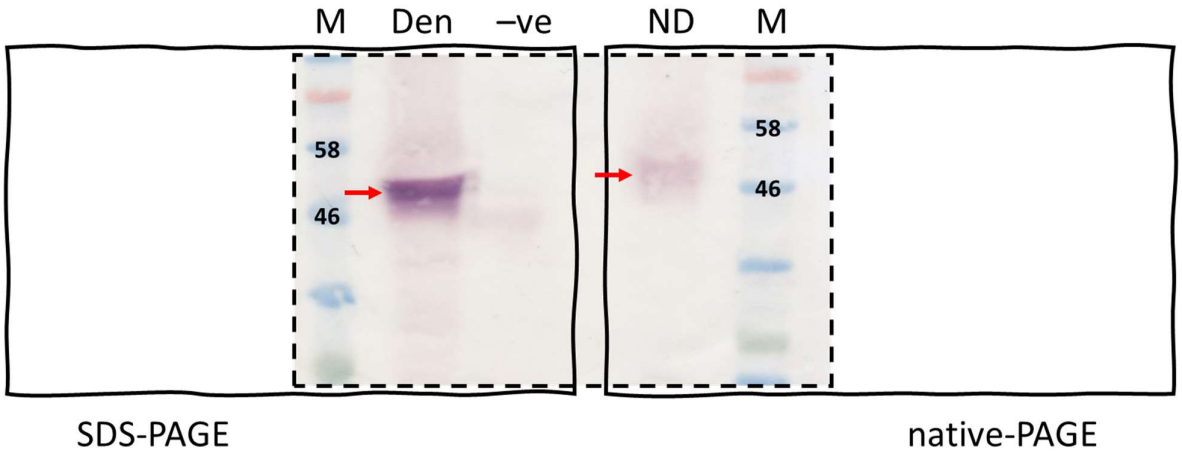

Figure 3.1: Western blot screening of CHIKV E2ΔTM-HIS proteins in their native conformation. A comparison of anti-6xHis antibody probing ability to denatured and non-denatured variants of the recombinant CHIKV E2ΔTM-HIS protein. M, MW marker; Den, denatured CHIKV E2ΔTM-HIS resolved by SDS-PAGE; -ve, protein extract from leaves infiltrated with empty pTRAkc-ERH; ND, non-denatured CHIKV E2ΔTM-HIS variant resolved by native-PAGE; red arrow; expected size of CHIKV E2ΔTM-HIS.

3.3.3 pH Precipitation

To investigate if pH precipitation can either partially or fully purify CHIKV E2ΔTM-HIS proteins from the remaining plant proteins, pellet and supernatant samples of crude extracts, harvested from infiltrated leaves, at different pH levels were screened for CHIKV E2ΔTM-HIS proteins. The supernatants and pellets of the pH precipitation samples were visualised by Coomassie staining (Fig. 3.2).

Compared to the unprocessed crude extract (C), none of the pH conditions investigated allowed for sufficient CHIKV E2ΔTM-HIS purification as other plant proteins were present in all samples containing CHIKV E2ΔTM-HIS, evidenced by the presence of multiple bands in all the samples tested. CHIKV E2ΔTM-HIS was present in all samples except for in the supernatants of the pH 5.5 and pH 3.5 samples and the pellet of the pH 5.0 sample. It

was evident that higher levels of CHIKV E2ΔTM-HIS can be found in the opposite samples to those mentioned above i.e. the pellets of the pH 5.5 and pH 3.5 samples and the supernatant of the pH 5.0 sample. However, while this is desirable, there were still other proteins present in the samples screened. In addition, RuBisCo and recombinant CRT co-purified with CHIKV E2ΔTM-HIS.

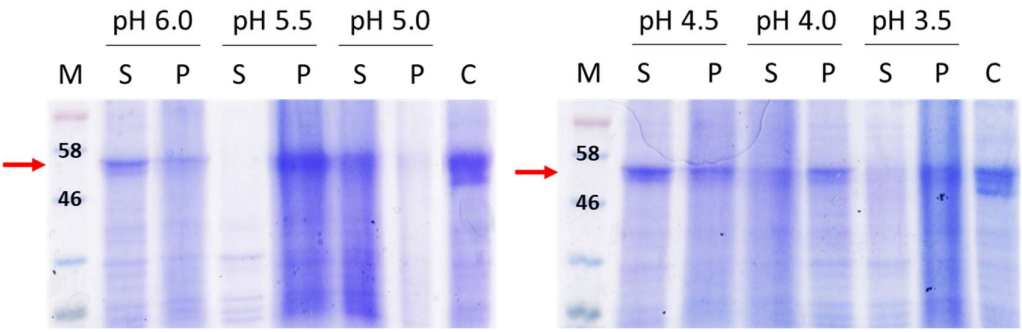

Figure 3.2: Coomassie staining of pH precipitated crude extract. Supernatant and pellet samples of crude extract that have undergone pH precipitation between a pH range of 6.0 and 3.5. M, MW marker; S, supernatant; P, pellet; C, unprocessed crude extract; red arrow; expected size of CHIKV E2ΔTM-HIS.

3.3.4 Ammonium Sulphate Precipitation

In order to purify CHIKV E2ΔTM-HIS, ammonium sulphate precipitation was performed on crude extracts from harvested leaves co-infiltrated with CHIKV E2ΔTM-HIS and CRT (Fig. 3.3A).

Western blotting of both the 0 -10% and 80+% ammonium sulphate fractions indicated undetectable levels of CHIKV E2ΔTM-HIS proteins. Fractions 30-40%, 40-50% and 50-60% ammonium sulphate precipitated the most CHIKV E2ΔTM-HIS protein and fractions 10-20%, 20-30%, 60-70% and 70-80% ammonium sulphate precipitated lower but detectable amounts of CHIKV E2ΔTM-HIS protein.

Since the major plant contaminant RuBisCo precipitates at approximately 40% (Delgado et al., 1995, Uemura et al., 1996), only recombinant CRT was screened to identify if CRT co-purified with CHIKV E2ΔTM-HIS in samples 30-40%, 40-50%, and 50-60% ammonium

sulphate (Fig. 3.3B). CRT co-precipitated with CHIKV E2ΔTM-HIS in all samples screened as evidenced by the presence of the intense ~49 kDa bands.

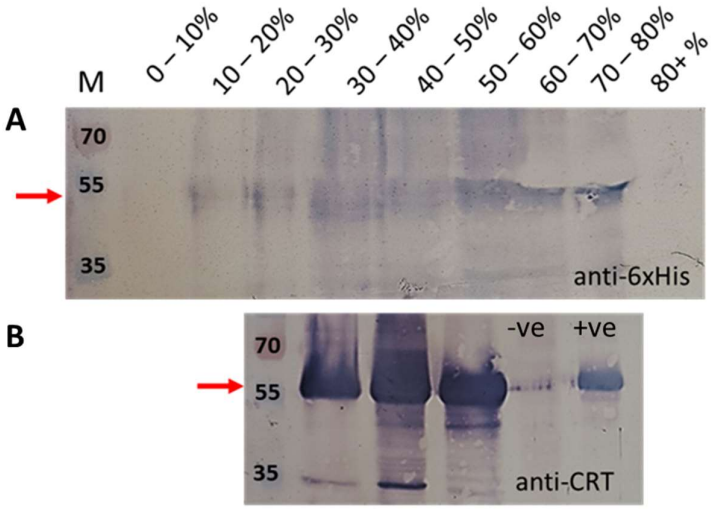

Figure 3.3: Western blot screening of ammonium sulphate-precipitation samples. Anti-6xHis and anti-CRT probing of resolved crude extract saturated from 0 – 80+% and 30 – 60% with ammonium sulphate for CHIKV E2ΔTM-HIS (A) and CRT (B) screening, respectively. M, MW marker; red arrow, expected size of recombinant protein of interest; -ve, no-CRT negative control; +ve, positive control generated by anti-CRT probing of crude extract harvested from leaves infiltrated with pEAQ-*HT* CRT.

3.4 Discussion

Industrial processes for the downstream processing of microbial- and mammalian cell-produced recombinant proteins are a well-established practice; however, DSP remains the majority of the overall cost (Schillberg et al., 2019). While plant-expression systems potentially offer a cheaper alternative (due to low-cost upstream cultivation) to recombinant protein production (Tschofen et al., 2016), whole plant-produced recombinant proteins are extracted simultaneously with a large quantity of insoluble proteins of the plant matrix and other soluble plant host proteins. This requires the elimination of large quantities of contaminants during DSP and this has the potential to complicate purification of the protein of interest and further increase DSP costs (Schillberg et al., 2019). Cheap and efficient downstream processing protocols are critical for the development of cost-effective biopharmaceuticals in plants. A one-step purification strategy is favoured over partial purification since partial purification takes

multiple steps to eventually fully purify the recombinant protein of interest and, as a result, increases the costs and labours of DSP.

Previous studies have shown the efficiency of using Ni-NTA chromatography to purify recombinant 6xHIS-tagged CHIKV E2 glycoproteins in insect (Cho et al., 2008, Bagno et al., 2020), *E. coli* (Verma et al., 2016, Bagno et al., 2020) and mammalian cells (Bagno et al., 2020). These studies present further evidence for how the 6xHis-tag and Ni-NTA chromatography provide a universal purification system of recombinant antibodies: if recombinant proteins are insoluble, located in inclusion bodies and/or have a tertiary structure such that the 6xHis-tag is hidden within, proteins can be extracted under denaturing conditions and subsequently purified successfully. Another strategy used that has the potential to bypass denaturing conditions is the movement of the 6xHis-tag to the opposite terminus of the protein (Crowe et al., 1994). An attempt at one-step purification using Ni-NTA affinity chromatography was unsuccessful under multiple conditions. Most of the CHIKV E2ΔTM-HIS protein washed out in equilibration wash S1 and the rest in washes S2 – S3, with zero to minimal elution of CHIKV E2ΔTM-HIS proteins (Table 3.1). This indicates a lack of protein binding to the Ni-resin by way of the fused 6xHis-tag. It is possible that nickel affinity purification was unsuccessful due to the 6xhis-tag being insufficiently exposed on the native CHIKV E2ΔTM-HIS protein to bind to immobilised Ni-ions of the Ni-affinity resin. In rare cases, the 6xHis-tag becomes hidden by the tertiary structure of the protein (Crowe et al., 1994). It is known that the stability of the interaction of the 6xHis-tag to the immobilised Ni-resin is dependent on how accessible the tag is on the protein of interest (Block et al., 2009). To investigate this, CHIKV E2ΔTM-HIS proteins were resolved by non-denaturing as well as denaturing polyacrylamide gels and probed by western blotting using anti-6xHis antibodies to compare anti-6xHis antibody binding by assessing the intensity of the CHIKV E2ΔTM-HIS protein bands (Fig. 3.1). Analysis of CHIKV E2ΔTM-HIS in its native conformation by native-PAGE showed that the 6xHis-tag is not exposed as anti-6xHis antibodies were unable to bind the 6xHis-tag when CHIKV E2ΔTM-HIS is in its native conformation (Fig. 3.1). Consequently, the problems with the binding of CHIKV E2ΔTM-HIS to the immobilised Ni-resin was credited to the unexposed 6xHis-tag. While denaturing conditions were used to remedy these problems, they were not satisfactory. It may be

worth attempting to purify the recombinant CHIKV E2ΔTM protein with the 6xHis-tag on the opposite terminus as this may bypass the need to alter conditions and may even allow for purification of the recombinant protein under non-denaturing conditions (Crowe et al., 1994). While this may have shown promising results, the time-constraints present at this stage of the study did not permit additional cloning, expression time-trials and further purification experimentation. Thus, alternative purification strategies were tested.

Since the pI of a protein of interest can be utilised to precipitate proteins at a pH level closest to that of the protein's pI (Pelegrine and Gasparetto, 2005), a pH precipitation protocol was performed on crude extract. This investigated whether CHIKV E2ΔTM-HIS, or a large portion of the other plant proteins (the large subunit of RuBisCo and CRT especially), would precipitate under different pH conditions in order to either fully or partially purify CHIKV E2ΔTM-HIS (Fig. 3.2). None of the pH conditions were able to purify CHIKV E2ΔTM-HIS fully or sufficiently enough for partial purification. Furthermore, the large subunit of RuBisCo and CRT contaminants were present in all samples that contained CHIKV E2ΔTM-HIS. Thus, pH purification is not an efficient strategy for CHIKV E2ΔTM-HIS purification.

Ammonium sulphate precipitation has remained the most widely used precipitation method. Ammonium sulphate is stabilising to the structure of proteins, very soluble, relatively cheap and has readily available pure material (Burgess, 2009). Since ammonium sulphate precipitation is a tried and tested purification strategy, crude extract underwent this protocol to investigate if CHIKV E2ΔTM-HIS proteins can be efficiently purified. Fractions were screened for CHIKV E2ΔTM-HIS and possible protein contaminants, like recombinant CRT and RuBisCo, in samples where CHIKV E2ΔTM-HIS precipitated to investigate the purity of the fractions (Fig. 3.3). While CHIKV E2ΔTM-HIS precipitated in a large saturation range (10 – 80%), fractions 30 – 60% precipitated most of the protein with no detectable protein precipitation seen in fractions 0-10% and 80+% ammonium sulphate. In studies where RuBisCo was purified by ammonium sulphate precipitation (Delgado et al., 1995, Uemura et al., 1996), and in-house ammonium sulphate precipitation results, RuBisCo precipitates at approximately 40% ammonium sulphate i.e. one of the saturation levels where CHIKV E2ΔTM-HIS more predominantly

precipitated. However, CHIKV E2ΔTM-HIS was still predominantly precipitated in fractions higher than 40% ammonium sulphate (40 – 70%). Therefore, only the presence of recombinant CRT was screened for since literature (Delgado et al., 1995, Uemura et al., 1996) provided RuBisCo's ammonium sulphate precipitation results. Screening of recombinant CRT in fractions 30 – 80% indicated that the highest levels of CRT co-purified in the fractions where CHIKV E2ΔTM-HIS most predominantly precipitated. Therefore, while ammonium sulphate is able to largely precipitate CHIKV E2ΔTM-HIS between 30 – 60% ammonium sulphate, RuBisCo and CRT proteins contaminate these fractions.

While CHIKV E2ΔTM-HIS was partially purified by ammonium sulphate precipitation, another purification step is needed thereafter to remove RuBisCo and CRT contaminants. Whether CRT and plant-protein contaminants will influence background in a downstream diagnostic assay is unknown; however, it is ideal that they be removed during the purification process. It would be worth investigating a second purification step by means of ion-exchange chromatography. This was not performed due to time constraints. In addition, it is worth investigating the exposure of a 6xHis-tag when fused to the C-terminus of CHIKV E2ΔTM, as opposed to the N-terminus in this study, as this conformation may expose the 6xHis-tag and allow for one-step purification by Ni-NTA chromatography. If the elution is still not pure enough, ammonium sulphate precipitation is often performed before Ni-NTA chromatography to remove the majority of the other plant protein contaminants. Otherwise, optimising an Ni-NTA chromatography technique under denaturing conditions of CHIKV E2ΔTM-HIS to expose the 6xHis-tag could potentially bypass the need for further purification strategies.

Chapter 4: Conclusion

In this study, two variants of the CHIKV E2 envelope glycoprotein were expressed in *Nicotiana* benthamiana plants for the development of a diagnostic reagent for the detection of CHIKV antibodies. The CHIKV E2 protein is an appropriate candidate for a CHIKV diagnostic reagent owing to its conservation and the fact that it is a major target for neutralizing antibodies against CHIKV during early infection.

The two synthesised CHIKV E2 gene variants were successfully cloned into pTRAkc-ERH plant expression vector. However, cloning of the same gene variants into pEAQ-*HT* plant expression vector was unsuccessful, despite the generation of two different cloning strategies. This may have been due to the size of the pEAQ-*HT* vector and the CHIKV E2 variant-gene inserts that were to be ligated into it. Likewise, the size of the resulting expression construct may have been too big to successfully transform into bacteria, as the bigger the construct to be transformed, the less likely transformation will be successful. Additionally, the pTRAkc-ERH expression constructs were successfully transformed into *A. tumefaciens* – as the pTRAkc-ERH constructs were smaller than the pEAQ-*HT* constructs - and subsequently expressed in *Nicotiana* plants.

The expression of both recombinant CHIKV E2 glycoprotein variants in *Nicotiana benthamiana* plants was successful. Previous work on CHIKV E2 and CHIKV E2ΔTM in insect cells has shown similar results: the transmembrane-deficient variant expressed higher levels of protein than the full-length variant. It is known that transmembrane proteins are often difficult to express in heterologous hosts and are often toxic to the host cell. Elimination of the transmembrane domain of proteins has been shown to allow or improve expression of the truncated protein. This may be the reason for the transmembrane-deficient variant of the CHIKV E2 glycoprotein expressing at higher levels than the full-length CHIKV E2 glycoprotein.

Although the CHIKV E2 protein variants produced in plants were detectable, yields were still not very high in comparison with the levels of many other recombinant proteins produced in plants. Since one of the most desirable outcomes of utilising plant-produced proteins is the generation of cost-effective reagents, it would be prudent to increase expression levels of CHIKV E2 variants. There are several approaches that one could

employ to do this. We showed with this work, that the co-expression of the chaperone protein CRT with CHIKV E2ΔTM-HIS increased its production. Chaperone proteins are responsible for the correct folding, during or after synthesis, of proteins and recombinant NSs can be used as a tool to suppress the deleterious effects of RNA silencing of exogenous expression of recombinant proteins and improve their expression. In addition, the recombinant chaperones are theorised to bridge the gap between the inherent differences in endogenous chaperone machinery between the plant and natural mammalian cell hosts - where viral glycoproteins are naturally produced - in addition to a potential bottleneck in the plant folding machinery to produce exogenous viral glycoproteins. The improved expression seen for the CHIKV E2-HIS variants when co-expressed with these proteins (co-expression of CHIKV E2ΔTM-HIS and CRT especially) may be credited to these reasons. This provided further evidence for the benefits of reducing the limitations of plant-expression technology by means of further optimisation strategies. Selection, optimisation, and extraction of the best expressed CHIKV E2 variant was successful and the conditions were scaled up. Co-expression of CHIKV E2ΔTM-HIS and CRT resulted in the highest yield of recombinant CHIKV E2ΔTM-HIS proteins when infiltrated plant leaves were harvested 3 days post-infiltration (dpi) and total soluble proteins extracted in a 1 x PBS, pH 7.4 buffer. While co-expression of the CHIKV E2 protein variants with NSs and CNX did improve its expression on certain dpi, the expression levels were still not sufficient to achieve the yields needed for a diagnostic reagent.

It is possible that the low levels of expression may be due to degradation of CHIKV E2-HIS protein variants as a result of underglycosylation. Co-expressing the CHIKV E2 glycoprotein variants with a recombinant oligosaccharyltransferase from *Leishmania major* would have been worth investigating as this has been shown to substantially increase N-glycosylation occupancy on N-glycosylation sites of other recombinant glycoproteins before. This is vital since glycosylation facilitates protein folding, stability, and the modulation of protein-protein interactions. Co-expression of CHIKV E2-HIS variants in the glycol-engineered ΔXT/FT *N. benthamiana* line – a plant line that lacks plant-specific β-1,2-linked xylose and core α-1,3-linked fucose - with a recombinant

oligosaccharyltransferase from *Leishmania major* may have had similar if not better results.

Fusing a 6xHis-tag onto the C-terminus of each variant provided a suitable detection and verification strategy using anti-HIS antibodies in lieu of anti-CHIKV E2 antibodies. While anti-CHIKV E2 antibodies are ideal, their availability during the research project were limited and expensive. Additionally, the 6xHis-tag is a cheaper alternative that contributes to the cost-effective advantage of biopharmaceutical development in plants. The 6xHis-tag is also a widely-used technology for detection and purification of recombinant proteins. However, while detection of recombinant CHIKV E2 proteins was successful using the 6xHis-tag, purification of plant-produced CHIKV E2ΔTM-HIS proteins by Ni-NTA affinity chromatography proved troublesome owing to the unexposed 6xHis-tag hidden within the recombinant proteins tertiary structure. This is not ideal since, while plant-expression systems boast lower upstream costs compared to the more traditional expression systems, downstream-processing remains the largest contributor to overall development costs. Therefore, since one-step purification of the CHIKV E2ΔTM-HIS protein by Ni-NTA affinity chromatography failed, additional cost-raising steps need to be included in the purification of the recombinant proteins.

Purification of CHIKV E2ΔTM-HIS proteins by pH precipitation was unable to separate CHIKV E2ΔTM-HIS proteins from other soluble plant proteins. This protocol only relies on the difference of pI between the soluble proteins and is not as specific as Ni-NTA chromatography to bind and elute only the proteins containing a 6xHis-tag. Ammonium sulphate precipitation was able to best precipitate CHIKV E2ΔTM-HIS proteins between 30-60% ammonium sulphate. However, RuBisCo and large amounts of recombinant CRT contaminated these samples in the fractions below 40% ammonium sulphate and 30-60% ammonium sulphate, respectively. Recombinant CRT proteins may be precipitating with the CHIKV E2ΔTM-HIS proteins as their solubility in the presence of varying salt ($(NH_4)_2SO_4$) concentrations may be similar. This would require a further purification step to separate CHIKV E2ΔTM-HIS proteins from the remaining plant contaminants. Future research should be done on fully purifying CHIKV E2ΔTM-HIS proteins in the dialysed 30-60% ammonium sulphate fractions by ion-exchange chromatography. In addition, it should be investigated and compared whether fusing the 6xHis-tag onto the C-terminus,

as well as or in lieu of the N-terminal 6xHis-tag, of recombinant CHIKV E2ΔTM will expose the tag(s) sufficiently and allow for the one-step purification of CHIKV E2ΔTM-HIS by Ni-NTA affinity chromatography.

In conclusion, this research investigates the feasibility and functionality of using recombinant plant-produced CHIKV E2 glycoprotein variants to develop a cheap and efficacious diagnostic assay for the rapid detection of CHIKV infection. This study further corroborates the merits of using a tobacco plant-expression technology for the development of a cost-effective, efficacious biopharmaceuticals. In addition, it highlights the limitations of the respective stages of the technology and, in some cases, how to overcome them.

References

AKAHATA, W., YANG, Z. Y., ANDERSEN, H., SUN, S., HOLDAWAY, H. A., KONG, W. P., LEWIS, M. G., HIGGS, S., ROSSMANN, M. G., RAO, S. & NABEL, G. J. 2010. A virus-like particle vaccine for epidemic Chikungunya virus protects nonhuman primates against infection. *Nat Med,* 16**,** 334-8.

APPASSAKIJ, H., KHUNTIKIJ, P., KEMAPUNMANUS, M., WUTTHANARUNGSAN, R. & SILPAPOJAKUL, K. 2013. Viremic profiles in asymptomatic and symptomatic chikungunya fever: a blood transfusion threat? *Transfusion,* 53**,** 2567-2574.

ARNDT, C., KORISTKA, S., BARTSCH, H. & BACHMANN, M. 2012. Native polyacrylamide gels. *Methods Mol Biol,* 869**,** 49-53.

BAGNO, F. F., GODÓI, L. C., FIGUEIREDO, M. M., SÉRGIO, S. A. R., MORAES, T. D. F. S., SALAZAR, N. D. C., KIM, Y. C., REYES-SANDOVAL, A. & DA FONSECA, F. G. 2020. Chikungunya E2 Protein Produced in E. coli and HEK293-T Cells—Comparison of Their Performances in ELISA. *Viruses,* 12.

BAO, H., RAMANATHAN, A. A., KAWALAKAR, O., SUNDARAM, S. G., TINGEY, C., BIAN, C. B., MURUGANANDAM, N., VIJAYACHARI, P., SARDESAI, N. Y., WEINER, D. B., UGEN, K. E. & MUTHUMANI, K. 2013. Nonstructural protein 2 (nsP2) of Chikungunya virus (CHIKV) enhances protective immunity mediated by a CHIKV envelope protein expressing DNA Vaccine. *Viral Immunol,* 26**,** 75-83.

BLOCK, H., MAERTENS, B., SPRIESTERSBACH, A., BRINKER, N., KUBICEK, J., FABIS, R., LABAHN, J. & SCHÄFER, F. 2009. Chapter 27 Immobilized-Metal Affinity Chromatography (IMAC). *Guide to Protein Purification, 2nd Edition.*

BURDINO, E., CALLERI, G., CARAMELLO, P. & GHISETTI, V. 2016. Unmet Needs for a Rapid Diagnosis of Chikungunya Virus Infection. *Emerg Infect Dis,* 22**,** 1837-9.

BURGESS, R. R. 2009. Chapter 20 Protein Precipitation Techniques. *Guide to Protein Purification, 2nd Edition.*

BURT, F. J., CHEN, W., MINER, J. J., LENSCHOW, D. J., MERITS, A., SCHNETTLER, E., KOHL, A., RUDD, P. A., TAYLOR, A., HERRERO, L. J., ZAID, A., NG, L. F. P. & MAHALINGAM, S. 2017. Chikungunya virus: an update on the biology and pathogenesis of this emerging pathogen. *The Lancet Infectious Diseases,* 17**,** e107-e117.

CAMPOS, G. S., ALBUQUERQUE BANDEIRA, A. C., DINIZ ROCHA, V. F., DIAS, J. P., CARVALHO, R. H. & SARDI, S. I. 2017. First Detection of Chikungunya Virus in Breast Milk. *Pediatr Infect Dis J,* 36**,** 1015-1017.

CARDONA-OSPINA, J. A., SEPULVEDA-ARIAS, J. C., MANCILLA, L. & GUTIERREZ-LOPEZ, L. G. 2016. Plant expression systems, a budding way to confront chikungunya and Zika in developing countries? *F1000Res,* 5**,** 2121.

CAREY, D. E. 1971. Chikungunya and Dengue: A Case of Mistaken Identity? *J Hist Med Allied Sci,* 26**,** 243-262.

CASTILHO, A., BEIHAMMER, G., PFEIFFER, C., GORITZER, K., MONTERO-MORALES, L., VAVRA, U., MARESCH, D., GRUNWALD-GRUBER, C., ALTMANN, F., STEINKELLNER, H. & STRASSER, R. 2018. An oligosaccharyltransferase from Leishmania major increases the N-glycan occupancy on recombinant glycoproteins produced in Nicotiana benthamiana. *Plant Biotechnol J,* 16**,** 1700-1709.

CHO, B., JEON, B. Y., KIM, J., NOH, J., KIM, J., PARK, M. & PARK, S. 2008. Expression and evaluation of Chikungunya virus E1 and E2 envelope proteins for serodiagnosis of Chikungunya virus infection. *Yonsei Med J,* 49**,** 828-35.

CHRISTOU, P. 1996. Transformation technology. *Trends in Plant Science,* 1**,** 423-431.

CROWE, J., DOBELI, H., GENTZ, R., HOCHULI, E., STUBER, D. & HENCO, K. 1994. 6xHis-Ni-NTA chromatography as a superior technique in recombinant protein expression/purification. *Methods Mol Biol,* 31**,** 371-87.

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.

DELGADO, E., MEDRANO, H., KEYS, A. J. & PARRY, M. A. J. 1995. Species variation in Rubisco specificity factor. *J. Exp. Bot.,* 46**,** 1775-1777.

DUONG-LY, K. C. & GABELLI, S. B. 2014. Salting out of proteins using ammonium sulfate precipitation. *Methods Enzymol,* 541**,** 85-94.

DUVENAGE, L., HITZEROTH, II, MEYERS, A. E. & RYBICKI, E. P. 2013. Expression in tobacco and purification of beak and feather disease virus capsid protein fused to elastin-like polypeptides. *J Virol Methods,* 191**,** 55-62.

EDELMAN, R., TACKET, C. O., WASSERMANS, S. S., BODISON, A., PERRY, J. G. & MANGIAFICO, J. A. 2000. Phase II safety and immunogenicity study of live chikungunya virus vaccine TSI-GSD-218. *Am J Trop Med Hyg* 62**,** 681-685.

EDWARDS, C. J., WELCH, S. R., CHAMBERLAIN, J., HEWSON, R., TOLLEY, H., CANE, P. A. & LLOYD, G. 2007. Molecular diagnosis and analysis of Chikungunya virus. *J Clin Virol,* 39**,** 271-5.

EHTESHAMI, M., TAO, S., ZANDI, K., HSIAO, H. M., JIANG, Y., HAMMOND, E., AMBLARD, F., RUSSELL, O. O., MERITS, A. & SCHINAZI, R. F. 2017. Characterization of beta-d-N(4)-Hydroxycytidine as a Novel Inhibitor of Chikungunya Virus. *Antimicrob Agents Chemother,* 61.

EINHAUER, A. & JUNGBAUER, A. 2001. The FLAG™ peptide, a versatile fusion tag for the purification of recombinant proteins. *J Biochem. Biophys. Methods,* 49**,** 455-465.

FISCHER, R. & EMANS, N. 2000. <Fischer-Emans2000_Article_MolecularFarmingOfPharmaceutic.pdf>. *Transgenic Research,* 9**,** 279-299.

FISCHER, R., STOGER, E., SCHILLBERG, S., CHRISTOU, P. & TWYMAN, R. M. 2004. Plant-based production of biopharmaceuticals. *Curr Opin Plant Biol,* 7**,** 152-8.

FOX, J. M., LONG, F., EDELING, M. A., LIN, H., VAN DUIJL-RICHTER, M. K. S., FONG, R. H., KAHLE, K. M., SMIT, J. M., JIN, J., SIMMONS, G., DORANZ, B. J., CROWE, J. E., FREMONT, D. H., ROSSMANN, M. G. & DIAMOND, M. S. 2015. Broadly Neutralizing Alphavirus Antibodies Bind an Epitope on E2 and Inhibit Entry and Egress. *Cell.,* 163**,** 1095-1107.

FRIC, J., BERTIN-MAGHIT, S., WANG, C. I., NARDIN, A. & WARTER, L. 2013. Use of human monoclonal antibodies to treat Chikungunya virus infection. *J Infect Dis,* 207**,** 319-22.

FUGLSANG, A. 2003. Codon optimizer: a freeware tool for codon optimization. *Protein Exp. Purif.,* 31**,** 247-249.

FUMAGALLI, M. J., DE SOUZA, W. M., ESPOSITO, D. L. A., SILVA, A., ROMEIRO, M. F., MARTINEZ, E. Z., DA FONSECA, B. A. L. & FIGUEIREDO, L. T. M. 2018. Enzyme-linked immunosorbent assay using recombinant envelope protein 2 antigen for diagnosis of Chikungunya virus. *Virol J,* 15**,** 112.

GABERC-POREKAR, V. & MENART, V. 2001. Perspectives of immobilized-metal affinity chromatography. *Journal of Biochemical and Biophysical Methods,* 49**,** 335-360.

GALÁN-HUERTA, K. A., RIVAS-ESTILLA, A. M., FERNÁNDEZ-SALAS, I., FARFAN-ALE, J. A. & RAMOS-JIMÉNEZ, J. 2015. Chikungunya virus: A general overview. *Medicina Universitaria,* 17**,** 175-183.

GRIVARD, P., LE ROUX, K., LAURENT, P., FIANU, A., PERRAU, J., GIGAN, J., HOARAU, G., GRONDIN, N., STAIKOWSKY, F., FAVIER, F. & MICHAULT, A. 2007. Molecular and serological diagnosis of Chikungunya virus infection. *Pathol Biol (Paris),* 55**,** 490-4.

GRODZKI, A. C. & BERENSTEIN, E. 2010. Antibody purification: ammonium sulfate fractionation or gel filtration. *Methods Mol Biol,* 588**,** 15-26.

HASAN, M., ISLAM, A., CHAKMA, C., KHATUN, Z. & RASHID, H. 2018. Chikungunya virus pathogenesis: A global bioterrorism for public health. *The Pharma Journal* 7**,** 15-24.

HITZEROTH, I. I. & VAN ZYL, A. R. 2016. Transient Expression of Viral Proteins in Plants Using Agrobacterium tumefaciens. *In:* THOMAS, S. (ed.) *Vaccine Design: Methods and Protocols, Volume 2: Vaccines for Veterinary Diseases.* New York, NY: Springer New York.

JOZALA, A. F., GERALDES, D. C., TUNDISI, L. L., FEITOSA, V. A., BREYER, C. A., CARDOSO, S. L., MAZZOLA, P. G., OLIVEIRA-NASCIMENTO, L., RANGEL-YAGUI, C. O., MAGALHAES, P. O., OLIVEIRA, M. A. & PESSOA, A., JR. 2016. Biopharmaceuticals from microorganisms: from production to purification. *Braz J Microbiol,* 47 Suppl 1**,** 51-63.

KAM, Y. W., LUM, F. M., TEO, T. H., LEE, W. W., SIMARMATA, D., HARJANTO, S., CHUA, C. L., CHAN, Y. F., WEE, J. K., CHOW, A., LIN, R. T., LEO, Y. S., LE GRAND, R., SAM, I. C., TONG, J. C., ROQUES, P., WIESMULLER, K. H., RENIA, L., ROTZSCHKE, O. & NG, L. F. 2012. Early neutralizing IgG response to Chikungunya virus in infected patients targets a dominant linear epitope on the E2 glycoprotein. *EMBO Mol Med,* 4**,** 330-43.

KAUR, P., THIRUCHELVAN, M., LEE, R. C., CHEN, H., CHEN, K. C., NG, M. L. & CHU, J. J. 2013. Inhibition of chikungunya virus replication by harringtonine, a novel antiviral that suppresses viral protein expression. *Antimicrob Agents Chemother,* 57**,** 155-67.

KHAN, M., DHANWANI, R., RAO, P. V. L. & PARIDA, M. 2012. Subunit vaccine formulations based on recombinant envelope proteins of Chikungunya virus elicit balanced Th1/Th2 response and virus-neutralizing antibodies in mice. *Virus Res.,* 167**,** 236-246.

KIMPLE, M. E., BRILL, A. L. & PASKER, R. L. 2013. Overview of affinity tags for protein purification. *Curr Protoc Protein Sci,* 73**,** Unit 9 9.

KING, T. P. 1972. Separation of proteins by ammonium sulfate gradient solubilization. *Biochemistry,* 11**,** 367-71.

KNUDSEN, A. B. 1995. Global distribution and continuing spread of Aedes albopictus. *Parassitologia,* 37**,** 91-97.

KUO, S. C., CHEN, Y. J., WANG, Y. M., TSUI, P. Y., KUO, M. D., WU, T. Y. & LO, S. J. 2012. Cell-based analysis of Chikungunya virus E1 protein in membrane fusion. *J Biomed Sci,* 19**,** 44.

LEVITT, N. H., RAMSBURG, H. H., HASTY, S. E., REPIK, P. M., COLE, F. E. & LUPTON, H. W. 1986. Development of an attenuated strain of chikungunya virus for use in vaccine production. *Vaccine,* 4**,** 157-162.

LIÉNARD, D., SOURROUILLE, C., GOMORD, V. & FAYE, L. 2007. Pharming and transgenic plants.

MA, J. K., DRAKE, P. M. & CHRISTOU, P. 2003. The production of recombinant pharmaceutical proteins in plants. *Nat Rev Genet,* 4**,** 794-805.

MACDONALD, J., MILETIC, S., GAILDRY, T., CHIN-FATT, A. & MENASSA, R. 2017. Co-expression with the Type 3 Secretion Chaperone CesT from Enterohemorrhagic E. coli Increases Accumulation of Recombinant Tir in Plant Chloroplasts. *Front Plant Sci,* 8**,** 283.

MACLEAN, J., KOEKEMOER, M., OLIVIER, A. J., STEWART, D., HITZEROTH, II, RADEMACHER, T., FISCHER, R., WILLIAMSON, A. L. & RYBICKI, E. P. 2007. Optimization of human papillomavirus type 16 (HPV-16) L1 expression in plants: comparison of the suitability of different HPV-16 L1 gene variants and different cell-compartment localization. *J Gen Virol,* 88**,** 1460-9.

MARGOLIN, E., CHAPMAN, R., MEYERS, A. E., VAN DIEPEN, M. T., XIMBA, P., HERMANUS, T., CROWTHER, C., WEBER, B., MORRIS, L., WILLIAMSON, A. L. & RYBICKI, E. P. 2019. Production and Immunogenicity of Soluble Plant-Produced HIV-1 Subtype C Envelope gp140 Immunogens. *Front Plant Sci,* 10**,** 1378.

MARGOLIN, E., CHAPMAN, R., WILLIAMSON, A. L., RYBICKI, E. P. & MEYERS, A. E. 2018. Production of complex viral glycoproteins in plants as vaccine immunogens. *Plant Biotechnol J*.

MATUSALI, G., COLAVITA, F., BORDI, L., LALLE, E., IPPOLITO, G., CAPOBIANCHI, M. R. & CASTILLETTI, C. 2019. Tropism of the Chikungunya Virus. *Viruses,* 11.

MAVALANKAR, D., SHASTRI, P. & RAMAN, P. 2007. Chikungunya epidemic in India: a major public-health disaster. *The Lancet Infect. Dis.,* 7**,** 306-307.

MBEWANA, S. 2017. *Development of Rift Valley fever virus candidate vaccines and reagents produced in Nicotiana benthamiana.* Doctor of Philosophy PhD Thesis, University of Cape Town.

METZ, S. W., GARDNER, J., GEERTSEMA, C., LE, T. T., GOH, L., VLAK, J. M., SUHRBIER, A. & PIJLMAN, G. P. 2013. Effective chikungunya virus-like particle vaccine produced in insect cells. *PLoS Negl Trop Dis,* 7**,** e2124.

METZ, S. W., GEERTSEMA, C., MARTINA, B. E., ANDRADE, P., HELDENS, J. G., VAN OERS, M. M., GOLDBACH, R. W., VLAK, J. M. & PIJLMAN, G. P. 2011. Functional processing and secretion of Chikungunya virus E1 and E2 glycoproteins in insect cells. *Virol J,* 8**,** 353.

MEYERS, A., CHAKAUYA, E., SHEPHARD, E., TANZER, F. L., MACLEAN, J., LYNCH, A., WILLIAMSON, A. L. & RYBICKI, E. P. 2008. Expression of HIV-1 antigens in plants as potential subunit vaccines. *BMC Biotechnol,* 8**,** 53.

MINER, J. J., YEANG, H. X. A., FOX, J. M., TAFFNER, S., MALKOVA, O. N., OH, S. T., KIM, A. H. J., DIAMOND, M. S., LENSCHOW, D. J. & YOKOYAMA, W. M. 2015. CHikungunya Viral Arthritis in the United States: A Mimic of Seronegative Rheumatoid Arthritis. *Arthritis & Rheumatology* 67**,** 1214-1220.

MOULAY, D., AZIZ-ALAOUI, M. A. & KWON, H. D. 2012. Optimal control of chikungunya disease: larvae reduction, treatment and prevention. *Math Biosci Eng,* 9**,** 369-92.

NAGELS, B., WETERINGS, K., CALLEWAERT, N. & VAN DAMME, E. J. M. 2012. Production of Plant Made Pharmaceuticals: From Plant Host to Functional Protein. *Critical Reviews in Plant Sciences,* 31**,** 148-180.

NANDI, S., KWONG, A. T., HOLTZ, B. R., ERWIN, R. L., MARCEL, S. & MCDONALD, K. A. 2016. Techno-economic analysis of a transient plant-based platform for monoclonal antibody production. *MAbs,* 8**,** 1456-1466.

PELEGRINE, D. H. G. & GASPARETTO, C. A. 2005. Whey proteins solubility as function of temperature and pH. *LWT - J Food Sci Technol,* 38**,** 77-80.

PERA, F. F., MUTEPFA, D. L., KHAN, A. M., ELS, J. H., MBEWANA, S., VAN DIJK, A. A., RYBICKI, E. P. & HITZEROTH, II 2015. Engineering and expression of a human rotavirus candidate vaccine in Nicotiana benthamiana. *Virol J,* 12**,** 205.

PETERSEN, L. R., JAMIESON, D. J., POWERS, A. M. & HONEIN, M. A. 2016. Zika Virus. *N Engl J Med,* 374**,** 1552-63.

PEYRET, H. & LOMONOSSOFF, G. P. 2013. The pEAQ vector series: the easy and quick way to produce recombinant proteins in plants. *Plant Mol Biol,* 83**,** 51-8.

PEZZI, L., REUSKEN, C. B., WEAVER, S. C., DREXLER, J. F., BUSCH, M., LABEAUD, A. D., DIAMOND, M. S., VASILAKIS, N., DREBOT, M. A., SIQUEIRA, A. M., RIBEIRO, G. S., KOHL, A., LECUIT, M., NG, L. F. P., GALLIAN, P., DE LAMBALLERIE, X., GLOPID-R CHIKUNGUNYA, O. N.-N. & MAYARO VIRUS WORKING, G. 2019. GloPID-R report on Chikungunya, O'nyong-nyong and Mayaro virus, part I: Biological diagnostics. *Antiviral Res,* 166**,** 66-81.

PHOOLCHAROEN, W., BHOO, S. H., LAI, H., MA, J., ARNTZEN, C. J., CHEN, Q. & MASON, H. S. 2011. Expression of an immunogenic Ebola immune complex in Nicotiana benthamiana. *Plant Biotechnol J,* 9**,** 807-16.

PIALOUX, G., GAÜZÈRE, B.-A., JAURÉGUIBERRY, S. & STROBEL, M. 2007. Chikungunya, an epidemic arbovirosis. *The Lancet Infect. Dis.,* 7**,** 319-327.

PLANTE, K. S., ROSSI, S. L., BERGREN, N. A., SEYMOUR, R. L. & WEAVER, S. C. 2015. Extended Preclinical Safety, Efficacy and Stability Testing of a Live-attenuated Chikungunya Vaccine Candidate. *PLoS Negl Trop Dis,* 9**,** e0004007.

PLATIS, D., DROSSARD, J., FISCHER, R., MA, J. K. & LABROU, N. E. 2008. New downstream processing strategy for the purification of monoclonal antibodies from transgenic tobacco plants. *J Chromatogr A,* 1211**,** 80-9.

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.

POWERS, A. M. & LOGUE, C. H. 2007. Changing patterns of chikungunya virus: re-emergence of a zoonotic arbovirus. *J Gen Virol,* 88**,** 2363-77.

PRAT, C. M., FLUSIN, O., PANELLA, A., TENEBRAY, B., LANCIOTTI, R. & LEPARC-GOFFART, I. 2014. Evaluation of commercially available serologic diagnostic tests for chikungunya virus. *Emerg Infect Dis,* 20**,** 2129-32.

RAMSAUER, K., SCHWAMEIS, M., FIRBAS, C., MÜLLNER, M., PUTNAK, R. J., THOMAS, S. J., DESPRÈS, P., TAUBER, E., JILMA, B. & TANGY, F. 2015. Immunogenicity, safety, and tolerability of a recombinant measles-virus-based chikungunya vaccine: a randomised, double-blind, placebo-controlled, active-comparator, first-in-man trial. *The Lancet Infect. Dis.,* 15**,** 519-527.

REWAR, S. & MIRDHA, D. 2014. Transmission of ebola virus disease: an overview. *Ann Glob Health,* 80**,** 444-51.

RUDOLPH, K. E., LESSLER, J., MOLONEY, R. M., KMUSH, B. & CUMMINGS, D. A. 2014. Incubation periods of mosquito-borne viral infections: a systematic review. *Am J Trop Med Hyg,* 90**,** 882-91.

RYBICKI, E. P. 2009. Plant-produced vaccines: promise and reality. *Drug Discov Today,* 14**,** 16-24.

RYBICKI, E. P. 2017. Plant-made vaccines and reagents for the One Health initiative. *Hum Vaccin Immunother,* 13**,** 2912-2917.

SAINSBURY, F., THUENEMANN, E. C. & LOMONOSSOFF, G. P. 2009. pEAQ: versatile expression vectors for easy and quick transient expression of heterologous proteins in plants. *Plant Biotechnol J,* 7**,** 682-93.

SCHILLBERG, S., RAVEN, N., SPIEGEL, H., RASCHE, S. & BUNTRU, M. 2019. Critical Analysis of the Commercial Potential of Plants for the Production of Recombinant Proteins. *Front Plant Sci,* 10**,** 720.

SCHILLBERG, S., ZIMMERMANN, S., VOSS, A. & FISCHER, R. 1999. Apoplastic and cytosolic expression of full-size antibodies and antibody fragments in Nicotiana tabacum. *Transgenic Res,* 8**,** 255-63.

SCHOLTHOF, H. B., SCHOLTHOF, K. B. & JACKSON, A. O. 1996. Plant virus gene vectors for transient expression of foreign proteins in plants. *Annu Rev Phytopathol,* 34**,** 299-323.

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

SEHR, P., ZUMBACH, K. & PAWLITA, M. 2001. A generic capture ELISA for recombinant proteins fused to glutathione S-transferase: validation for HPV serology. *J Immunol Methods,* 253**,** 153-162.

SELVARAJAH, S., SEXTON, N. R., KAHLE, K. M., FONG, R. H., MATTIA, K. A., GARDNER, J., LU, K., LISS, N. M., SALVADOR, B., TUCKER, D. F., BARNES, T., MABILA, M., ZHOU, X., ROSSINI, G., RUCKER, J. B., SANDERS, D. A., SUHRBIER, A., SAMBRI, V., MICHAULT, A., MUENCH, M. O., DORANZ, B. J. & SIMMONS, G. 2013. A neutralizing monoclonal antibody targeting the acid-sensitive region in chikungunya virus E2 protects from disease. *PLoS Negl Trop Dis,* 7**,** e2423.

SHENG, Y., MANCINO, V. & BIRREN, B. 1995. Transformation of Escherichia coli with large DNA molecules by electroporation. *Nucleic Acids Res,* 23**,** 1990-6.

SILVA, L. A. & DERMODY, T. S. 2017. Chikungunya virus: epidemiology, replication, disease mechanisms, and prospective intervention strategies. *J Clin Invest,* 127**,** 737-749.

STRATFORD, R., DOUCE, G., ZHANG-BARBER, L., FAIRWEATHER, N., ESKOLA, J. & DOUGAN, G. 2000. Influence of codon usage on the immunogenicity of a DNA vaccine against tetanus. *Vaccine,* 19**,** 810-815.

SUN, Q., DU, X. & CHENG, W. 2019. Structures Unveil the Invasion Mechanism of Chikungunya Virus. *Trends Microbiol,* 27**,** 656-658.

TAKEDA, A., SUGIYAMA, K., NAGANO, H., MORI, M., KAIDO, M., MISE, K., TSUDA, S. & OKUNO, T. 2002. Identification of a novel RNA silencing suppressor, NSs protein of Tomato spotted wilt virus. *FEBS Letters,* 532**,** 75-79.

THARMARAJAH, K., MAHALINGAM, S. & ZAID, A. 2017. Chikungunya: vaccines and therapeutics. *F1000Res,* 6**,** 2114.

THIBOUTOT, M. M., KANNAN, S., KAWALEKAR, O. U., SHEDLOCK, D. J., KHAN, A. S., SARANGAN, G., SRIKANTH, P., WEINER, D. B. & MUTHUMANI, K. 2010. Chikungunya: a potentially emerging epidemic? *PLoS Negl Trop Dis,* 4**,** e623.

THOMAS, D. R. & WALMSLEY, A. M. 2014. Improved expression of recombinant plant-made hEGF. *Plant Cell Rep,* 33**,** 1801-14.

TREGONING, J. S., NIXON, P., KURODA, H., SVAB, Z., CLARE, S., BOWE, F., FAIRWEATHER, N., YTTERBERG, J., VAN WIJK, K. J., DOUGAN, G. & MALIGA, P. 2003. Expression of tetanus toxin Fragment C in tobacco chloroplasts. *Nucleic Acids Res,* 31**,** 1174-9.

TSCHOFEN, M., KNOPP, D., HOOD, E. & STOGER, E. 2016. Plant Molecular Farming: Much More than Medicines. *Annu Rev Anal Chem (Palo Alto Calif),* 9**,** 271-94.

TSETSARKIN, K. A., VANLANDINGHAM, D. L., MCGEE, C. E. & HIGGS, S. 2007. A Single Mutation in Chikungunya Virus Affects Vector Specificity and Epidemic Potential. *PLOS Pathogens,* 3**,** e201.

TURNER, T. L., KOPP, B. T., PAUL, G., LANDGRAVE, L. C., HAYES, D., JR. & THOMPSON, R. 2014. Respiratory syncytial virus: current and emerging treatment options. *Clinicoecon Outcomes Res,* 6**,** 217-25.

UEMURA, K., SUZUKI, Y., SHIKANAI, T., WADANO, A., JENSEN, R. G., CHMARA, W. & YOKOTA, A. 1996. A Rapid and Sensitive Method for Determination of Relative Specificity of RuBisCO from Various Species by Anion-Exchange Chromatography. *Plant Cell Physiol,* 37**,** 325-331.

VERBEEK, M. 2019. *The Expression of the Shuni Virus Nucleocapsid Protein in Nicotiana benthamiana for Use as a Diagnostic Reagent* Master od Science Dissertation, University of Cape Town.

VERMA, A., CHANDELE, A., NAYAK, K., KAJA, M. K., ARULANDU, A., LODHA, R. & RAY, P. 2016. High yield expression and purification of Chikungunya virus E2 recombinant protein and its evaluation for serodiagnosis. *J Virol Methods,* 235**,** 73-79.

VOLK, S. M., CHEN, R., TSETSARKIN, K. A., ADAMS, A. P., GARCIA, T. I., SALL, A. A., NASAR, F., SCHUH, A. J., HOLMES, E. C., HIGGS, S., MAHARAJ, P. D., BRAULT, A. C. & WEAVER, S. C. 2010. Genome-Scale Phylogenetic Analyses of Chikungunya Virus Reveal Independent Emergences of Recent Epidemics and Various Evolutionary Rates. *J. Virol.,* 84**,** 6497-6504.

WAUGH, D. S. 2005. Making the most of affinity tags. *Trends Biotechnol,* 23**,** 316-20.

WEAVER, S. C. & LECUIT, M. 2015. Chikungunya virus and the global spread of a mosquito-borne disease. *N Engl J Med,* 372**,** 1231-9.

WEBER, C., BÜCHNER, S. M., SCHNIERLE, B. S. & WILLIAMS, M. 2015. A Small Antigenic Determinant of the Chikungunya Virus E2 Protein Is Sufficient to Induce Neutralizing Antibodies which Are Partially Protective in Mice. *PLoS Negl Trop Dis,* 9**,** e0003684.

WIMALASIRI-YAPA, B., STASSEN, L., HUANG, X., HAFNER, L. M., HU, W., DEVINE, G. J., YAKOB, L., JANSEN, C. C., FADDY, H. M., VIENNET, E. & FRENTIU, F. D. 2019. Chikungunya virus in Asia - Pacific: a systematic review. *Emerg Microbes Infect,* 8**,** 70-79.

ZENTGRAF, H., FREY, M., SCHWINN, S., TESSMER, C., WILLEMANN, B., SAMSTAG, Y. & VELHAGEN, I. 1995. Detection of histidine-tagged fusion proteins by using a high-specific mouse monoclonal anti-histidine tag antibody. *Nucleic Acids Res,* 23**,** 3347-8.

ZHENG, C.-F., SIMCOX, T., XU, L. & VAILLANCOURT, P. 1997. A new expression vector for high level protein production, one step purification and direct isotopic labeling of calmodulin-binding peptide fusion proteins. *Gene,* 186**,** 55-60.

Appendix

ACAAAGGATAATTTCAATGTTTACAAGGCTACCAGACCATACCTCGCACACTGCCCAGATTGCGGAGA
AGGACACAGTTGCCACTCTCCAGTTGCTCTTGAAAGGATTAGAAATGAGGCTACTGATGGTACACTTA
AGATTCAAGTTTCTTTGCAGATTGGAATTAAGACTGATGATTCACATGATTGGACAAAACTTAGGTAC
ATGGATAATCATATGCCAGCTGATGCTGAAAGGGCTGGTTTGTTTGTTAGAACTTCTGCTCCTTGTAC
AATTACTGGTACAATGGGACATTTTATTCTTGCTAGGTGCCCAAAGGGAGAGACTTTGACAGTTGGTT
TTACTGATTCAAGAAAAATTTCTCATTCATGTACACATCCTTTTCATCACGATCCTCCTGTTATTGGTAG
GGAAAAGTTTCATTCTAGACCACAACATGGAAAAGAGCTTCCTTGCTCAACATATGTTCAGAATACTG
CTGCTACAGCTGAAGAGATTGAAGTTCATATGCCTCCAGATACTCCTGATAGGACATTGATGTCTCAA
CAGTCTGGTAACGTTAAGATCACTGTTAATGGTCAAACAGTTAGATACAAATGTAATTGCGGTGGATC
TAATGAGGGACTTACTACAACTGATAAGGTTATTAATAACTGTAAAGTTGATCAATGCCATGCTGCTG
TTACTAACCATAAGAAGTGGCAGTACAACTCACCTCTTGTTCCAAGGAATGCTGAGTTGGGAGATAG
AAAGGGAAAAATTCATATCCCTTTCCCATTGGCTAATGTTACTTGTAGGGTTCCTAAGGCTAGAAATC
CAACTGTTACATACGGAAAGAATCAAGTTATTATGCTTTTGTACCCTGATCATCCAACACTTTTGTCTTA
CAGAAACATGGGAGAAGAGCCAAATTACCAGGAAGAGTGGGTTACTTATAAGAAAGAAGTTGTTCTT
ACTGTTCCTACAGAAGGTTTGGAGGTTACATGGGGAAATAACGAACCTTACAAGTATTGGCCACAGC
TTTCTGCTAATGGTACTGCTCATGGACATCCACATGAAATTATTCTTTACTACTACGAGTTGTACCCTAC
TATGACAGTTGTTGTTGTTTCTGTTGCTTCTTTTATTCTTTTGTCAATGGTTGGTATGGCTGTTGGAATG
TGTATGTGCGCTAGAAGGAGATGCATTACACCTTATGAGTTGACTCCAGGTGCTACAGTTCCTTTTTT
ATTGTCACTCATCTGTTGTTTTAGAACCGCAAAGGCT

Figure A1: Consensus sequence of full-length CHIKV E2. Three South African CHIKV isolates (GenBank: HM045792.1, Vereeniging, Gauteng province; HM045795.1, SAH2123 strain; HM045805.1, CHIKV AR 18211 1976) were used to generate a consensus sequence. Highlighted sequence represents the transmembrane domain that is removed to generate the CHIKV E2ΔTM variant.